Chapter 1

Foundations

Introduction

Welcome to Book A *Number theory*. This first chapter lays a firm foundation for the number theory components of the module. It examines in detail the fundamental ideas that form the basis from which to study the properties of numbers. We introduce, and make precise, the key notion of *divisibility*. For this we use a result concerning the division of one integer by another, which will be familiar to you. This result is known as the *Division Algorithm*. Its consequences, both practical and theoretical, make it a cornerstone of number theory. Through it we embark on a study of the properties of divisibility, and introduce the concept of the *highest common factor* – a topic that might already be familiar to you and one that is going to figure prominently in this part of the module. The chapter also reviews a powerful method of proof in number theory: *proof by mathematical induction*. You may well have met this idea before, but it is of such importance to us that we give a complete treatment here.

Before we get down to any theoretical ideas we need some families of numbers to which we might apply our techniques, and that is where we begin, after first giving a brief preview in Section 1 of some of the tasks that will confront us in the module.

Chapter guide

Section 1 gives a brief introduction to the kinds of problem that arise in number theory, which should require little study time. You may have seen aspects of the material in Sections 2 and 5 in other modules that you have studied, but here we take a more rigorous approach. The main focus of this chapter is on mathematical induction, covered in Section 3, and on the notion of divisibility, covered in Section 4. Section 5 introduces linear Diophantine equations, a topic that you will encounter again in Book C.

1 What is number theory?

> The elementary theory of numbers should be one of the very best subjects for early mathematical instruction. It demands very little previous knowledge; its subject matter is tangible and familiar; the processes of reasoning it employs are simple, general and few; and it is unique among the mathematical sciences in its appeal to natural human curiosity. A month's intelligent instruction in the theory of numbers ought to be twice as instructive, twice as useful, and ten times more entertaining as the same amount of 'calculus for engineers'.
>
> G.H. Hardy, *Bulletin of the AMS* (1929)

There is an irresistible fascination in searching for numbers with specified properties. Nearly every century, as far back as the history of mathematics can be traced, has witnessed new and exciting discoveries concerning properties of numbers. Many of the greatest mathematicians, despite having their major interests elsewhere, have at some time in their careers been drawn into problems of number theory and have contributed to the body of knowledge. So what is the appeal of this subject both for professional mathematicians and for thousands of amateurs?

Consider the following problems.

1. Find all the factors of 4 294 967 297.
2. A right-angled triangle has the property that all three of its sides have length equal to a whole number of units. If one of the sides is 24, find seven pairs of values for the other two sides.
3. Show that

   $$1 \times 2 \times 3 \times 4 \times \cdots \times (n-1) + 1$$

   is always divisible by n if n is prime, but is never divisible by n if n is composite.
4. Can every even integer from 4 onwards be expressed as the sum of two primes?
5. For which integers n is $2^n - 1$ a prime?

All these problems will confront us at some stage in this part of the module. Each one illustrates the most attractive feature of number theory: the problems can be understood by beginners of the subject. Perhaps you are not completely familiar with some of the terms (which will be explained later) such as prime, composite and factor, but you should have a feeling for what is involved in these problems. Indeed, you could well take pencil and paper and start exploring some of them; there is no substantial body of knowledge needed as a prerequisite to becoming involved in number theory.

Yet while there is no difficulty posing these problems in a readily intelligible form, the varying degrees of difficulty involved in solving them highlights the most intriguing feature of number theory.

Because of the size of the number involved, Problem 1 is quite tricky. However, any reader who is familiar with computers might quickly solve this problem. As it happens, the smallest positive factor of the given number is 641 – a fact which, in the seventeenth century, eluded one of the all-time great mathematicians, Fermat (1601–1665), who had a particular interest in this very problem.

Problem 2 asks for solutions in whole numbers of the famous equation of Pythagoras relating the edge lengths of a right-angled triangle, $x^2 + y^2 = z^2$ (illustrated in Figure 1.1). This is an example of a Diophantine equation, named after the Greek mathematician Diophantus of the early Christian era. We will meet other instances of Diophantine equations in this module, which we will show how to solve. Many books have been written about Diophantine equations. What emerges is a lack of a general theory for solving them; each Diophantine equation appears to be a problem in its own right, requiring its own method of solution. In the meantime, if you are familiar with the 'smallest' solution in integers of the Pythagorean equation, namely $3^2 + 4^2 = 5^2$, then you may have spotted two of the solutions required by Problem 2 by scaling, namely $18^2 + 24^2 = 30^2$ and $24^2 + 32^2 = 40^2$. But what about the other five solutions? You might discover them by patient trial and error – but is there a more systematic approach? We will explore this further at the start of Book C, Chapter 12.

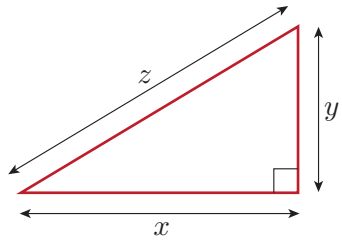


Figure 1.1 An illustration of the Pythagorean equation $x^2 + y^2 = z^2$

Problems 3, 4 and 5 are classics of number theory. Because the values involved in Problem 3 get very large, very quickly, it is not an easy problem to explore in depth by looking at special cases. But it turns out to be true, and we can give (and will do so in Chapter 4) a general proof of this result. Problem 4 is easier to explore and, on finding it very straightforward to write each even integer up to 200 (say) as a sum of two primes, you would probably be tempted to conclude that it also is a true result. But, perhaps surprisingly, nobody has ever managed to prove it; to this day it remains an unsolved problem of mathematics, despite assault by many great mathematicians. Problem 5 has its origin at the heart of another famous problem, which you will meet in Book C, Chapter 9. By the year 1951, only 12 numbers n were known for which $2^n - 1$ is prime; then computers were brought to bear on the problem, with the result that by the year 2014 forty-eight such numbers were known. The largest one, $n = 57\,885\,161$, led to the number $2^{57\,885\,161} - 1$ being the largest known prime at that time. And yet, despite the scarcity of solutions to Problem 5, most mathematicians still believe that there are infinitely many solutions to this problem – though proof of this seems, at the moment, hopelessly beyond reach.

Number theory is a subject that has developed over a long period of time and, as the previous paragraph suggests, remains very active today. We have already mentioned a few of the great mathematicians who contributed to the development of the subject; there are many, many more. In the *History Reader* for this topic, we give a flavour of the history of the subject. The *History Reader* is not an assessed part of the module and by no means gives a systematic account of the history of number theory. Its inclusion will, we hope, enliven the theoretical side of the material and should also reveal the stumbling way in which progress has been made.

The nature of number theory has changed dramatically in recent years, due to the advent of the computer. Computations that were nigh impossible just a few years ago can now be managed with ease. Consequently, exploration of alleged results and searches for numbers with prescribed properties are readily attacked with the assistance of a machine. Although the material of this module presents many challenges for those with access to (and enthusiasm for) a computer, we have written the module in such a way as to avoid the heavy computational side of the subject. We will be attacking problems in the way that they have been tackled historically, using pencil and paper. There are, however, many problems in the module involving 'arithmetic' that could be simplified with the help of a calculator, and although it is not essential, it is advisable that you have one.

Number theory is all about problem-solving. Nobody becomes competent at calculus by reading about it; it is essential to practise differentiation and integration on a large number of examples. The same is true of number theory; you cannot get to grips with the subject without solving lots of problems yourself. To this end we have included a good stock of problems in all the chapters, and while studying you should always have pencil and paper at hand. The Worked Exercises in the chapters are for you to read, but the Exercises are for you to *do*. Do not spend forever on an exercise that looks like defeating you. If you get stuck, you should refer to the solution – but do not give in too easily!

2 Numbers from patterns

In this section we are concerned with certain families of numbers that can be derived from patterns, namely the polygonal numbers and the pyramidal numbers.

2.1 Polygonal numbers

Figure 2.1 illustrates the first five of the so-called **triangular numbers**, T_n. The arrangement of T_4 will be familiar to anyone who indulges in ten-pin bowling and T_5 is, more or less, the arrangement of the pack of 15 red balls at the start of a game of snooker. It is evident that the triangular numbers are the sums of consecutive integers starting from 1:

$$T_1 = 1, \quad T_2 = 1 + 2, \quad T_3 = 1 + 2 + 3, \quad \ldots, \quad T_n = 1 + 2 + 3 + \cdots + n.$$

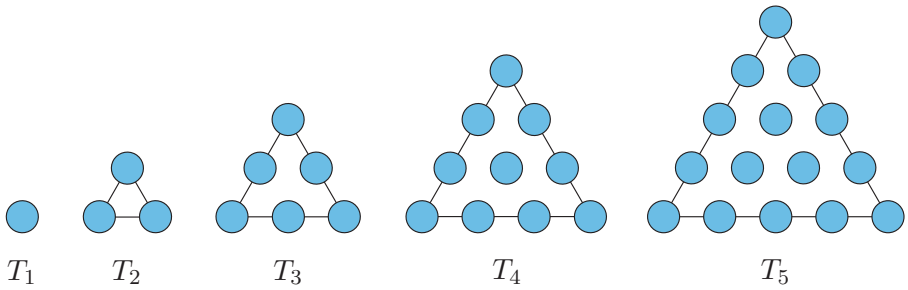


Figure 2.1 Triangular numbers

We can obtain a simple formula for T_n by writing the sum both forwards and backwards:

$$\begin{array}{rcccccccccc} T_n & = & 1 & + & 2 & + & 3 & + & \cdots & + & n \\ T_n & = & n & + & (n-1) & + & (n-2) & + & \cdots & + & 1 \end{array}$$

and adding the two sums by pairing each term in the first line with the term directly below it. Each corresponding pair has a sum of $(n + 1)$, so we obtain

$$2T_n = (n + 1) + (n + 1) + \cdots + (n + 1) = n(n + 1),$$

giving

$$T_n = 1 + 2 + 3 + \cdots + n = \tfrac{1}{2}n(n + 1).$$

This same argument can be used for any arithmetic series. An **arithmetic series** is the sum of a finite sequence of terms in which the difference between consecutive terms is constant. So the arithmetic series having n terms with first term a and common difference d is

$$a + (a + d) + (a + 2d) + \cdots + (a + (n - 1)d).$$

The sequence (rather than the sum) of terms

$$a,\ a + d,\ a + 2d,\ \ldots$$

is called an **arithmetic progression** (or arithmetic sequence). For example, the sequence of triangular numbers is an arithmetic progression with $a = d = 1$.

A formula for the arithmetic series can be obtained by reversing the sum and adding corresponding terms, as was done for the triangular numbers: each pair of corresponding terms sums to $2a + (n-1)d$, and so we have

$$a + (a+d) + (a+2d) + \cdots + (a+(n-1)d) = \tfrac{1}{2}n(2a+(n-1)d).$$

This formula can be rewritten as

$$n(a + \tfrac{1}{2}(n-1)d),$$

illustrating that there are n terms of average $a + \frac{1}{2}(n-1)d$.

Exercise 2.1

Determine the following sums.

(a) $3 + 10 + 17 + 24 + \cdots + 143$

(b) $100 + 98 + 96 + \cdots + 50$

Figure 2.2 shows that two consecutive triangular numbers add up to a **square number**, that is, a number of the form n^2 for some integer $n \geq 1$. In Figure 2.2, each square of dots is made up of a triangle of white dots and the next larger triangle of blue dots. Interpreting this observation in terms of triangular and square numbers, it would appear that

$$T_{n-1} + T_n = n^2, \quad n \geq 1.$$

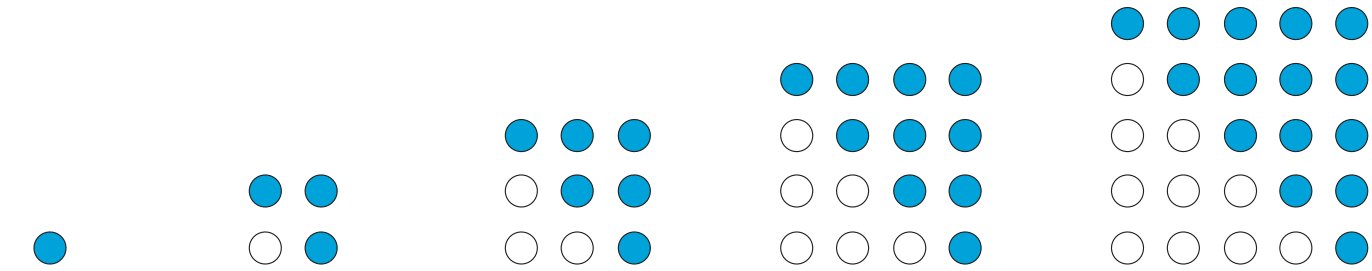


Figure 2.2 Adding consecutive triangular numbers

By convention we take $T_0 = 0$, which is the value given by the formula $\frac{1}{2}n(n+1)$. A formal proof that $T_{n-1} + T_n$ is equal to n^2 is readily obtained using the formula for T_n:

$$T_{n-1} + T_n = \tfrac{1}{2}(n-1)n + \tfrac{1}{2}n(n+1) = \tfrac{1}{2}n(n-1+n+1) = n^2.$$

In future we may refer to the square numbers merely as the squares. Square numbers are sometimes called perfect squares. Furthermore, we may shorten the phrase ‘is a triangular number’ to ‘is triangular’.

Exercise 2.2

(a) You have seen that the triangular numbers are precisely the numbers of the form $\frac{1}{2}n(n+1)$. Show that if k is a triangular number then so is $9k+1$.

(b) Show that an integer k is triangular if, and only if, $8k+1$ is square.

Note that since this is an 'if, and only if' statement, its proof will need to have two parts. You will need to show that if k is triangular then $8k+1$ is square and, conversely, that if $8k+1$ is square then k is triangular.

Hint: for the latter, note that the square number $8k+1$ is odd, and so it has to be the square of an odd number, that is, $8k+1=(2s+1)^2$, for some integer s.

Let us reconsider the hint given in Exercise 2.2(b). It advised you to make use of the fact that you were dealing with an odd number. The odd integers

$$\ldots, -3,\ -1,\ 1,\ 3,\ 5,\ \ldots$$

are those that can be written as $2m+1$, for some integer m. The even integers

$$\ldots, -4,\ -2,\ 0,\ 2,\ 4,\ \ldots$$

are those that can be written as $2m$, for some integer m. Hence all integers can be written as a multiple of 2 plus a remainder r that is either 0 or 1. Expressing an integer as a multiple of a positive integer k plus a remainder r (where $0 \leq r < k$) can be a useful approach, as you will see in Section 4.

One instance of the result of Exercise 2.2(b) is illustrated in Figure 2.3. A 9×9 square of dots is broken into eight copies of $T_4 = 10$ with one dot left over in the middle.

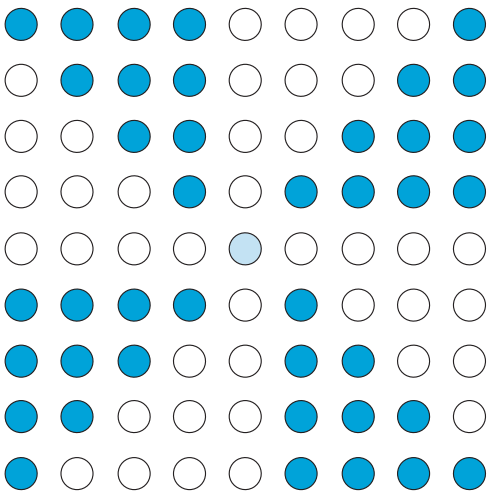


Figure 2.3 An illustration of $8T_4 + 1 = 9^2$

While the triangular numbers are sums of consecutive integers,

$$1,\ 1+2,\ 1+2+3,\ \ldots,$$

Figure 2.4 illustrates that the squares are sums of consecutive odd integers,

$$1,\ 1+3,\ 1+3+5,\ \ldots;$$

the next odd number is added to the current square as dots along the bottom and right-hand side to produce the next square.

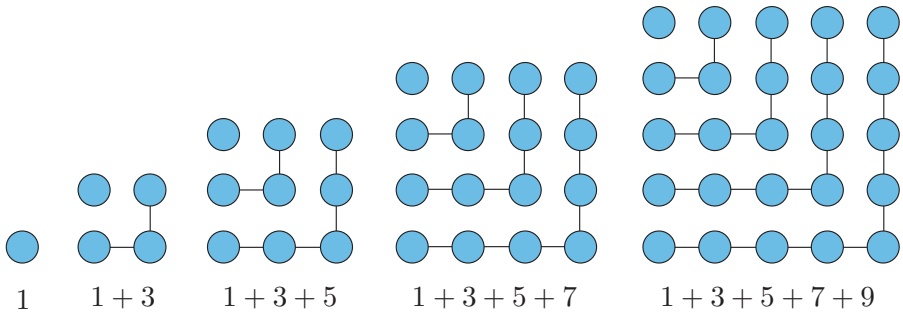


Figure 2.4 Squares as sums of odd integers

The general formula suggested by this sequence of diagrams is

$$1+3+5+\cdots+(2n-1)=n^2, \quad n \geq 1.$$

We will use 'formulas' rather than 'formulae' as the plural of 'formula'.

A proof of this result is immediate from the formula for an arithmetic series but we will leave it aside for the time being. Formulas like this, in which results are asserted for each positive integer n, are often tailor-made for proof by the method of mathematical induction, which we deal with in Section 3.

Higher polygonal numbers

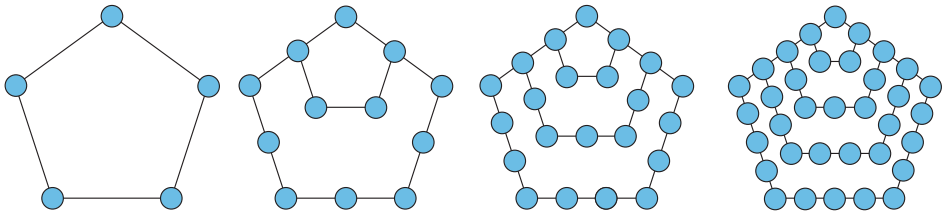
We have seen that triangular numbers are sums of consecutive integers, and squares are sums of consecutive odd integers. The pattern can be continued: **pentagonal numbers** are the sums obtained from going up in threes, namely

$$1,\ 1+4,\ 1+4+7,\ 1+4+7+10,\ \ldots.$$

Figure 2.5 illustrates the start of the sequence of pentagonal numbers, and Figure 2.6 goes one stage further and shows how the **hexagonal numbers**,

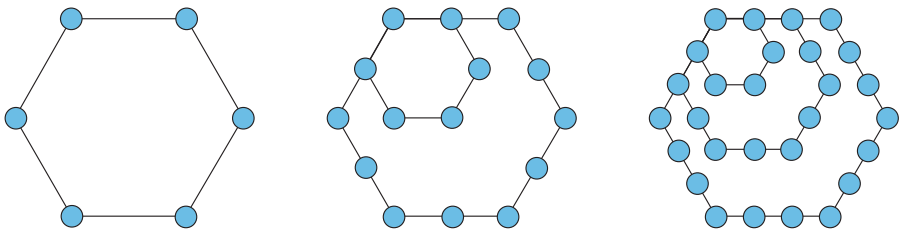
$$1,\ 1+5,\ 1+5+9,\ 1+5+9+13,\ \ldots,$$

arise from arrangements of dots forming hexagons.



$1+4$ $1+4+7$ $1+4+7+10$ $1+4+7+10+13$

Figure 2.5 Pentagonal numbers



$1+5$ $1+5+9$ $1+5+9+13$

Figure 2.6 Hexagonal numbers

Table 2.1 shows the beginnings of the sequences of **polygonal numbers**. The first row is the sequence of triangular numbers,

$$1,\ 1+2=3,\ 1+2+3=6,\ 1+2+3+4=10,\ \ldots.$$

We have extended the table beyond the numbers already seen by appealing to the emerging patterns. For example, the **octagonal numbers** form the sequence

$$1,\ 1+7,\ 1+7+13,\ 1+7+13+19,\ \ldots.$$

Table 2.1 Table of polygonal numbers

	Number of sides	Term 1	Term 2	Term 3	Term 4	Term 5	Term 6	Term 7	Term 8	...	Term n
Triangular	3	1	3	6	10	15	21	28	36	...	$\frac{1}{2}n(n+1)$
Square	4	1	4	9	16	25	36	49	64	...	n^2
Pentagonal	5	1	5	12	22	35	51	70	92	...	?
Hexagonal	6	1	6	15	28	45	66	91	120	...	?
Heptagonal	7	1	7	18	34	55	81	112	148	...	?
Octagonal	8	1	8	21	40	65	96	133	176	...	?
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
k-gonal	k	1	k	$3k-3$	?	?	?	?	?	...	$P(k,n)$

In the right-hand column we have listed the discovered formulas for the nth triangular number and for the nth square. The rest are missing, but could be determined from the appropriate arithmetic series. Along the bottom row we have begun listing the terms for the k-gonal numbers. You may be able to spot formulas (in terms of k) for these numbers from the patterns that are present in the table.

Of course, we could readily fill in the right-hand column and the bottom row if we could find a general formula for the nth term in the sequence of k-gonal numbers, which we will call $P(k, n)$. The next exercise invites you to provide this formula. Remember that $P(3, n)$, the nth triangular number, is the sum of n terms of the arithmetic progression with first term 1 and common difference 1; $P(4, n)$, the nth square, is the sum of n terms of the arithmetic progression with first term 1 and common difference 2; and, in general, $P(k, n)$ is the sum of n terms of the arithmetic progression with first term 1 and common difference $k - 2$.

Exercise 2.3

(a) Obtain a formula for $P(k, n)$, the nth k-gonal number, and hence fill in the missing entries in Table 2.1.

(b) The numbers in each of the columns of Table 2.1 appear to be in arithmetic progression. The common difference between two successive terms in each column is equal to the triangular number at the head of the previous column. That is,

$$P(k, n) = P(k - 1, n) + P(3, n - 1).$$

Prove this formula algebraically.

2.2 Pyramidal numbers

The polygonal numbers arise as sums of arithmetic progressions with first term 1. We can take this a stage further: the **pyramidal numbers** are those that arise as sums of finite sequences of polygonal numbers. For instance, from the sequence of triangular numbers,

$$1,\ 3,\ 6,\ 10,\ 15,\ \ldots,$$

we obtain the **triangle-based pyramidal numbers** as follows:

$$1,\ 1 + 3 = 4,\ 1 + 3 + 6 = 10,\ 1 + 3 + 6 + 10 = 20,\ \ldots.$$

Figure 2.7 illustrates how these numbers arise from layers of a triangle-based pyramid.

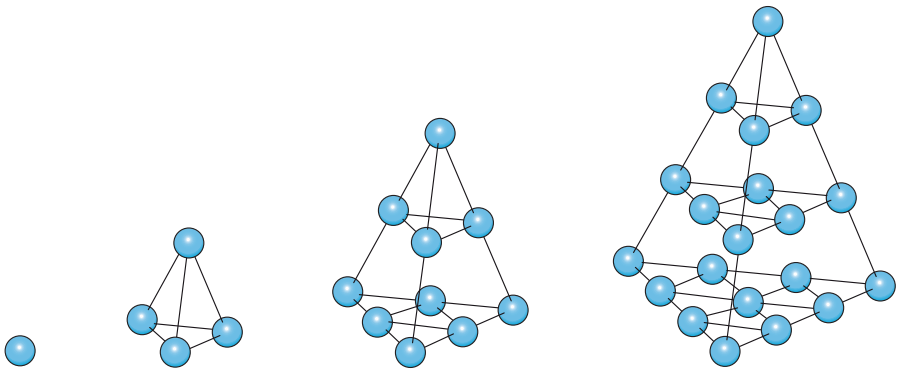


Figure 2.7 Triangle-based pyramidal numbers

Similarly, by adding consecutive terms of the sequence of squares we get the **square-based pyramidal numbers**

$$1,\ 1+4=5,\ 1+4+9=14,\ 1+4+9+16=30,\ \ldots,$$

illustrated in Figure 2.8.

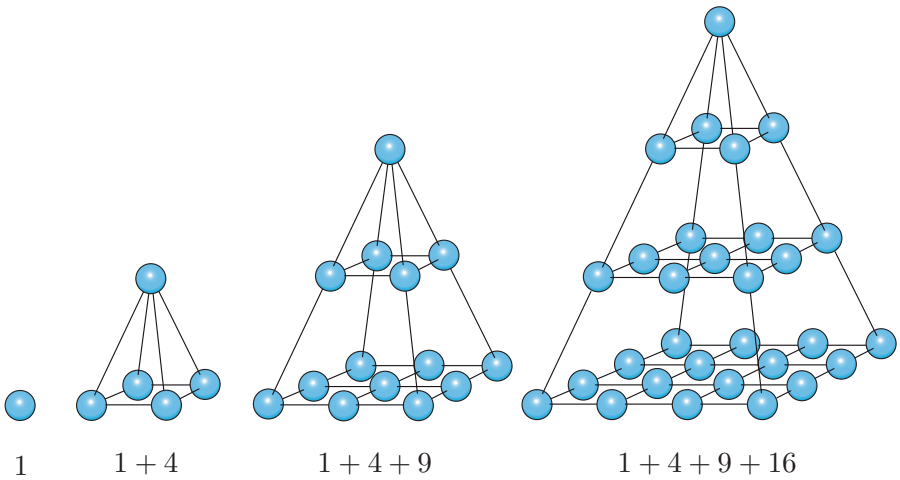


Figure 2.8 Square-based pyramidal numbers

The pyramidal numbers are readily generalised. For example, the **heptagonal-based pyramidal numbers** are generated as sums of heptagonal numbers,

$$1,\ 1+7=8,\ 1+7+18=26,\ 1+7+18+34=60,\ \ldots.$$

Table 2.2 gives some of the pyramidal numbers. We have started, but not completed, the row for the general k-gon and the column for the general nth term. The formulas that occur in the latter are not so easily proved; we return to them when we have discussed proof by mathematical induction.

Table 2.2 Table of pyramidal numbers

Base	Term 1	Term 2	Term 3	Term 4	Term 5	Term 6	Term 7	Term 8	Term n
Triangle	1	4	10	20	35	56	84	120	$\frac{1}{6}n(n+1)(n+2)$
Square	1	5	14	30	55	91	140	204	$\frac{1}{6}n(n+1)(2n+1)$
Pentagon	1	6	18	40	75	126	196	288	?
Hexagon	1	7	22	50	95	161	252	372	?
Heptagon	1	8	26	60	115	196	308	456	?
Octagon	1	9	30	70	135	231	364	540	?
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
k-gon	1	$k+1$	$4k-2$	?	?	?	?	?	$Q(k,n)$

There may not appear to be anything complicated about these numbers derived from geometric configurations. However, they give rise to many interesting and difficult problems. For example, the first two triangular pyramidal numbers, 1 and 4, are both square. Are there any other triangular pyramidal numbers that are square? It turns out that there is just one more:

$$140^2 = 1 + 3 + 6 + 10 + \cdots + 1176.$$

However, the proof that this is the only one is difficult.

The table of polygonal numbers (Table 2.1) reveals two numbers that are both square and triangular, namely 1 and 36. Are there any more? As it happens the next one is 1225, followed by 41 616 and then 1 413 721. Exercise 2.2(b) gives us a clue to one way of attacking this problem. The number x^2 will be triangular as well as square if, and only if, $8x^2+1$ is a square. In other words, triangular squares will arise from integer solutions of the equation $8x^2 + 1 = y^2$. If x and y are integers satisfying this equation then x^2 is a triangular square. In fact there are infinitely many solutions of this equation, and hence infinitely many triangular squares. You will see how to find them all later in the module, in Book C, Chapter 12.

Exercise 2.4

(a) In Exercise 2.3(b) we discovered the formula

$$P(k,n) = P(k-1,n) + P(3,n-1).$$

Examination of Table 2.2 shows that the same relationship appears to hold there, namely that each entry in the body of the table is equal to the one above it added to the triangle-based pyramidal number at the top of the previous column. That is,

$$Q(k,n) = Q(k-1,n) + Q(3,n-1).$$

Assuming this formula is true, show that

$$Q(5, n) = \tfrac{1}{2}n^2(n + 1)$$

and determine $Q(6, n)$.

(b) If either $2n^2 + 1$ or $2n^2 - 1$ is a square, say m^2, show that $(nm)^2$ is a triangular number. Hence find three examples of squares that are also triangular.

(c) Show that no triangular number can have 2, 4, 7 or 9 as its last digit.

Hint: you might find a case-by-case analysis of final digits occurring in numbers of the form $k(k + 1)$ helpful.

3 Mathematical induction

Although you may have encountered mathematical induction before now, you may not be familiar with the more formal approach taken to it in this section.

3.1 Notation

Before progressing further we introduce some standard notation to be used throughout this module. In number theory we are primarily interested in the positive integers, also called the *natural numbers*.

A collection of distinct objects (such as numbers) is known as a **set**. The set of all **integers**, positive, negative or zero, is denoted by $\mathbb{Z}$:

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$

and the set of **natural numbers** is denoted by $\mathbb{N}$:

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

We have described each of these sets by (partially) listing its elements within 'curly brackets', which are known more formally as **braces**. We often describe a set by giving a property that characterises its elements. For example, the set of all integers that are multiples of 3, and which is denoted by $3\mathbb{Z}$, can be described either by

$$3\mathbb{Z} = \{\ldots, -6, -3, 0, 3, 6, \ldots\}$$

or by

$$3\mathbb{Z} = \{3n : n \in \mathbb{Z}\}.$$

We will also have occasion to refer to the set $\mathbb{Q}$ of **rational numbers** and the set $\mathbb{R}$ of **real numbers**. The rational numbers are the numbers $\frac{a}{b}$, where $a \in \mathbb{Z}$ and $b \in \mathbb{N}$, and the real numbers can be thought of as values representing all the points along an infinite number line.

Finally, we will make extensive use of the **modulus** or *absolute value* function, which is defined for any element n of $\mathbb{Z}$ (or $\mathbb{Q}$ or $\mathbb{R}$) by

$$|n| = \begin{cases} n, & \text{if } n \geq 0, \\ -n, & \text{if } n < 0. \end{cases}$$

Exercise 3.1

(a) Which set of numbers is described by each of the following?

(i) $\{n \in \mathbb{Z} : |n| > 20\}$

(ii) $\{n : n = 2m^2, \text{ for some } m \in \mathbb{N}\}$

(b) Describe each of the following sets in the notation used in part (a).

(i) The set of all odd natural numbers.

(ii) The set of all integers lying between -100 and 100 inclusive.

3.2 Mathematical induction

One property of $\mathbb{N}$ that is not shared by some other number sets such as $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$ is worth recording. It may seem a rather obvious property, but its presence is crucial to establishing less apparent properties that we are going to need.

The Well-Ordering Principle for $\mathbb{N}$

Every non-empty subset of $\mathbb{N}$ has a least member. In other words, if S is a non-empty subset of $\mathbb{N}$ then there exists $b \in S$ such that $b \leq n$ for all $n \in S$.

In future we will refer to this just as the Well-Ordering Principle. It says that any collection of natural numbers we care to describe, as long as it has some elements in it, must have a least member. Notice that the set of positive real numbers does not have this property. For example, the set of positive real numbers itself has no least element since, if x is *any* positive real, then

$$0 < \tfrac{1}{2}x < x,$$

showing that $\frac{1}{2}x$ is a positive real number smaller than x. Hence x cannot be the smallest positive real number.

From the Well-Ordering Principle we can quickly deduce the result that is the backbone of the method of proof by *mathematical induction*.

Theorem 3.1 *Principle of Induction*

If S is a set of natural numbers with the following two properties:

(a) 1 is a member of S

(b) if $k \in S$, then the next integer $k+1 \in S$

then $S = \mathbb{N}$.

Proof We give a proof by contradiction. We assume that the theorem is not true and that there is a set S of natural numbers satisfying (a) and (b) that is not the whole of $\mathbb{N}$. We show that this assumption leads to a contradiction and therefore the theorem must be true.

Let A be the set of all natural numbers that do not belong to S. By the assumption that S is not the whole of $\mathbb{N}$, we know that A is non-empty. Hence, by the Well-Ordering Principle, A must contain a least member. Let this least member of A be a.

By property (a) the integer 1 belongs to S, and so 1 is not in A. Therefore $a > 1$. Now consider the integer $a-1$, which must be positive as $a > 1$. Furthermore, as $a - 1 < a$ and a is the least member of A, it follows that $a - 1$ does not belong to A. Therefore $a - 1$ is a member of S.

But property (b) now tells us that the integer following $a - 1$, namely a itself, belongs to S. This contradicts the fact that a belongs to A. This contradiction means that the one assumption we made, namely that A is non-empty, must be false. So A is empty and hence S is the set of all natural numbers. ■

Worked Exercise 3.2

Prove that

$$1 + 3 + 5 + \cdots + (2n-1) = n^2, \quad \text{for all natural numbers } n.$$

We discovered this formula in Section 2 but did not give a formal proof. We rectify that omission now.

Solution

With an eye on the Principle of Induction, let S be the set of natural numbers n for which the formula holds. Our goal is to show that $S = \mathbb{N}$.

Putting $n = 1$ in the formula, we observe that the left-hand side has only one term so that the formula reduces to $1 = 1^2$, which is true. So $1 \in S$.

It remains to prove property (b). To that end, suppose $k \in S$, where k is some natural number. Since $k \in S$ we have that

$$1 + 3 + 5 + \cdots + (2k-1) = k^2.$$

To deduce that $k+1 \in S$ we must show that the given formula is true for the case $n = k+1$. Working on the left-hand side of the formula for $n = k+1$:

$$\begin{aligned} 1+3+5+\cdots+(2k-1)+(2k+1) & \\ &= k^2+(2k+1), \quad \text{from the assumption that } k \in S, \\ &= (k+1)^2, \quad \text{which is the required right-hand side.} \end{aligned}$$

This reasoning shows that if $k \in S$ then $k+1 \in S$, which confirms property (b).

Hence $S = \mathbb{N}$, and the formula is true for all natural numbers.

We presented Worked Exercise 3.2 by applying the Principle of Induction to the set of natural numbers for which the proposition was true. In future we will present such arguments less formally and will refer to them as proof by **mathematical induction**. The propositions for which we attempt such proofs are ones that are given in terms of a general natural number n, and that we hope to be true for all $n \in \mathbb{N}$.

Let $P(n)$ be a proposition whose truth we wish to prove by mathematical induction. Such a proposition can take various forms. It might be a formula as in Worked Exercise 3.2, an inequality such as $2^n > n^3$, which we will discuss in Worked Exercise 3.4, or a statement such as 'n can be written as a sum of distinct powers of 2', which you will meet in Worked Exercise 3.6. To prove the truth of $P(n)$, let S be the set of natural numbers n for which $P(n)$ is true. If we know that $P(1)$ is true then $1 \in S$. Further, if we can show that whenever $P(k)$ is true it follows that $P(k+1)$ is true, then $k \in S$ implies $k+1 \in S$. Hence, by the Principle of Induction, $S = \mathbb{N}$, and $P(n)$ is true for all natural numbers.

We state this formally as follows.

Principle of Mathematical Induction

Let $P(n)$ be a proposition depending on a natural number n. If:

(a) $P(1)$ is true

(b) for any integer $k \geq 1$, if $P(k)$ is true then $P(k+1)$ is true

then $P(n)$ is true for all $n \in \mathbb{N}$.

Step (a), showing that the proposition is true for the first value, is called the **basis for the induction**. Step (b) is called the **induction step**. The assumption made in this step, that $P(k)$ is true for some integer k, is called the **induction hypothesis**.

Worked Exercise 3.3

Prove the following formula for the sum of the first n triangular numbers.

$$1 + 3 + 6 + 10 + \cdots + \tfrac{1}{2}n(n+1) = \tfrac{1}{6}n(n+1)(n+2) \qquad P(n)$$

By 'prove' we mean 'prove for all integers $n \geq 1$'. The 'for all integers $n \geq 1$' is taken for granted.

Solution

We use mathematical induction to prove that $P(n)$ is true for all integers $n \geq 1$ by establishing (a) and (b) above.

First the basis for the induction. Putting $n = 1$ in $P(n)$ gives

$$1 = \tfrac{1}{6} \times 1(1+1)(1+2).$$

This is correct, showing that $P(1)$ is true.

Now the induction step. We assume that $P(k)$ is true for some natural number k. That is, we have the induction hypothesis

$$1 + 3 + 6 + 10 + \cdots + \tfrac{1}{2}k(k+1) = \tfrac{1}{6}k(k+1)(k+2).$$

We need to deduce the truth of $P(k+1)$ from this, that is

$$1 + 3 + 6 + 10 + \cdots + \tfrac{1}{2}k(k+1) + \tfrac{1}{2}(k+1)(k+2) = \tfrac{1}{6}(k+1)(k+2)(k+3).$$

The last term on the left-hand side is the $(k+1)$th triangular number.

Taking the left-hand side, we have

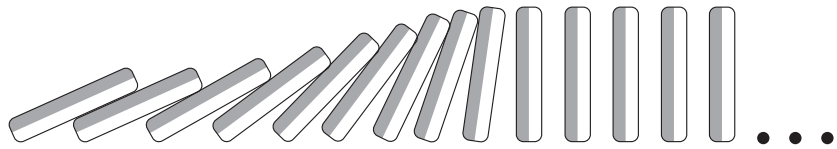
$$\begin{aligned} &1 + 3 + 6 + 10 + \cdots + \tfrac{1}{2}k(k+1) + \tfrac{1}{2}(k+1)(k+2) \\ &= \tfrac{1}{6}k(k+1)(k+2) + \tfrac{1}{2}(k+1)(k+2), \\ &\qquad\qquad \text{by the induction hypothesis,} \\ &= (k+1)(k+2)(\tfrac{1}{6}k + \tfrac{1}{2}) \\ &= \tfrac{1}{6}(k+1)(k+2)(k+3), \quad \text{which is the RHS of } P(k+1). \end{aligned}$$

RHS is short for right-hand side and LHS is short for left-hand side of an equation.

This establishes the truth of $P(k+1)$ and completes the induction step.

Hence, by the Principle of Mathematical Induction, $P(n)$ is true for all natural numbers.

Mathematical induction can be likened to a line of dominoes (Figure 3.1), arranged so that if any one falls over, it will knock over the next one along the line (the induction step). Then, if the first domino is knocked over (the basis for the induction), they all fall over.



The first domino is knocked over

If the kth domino is knocked over, so too is the $(k+1)$th

Figure 3.1 Illustrating mathematical induction

Of course, the induction step is at the heart of any induction proof. Quite often, as in both the preceding worked exercises, some algebraic manipulation is involved. This may be straightforward or quite complex. The next exercise provides you with an opportunity to practise proof by induction.

Exercise 3.2

(a) Use mathematical induction to prove that the formula

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \tfrac{1}{6}n(n+1)(2n+1)$$

is true for all natural numbers.

(b) A **geometric series** is the sum of a finite sequence of terms in which the ratio of successive terms is constant, r in the following example.

Use mathematical induction to prove the formula for the geometric series

$$a + ar + ar^2 + ar^3 + \cdots + ar^{n-1} = \frac{a(r^n - 1)}{r - 1}, \quad r \neq 1.$$

3.3 More general induction

The examples that we have seen using mathematical induction have all concerned propositions $P(n)$ that are true for all natural numbers. In this case $n = 1$, the smallest integer for which the proposition was true, provided the basis for the induction. In fact there is nothing special about the number 1 here. The Principle of Induction is readily adapted to provide proofs of results that are true for all integers greater than some integer n_0.

Principle of Mathematical Induction (generalised)

Let $P(n)$ be a proposition depending on an integer n. If:

(a) $P(n_0)$ is true

(b) for any integer $k \geq n_0$, if $P(k)$ is true then $P(k+1)$ is true

then $P(n)$ is true for all integers $n \geq n_0$.

As before, step (a) of the generalised principle is called the basis for the induction, and step (b) is called the induction step.

In the next worked exercise we use this generalised form of induction to prove a result that is true for all integers from 10 onwards. In this example the induction step involves more reasoning than the simple algebraic manipulations we have met so far.

Worked Exercise 3.4

For which natural numbers n is it true that $2^n > n^3$?

Solution

Exploration of the relative values of 2^n and n^3 for small values of n shows that 2^n is larger when $n = 1$, but then n^3 is larger for $n = 2, 3, 4, \ldots, 9$. For $n = 10$, 2^n becomes the larger value once again:

$$2^{10} = 1024 > 1000 = 10^3.$$

As n now increases it appears that 2^n grows more quickly than n^3, which leads us to conjecture:

$$2^n > n^3 \text{ for all integers } n \geq 10.$$

Preparing the way for induction, we let $P(n)$ be the statement that $2^n > n^3$.

Having seen that $P(10)$ is true, the basis for the induction is established (as $n_0 = 10$).

It remains to show that, for any integer $k \geq 10$, if $P(k)$ is true then $P(k+1)$ is true. So the induction hypothesis is

$$2^k > k^3, \quad \text{for some } k \geq 10,$$

and from this assumption we need to show that

$$2^{k+1} > (k+1)^3.$$

Now, $2^{k+1} = 2 \times 2^k$ and so the induction hypothesis gives

$$2^{k+1} > 2k^3.$$

Therefore it suffices to show that, for any $k \geq 10$,

$$2k^3 \geq (k+1)^3,$$

or, rearranging,

$$\left(1 + \tfrac{1}{k}\right)^3 \leq 2.$$

This is certainly true by the following argument. As $k > 1$, the largest value of $\left(1 + \frac{1}{k}\right)^3$ comes from the largest value of $1 + \frac{1}{k}$. This in turn comes from the smallest value of k, namely $k = 10$. However, $\left(1 + \frac{1}{10}\right)^3 = 1.331$, which is less than 2. This completes the induction step.

Hence, by the Generalised Principle of Mathematical Induction, the proposition is true for all integers $n \geq 10$.

Worked Exercise 3.4 has some points of interest. Investigation of the induction step led us to the inequality

$$\left(1 + \tfrac{1}{k}\right)^3 \leq 2.$$

In fact this inequality holds true for all integers $k \geq 4$ and so the induction step works for all integers $k \geq 4$ (although we were concerned only with integers $k \geq 10$). Drawing on the earlier analogy with dominoes, the situation in Worked Exercise 3.4 is as depicted in Figure 3.2. If any domino from the fourth onwards were to fall it would knock the next over, but the first one of these that *does* fall is the tenth.

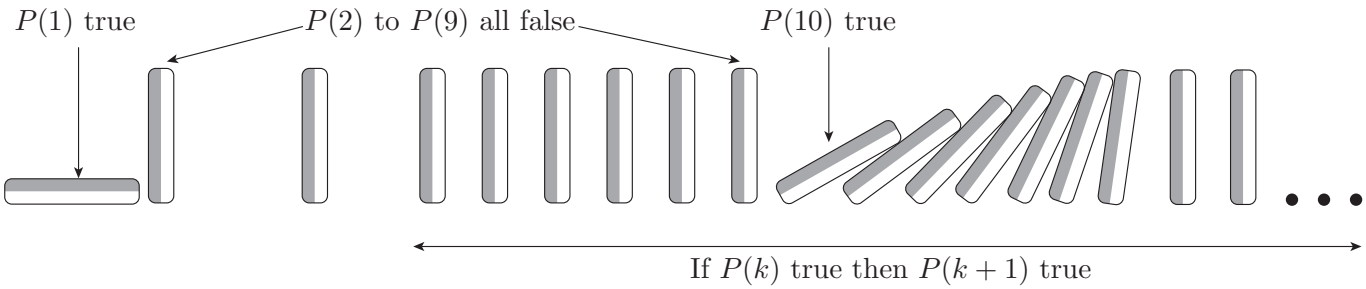


Figure 3.2 $P(n) : 2^n > n^3$ for $n \geq 10$

In the examples so far, when establishing the induction step, we assumed that the proposition under question was true for some integer $k \geq n_0$. On occasion it can be helpful to make a more general assumption. The *Second Principle of Mathematical Induction* employs a different induction step.

Second Principle of Mathematical Induction

Let $P(n)$ be a proposition depending on an integer n. If:

(a) $P(n_0)$ is true

(b′) for any integer $k \geq n_0$, if $P(n_0)$, $P(n_0+1)$, …, $P(k)$ are all true, then $P(k+1)$ is true

then $P(n)$ is true for all integers $n \geq n_0$.

The Second Principle of Mathematical Induction seems logically sound if one thinks of the domino analogy, and can be established from the Well-Ordering Principle. We will not do so here but proof of this can be found in the solution to Exercise 3.4(d).

The next two worked exercises both make use of the Second Principle of Mathematical Induction.

Worked Exercise 3.5

A sequence of integers is defined as follows:

$$x_0 = 1; \quad x_n = x_0 + x_1 + \cdots + x_{n-1}, \quad \text{for all integers } n \geq 1.$$

Prove that $x_n = 2^{n-1}$, for all integers $n \geq 1$.

Solution

Denote the proposition $x_n = 2^{n-1}$ by $P(n)$.

Then $x_1 = x_0 = 1 = 2^0$, so $P(1)$ is true and we have the basis for the induction.

Heading towards the (new) induction step, suppose that $P(r)$ is true for all $1 \leq r \leq k$; that is,

$$x_r = 2^{r-1}, \quad \text{for } 1 \leq r \leq k.$$

Then

$$\begin{aligned} x_{k+1} &= x_0 + x_1 + \cdots + x_k \\ &= 1 + \left(1 + 2 + 4 + \cdots + 2^{k-1}\right), \text{ by the induction hypothesis,} \\ &= 1 + \frac{2^k - 1}{2 - 1}, \text{ by formula for geometric series, Exercise 3.2(b),} \\ &= 2^k. \end{aligned}$$

This establishes the truth of $P(k+1)$ and completes the induction step.

Hence the result follows by the Second Principle of Mathematical Induction.

Although there are other ways of proving the result in Worked Exercise 3.5, it is not easy to prove using mathematical induction without making use of the Second Principle. It is the fact that each number in the sequence is defined in terms of *all* the preceding numbers that makes the Second Principle relevant. In Worked Exercise 3.6 we employ the Second Principle of Mathematical Induction again, but this time it will be used differently.

Worked Exercise 3.6

Prove that every natural number n can be written as a sum of distinct (integer) powers of 2. For example, $13 = 2^0 + 2^2 + 2^3$ and $20 = 2^2 + 2^4$.

You might be familiar with the fact that each natural number is expressed *uniquely* in this way: we are essentially representing n in binary. However, we omit the proof of uniqueness here.

Solution

Let $P(n)$ be the proposition that n can be written as a sum of distinct powers of 2.

Then, since $1 = 2^0$, $P(1)$ is true and we have the basis for the induction.

To make use of the Second Principle we assume that, for some natural number k, $P(r)$ is true for each integer r in the range $1 \leq r \leq k$. That is, we assume that each r in this range can be expressed in the required way. To complete the induction proof we show that, under this assumption, $k+1$ can be expressed as a sum of distinct powers of 2, thereby establishing the truth of $P(k+1)$.

We consider two cases, depending on whether $k+1$ is even or odd.

Case $k+1$ even

Since $k+1$ is even we can write $k+1 = 2s$ for some natural number s. Now $s < k+1$ so, by the induction hypothesis, s can be written as a sum of distinct powers of 2, say

$$s = 2^a + 2^b + \cdots + 2^f, \quad a, b, \ldots, f \text{ distinct.}$$

Multiplying this sum by 2 gives

$$k+1 = 2s = 2^{a+1} + 2^{b+1} + \cdots + 2^{f+1}, \quad a+1, b+1, \ldots, f+1 \text{ distinct,}$$

as required.

Case $k+1$ odd

Since $k+1$ is odd we can write $k+1 = 2s+1$ for some natural number s. Now $2s < k+1$ so, by the induction hypothesis, $2s$ can be written as a sum of distinct powers of 2, say

$$2s = 2^a + 2^b + \cdots + 2^f, \quad a, b, \ldots, f \text{ distinct.}$$

Since $2s$ is even, and the exponents in this representation of $2s$ are distinct, no exponent can be 0 (because the only power of 2 giving an odd integer is $2^0 = 1$).

Hence

$$\begin{aligned} k+1 &= 2s+1 \\ &= 2^a + 2^b + \cdots + 2^f + 1, \quad a, b, \ldots, f \text{ distinct and non-zero,} \\ &= 2^a + 2^b + \cdots + 2^f + 2^0, \quad a, b, \ldots, f, 0 \text{ distinct,} \end{aligned}$$

as required.

This completes the induction step and the result follows by the Second Principle of Mathematical Induction.

Exercise 3.3

In the sequence

$$1,\ 3,\ 4,\ 7,\ 11,\ 18,\ 29,\ 47,\ \ldots,$$

each term, from the third onwards, is the sum of the previous two terms. That is, the sequence $\{L_n\}$ is defined by

$$L_1 = 1; \quad L_2 = 3; \quad L_n = L_{n-1} + L_{n-2}, \quad \text{for } n \geq 3.$$

Use the Second Principle of Mathematical Induction to prove that

$$L_n < \left(\tfrac{7}{4}\right)^n, \quad \text{for all } n \geq 1.$$

Mathematical induction provides a powerful method of proof, whose applications include proving the truth of formulas for all integers from some integer onwards. But while mathematical induction is a great help in proving formulas, it does not help us to find such formulas. Discovering the formulas in the first place is a different matter.

We finish this section with an exercise that sets several problems on proof by mathematical induction.

Exercise 3.4

(a) Use the Principle of Mathematical Induction to prove that the following formulas are true for all natural numbers n.

(i) $1^2 + 3^2 + 5^2 + \cdots + (2n-1)^2 = \frac{1}{3}n(4n^2 - 1)$

(ii) $1 + 5 + 12 + 22 + \cdots + \frac{1}{2}n(3n-1) = \frac{1}{2}n^2(n+1)$

Note that part (ii) establishes the formula for sums of pentagonal numbers derived as $Q(5, n)$ in Exercise 2.4(a) of Section 2 above.

(b) The **factorial** of $k \in \mathbb{N}$ is $k! = 1 \times 2 \times \cdots \times k$.

Prove the following formula for all natural numbers n.

$$1 \times (1!) + 2 \times (2!) + 3 \times (3!) + \cdots + n \times (n!) = (n+1)! - 1$$

(c) For which natural numbers n is it true that $n! > 6n^2$? Prove your result using mathematical induction.

(d) Show how the Second Principle of Mathematical Induction can be derived from the Well-Ordering Principle, as follows.

Suppose that the conditions (a) and (b′) stated in the Second Principle of Mathematical Induction hold. Let S be the set of integers k, $k \geq n_0$, for which $P(k)$ does not hold, that is,

$$S = \{k \in \mathbb{Z} : k \geq n_0 \text{ and } P(k) \text{ is false}\}.$$

Suppose that S is non-empty and use this assumption to deduce that the set

$$T = \{s - n_0 : s \in S\}$$

is a non-empty set of natural numbers. Show how the existence of a least member of T leads to a contradiction.

4 Divisibility

This section introduces some useful properties of division, as well as the connected notions of highest common factor and least common multiple.

4.1 Integer division

From an early age we learn how to divide one natural number by another, obtaining a quotient and a remainder. For example, we can divide 23 by 4 to obtain a quotient of 5 and a remainder of 3 or, rather more formally, $23 = 4 \times 5 + 3$. In fact, if we stipulate that when dividing by 4 the only permitted remainders are 0, 1, 2 and 3 then the quotient and remainder in any division by 4 turn out to be unique. That is, the only way that we can write $23 = 4 \times q + r$, where q and r are integers with $0 \leq r < 4$, is $q = 5$ and $r = 3$. Such uniqueness holds for division by any natural number. The familiar result exemplified here, known as the *Division Algorithm*, is of enormous theoretical importance. In what follows we will discuss certain consequences of it that are fundamental to our treatment of numbers.

Theorem 4.1 *The Division Algorithm*

For any two integers a and b, where $b > 0$, there exist unique integers q and r such that

$$a = bq + r, \quad \text{where } 0 \leq r < b.$$

Proof Consider the set of integers

$$S = \{a - bn \geq 0 : n \in \mathbb{Z}\}.$$

S is the set of non-negative integers in the set $\{\ldots, a-2b, a-b, a, a+b, \ldots\}$.

The set S is non-empty because, as b is positive, $a - bn$ will certainly be positive for large negative values of n. Either 0 is a member of S, and hence the least member of S, or S is a non-empty set of natural numbers and, by the Well-Ordering Principle, again has a least member. So in either case there exists an integer q such that $a - bq = r$ is the least member of S. Hence we have integers q and r such that

$$a = bq + r, \quad \text{where } 0 \leq r.$$

If $r \geq b$ then $r - b \geq 0$. In addition,

$$r - b = a - bq - b = a - b(q + 1)$$

and so $r - b$ is a member of S. However, $r - b$ is smaller than r, which contradicts the minimality of r, and so $r < b$.

It remains to prove the uniqueness of q and r. Suppose, therefore, that

$$a = bq + r, \quad 0 \leq r < b,$$

and

$$a = bq' + r', \quad 0 \leq r' < b.$$

Subtracting gives

$$0 = b(q - q') + (r - r').$$

Now if $q = q'$ this equation gives $r = r'$ and so the expressions for a are the same.

On the other hand, if $q \neq q'$, we may assume that $q > q'$ and hence that $q - q' > 0$. It follows that $q - q' \geq 1$. Since $b > 0$ we have

$$\begin{aligned} r' - r &= b(q - q') \\ &\geq b \times 1 = b. \end{aligned}$$

Adding r to both sides gives

$$r' \geq b + r \geq b,$$

contradicting the definition of r'.

Therefore the values of q and r are unique. ■

Essentially the same argument is illustrated as follows. Imagine multiples of b stepped off along the number line, as shown in Figure 4.1. As the intervals marked off in this way cover the whole line, the number a lies in some interval

$$bq \leq a < b(q + 1).$$

This interval containing a will be unique. Notice that if a is a multiple of b then a coincides with one of the marks between the intervals. In this case, the interval to either side of a could be chosen to contain a. However, the inequalities decree that we choose the interval that has a at its left-hand end.

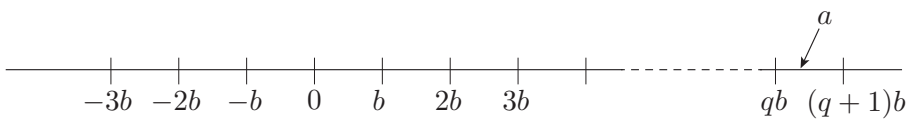


Figure 4.1 Illustrating the Division Algorithm

Subtracting bq through the above inequality shows us that there is a unique integer q such that $0 \leq a - bq < b$. Now putting $a - bq = r$ gives the claimed result.

The way of representing numbers suggested by the Division Algorithm will be important to us, so let us pause to look at some simple applications.

Worked Exercise 4.2

Show that when any square is divided by 4 the remainder is either 0 or 1.

Solution

The Division Algorithm, applied to division by 4, tells us that any integer can be written uniquely in one of the forms $4n$, $4n+1$, $4n+2$ or $4n+3$ for some integer n. So any square number is the square of an integer of one of these four forms. We write each square in its unique form $4q+r$ and hope to find that the only possible values for r are 0 and 1.

$$(4n)^2 = 16n^2 = 4(4n^2) + 0$$
$$(4n+1)^2 = 16n^2 + 8n + 1 = 4(4n^2 + 2n) + 1$$
$$(4n+2)^2 = 16n^2 + 16n + 4 = 4(4n^2 + 4n + 1) + 0$$
$$(4n+3)^2 = 16n^2 + 24n + 9 = 4(4n^2 + 6n + 2) + 1$$

In each case the remainder on division by 4 is either 0 or 1, as claimed.

In fact we could have made the calculations simpler by working with remainders on division by 2. Any number is either of form $2n$ or of form $2n+1$ and the respective squares are

$$(2n)^2 = 4n^2 + 0,$$
$$(2n+1)^2 = 4n^2 + 4n + 1 = 4(n^2 + n) + 1,$$

showing that the remainder on division by 4 is either 0 or 1.

Exercise 4.1

The Division Algorithm tells us that any integer can be written in one of the forms $3n$, $3n+1$ or $3n+2$. Use this fact to deduce that when any cube is divided by 9 the remainder is one of 0, 1 or 8.

Before moving on, we would like to highlight two formulas with which you are likely to be familiar and which can readily be verified by expanding the brackets.

$$(an+b)^2 = a^2n^2 + 2abn + b^2$$
$$(an+b)^3 = a^3n^3 + 3a^2bn^2 + 3ab^2n + b^3$$

Knowledge of these formulas can be useful, as was the case in the solutions to Worked Exercise 4.2 and Exercise 4.1.

The Division Algorithm provides us with our definition of divisibility: one number is divisible by another when the unique remainder is 0.

Definition 4.3 *Factors and multiples*

An integer a is divisible by the natural number b (or, for brevity, b **divides** a) if there exists some integer q such that $a = bq$.

When b divides a we say that b is a **factor** of a or that a is a **multiple** of b.

We write $b \mid a$ as a shorthand for b divides a, and $b \nmid a$ for b does not divide a.

In other texts you may find the phrase 'b is a **divisor** of a' rather than 'b is a factor of a'.

Be careful not to confuse the statement $b \mid a$ with the fraction $\frac{b}{a}$. For example, $3 \mid 21$ since $21 = 3 \times 7$ and $6 \mid -24$ since $-24 = 6 \times (-4)$. On the other hand, $6 \nmid 16$ since $16 = 6 \times 2 + 4$.

Notice that we consider division only by positive integers. The definition could be adapted to permit division by negative numbers by declaring that a is divisible by b (positive or negative) when there exists an integer q such that $a = bq$. But if $a = bq$ it must be the case that $a = (-b)(-q)$ and so this definition would lead to b being a factor of a if, and only if, $-b$ is a factor of a. There is therefore nothing essentially complicated in dividing by negative numbers, but we will have no occasion to do so. Therefore, if we ask for the factors of an integer, we expect answers that are positive.

Notice also that 0 is excluded as a factor, as the only number for which 0 is a factor is 0 itself. Additionally, since $0 = n \times 0$, every natural number is a factor of 0; in fact 0 is unique in having an infinite number of factors, as will be shown in the proof of Theorem 4.4(b).

Exercise 4.2

(a) List all the factors of 18 and -24. Do these two numbers have any factors in common?

(b) Show that if n is not a multiple of 3 then $n^2 - 1$ is divisible by 3.

Exercise 4.2(a) highlights one property of divisibility, namely that any integer greater than 1 has at least two factors, namely itself and 1. This is just one of a number of simple properties that are direct consequences of the definition. Such properties are not particularly exciting in themselves but some of them will prove to be useful as tools for deriving further information. A selection of the more important properties have been grouped together in the following theorem.

Theorem 4.4 *Properties of division*

Let a and b be natural numbers and c and d be any integers.

(a) If $a \mid c$ then $a \mid (c + na)$ for any integer n.

(b) If $c \neq 0$ and $a \mid c$ then $a \leq |c|$.

(c) If $a \mid b$ and $b \mid a$ then $a = b$.

(d) If $a \mid b$ and $b \mid c$ then $a \mid c$.

(e) If $a \mid c$ and $a \mid d$ then $a \mid (mc + nd)$ for any integers m and n.

Proof

(a) As $a \mid c$ there is an integer q such that $c = aq$. But then

$$c + na = aq + na = a(q + n) = ax,$$

and as $x = q + n$ is an integer, this confirms that a divides $c + na$.

(b) As $c = aq$ as in (a), $|c| = |aq| = a|q|$, and the result follows since $|q| \geq 1$, q being a non-zero integer.

Notice that this result shows that the non-zero integer c can have only a finite number of factors, since they must all be less than or equal to $|c|$.

(c) If $a \mid b$ and $b \mid a$ then, from (b), $a \leq b$ and $b \leq a$. The required equality follows.

(d) If $a \mid b$ and $b \mid c$ then there are integers s and t such that $b = as$ and $c = bt$. Substituting for b gives $c = ast = a(st)$, and as st is an integer this shows that $a \mid c$.

(e) If $a \mid c$ and $a \mid d$ then there are integers p and q such that $c = ap$ and $d = aq$. But then, for any integers m and n,

$$mc + nd = map + naq = a(mp + nq) = ax,$$

where $x = mp + nq$ is an integer. So $a \mid (mc + nd)$. ■

4.2 The highest common factor

In Exercise 4.2(a) we saw that the two numbers 18 and -24 had four common factors, namely 1, 2, 3 and 6. It is clear that any two integers will have at least one common factor because 1 divides every integer. What is more, property (b) of the previous theorem guarantees that any integer (other than 0) has only finitely many factors. Therefore any two integers, not both of which are 0, have a non-empty, finite set of common factors. We are going to be interested in the *highest common factor* of two integers, that is, the largest member of the finite set of common factors.

Highest common factors are also known as **greatest common divisors** with the corresponding notation $\gcd(a, b)$.

When we write $\text{hcf}(a, b)$ it will be understood that a and b are integers not both of which are zero.

Definition 4.5 *Highest common factor*

The **highest common factor** of two integers a and b, not both of which are zero, is the natural number n satisfying the following:

(a) $n \mid a$ and $n \mid b$

(b) if $d \mid a$ and $d \mid b$ then $d \leq n$.

The highest common factor of a and b is denoted by $\text{hcf}(a, b)$.

Property (a) of this definition says that $\text{hcf}(a, b)$ *is* a common factor of a and b, while property (b) says that it is *maximal* amongst the common factors.

Exercise 4.2(a) establishes that $\text{hcf}(18, -24) = 6$. Similarly, as you may wish to verify, $\text{hcf}(3, 20) = 1$ and $\text{hcf}(14, 35) = 7$.

Two properties that follow immediately from the definition are worth recording.

Since zero is divisible by every natural number, it follows that

$$\text{hcf}(0, b) = |b|.$$

Since n and $-n$ have the same set of positive factors,

$$\text{hcf}(a, b) = \text{hcf}(-a, b) = \text{hcf}(a, -b) = \text{hcf}(-a, -b).$$

Property (e) of Theorem 4.4 shows that any common factor of two integers c and d also divides any integer of the form $mc + nd$. Numbers of this latter form merit their own name.

Definition 4.6 *Integer combination*

If a and b are integers, then any integer of the form

$$ma + nb, \quad m, n \in \mathbb{Z},$$

is called an **integer combination** of a and b.

So $\text{hcf}(a, b)$ divides every integer combination of a and b. In fact more is true: one of these integer combinations *is* $\text{hcf}(a, b)$.

To illustrate that claim, take the three examples of highest common factor mentioned above.

$$\text{hcf}(18, -24) = 6 = 3 \times 18 \ + \ 2 \times (-24)$$
$$\text{hcf}(3, 20) = 1 = 7 \times 3 \ + \ (-1) \times 20$$
$$\text{hcf}(14, 35) = 7 = (-2) \times 14 \ + \ 1 \times 35$$

Proposition 4.7

Given any integers a and b, not both zero, there exist integers m and n such that

$$\text{hcf}(a, b) = ma + nb.$$

In other words, $\text{hcf}(a, b)$ is an integer combination of a and b.

Proof Consider the set

$$S = \{xa + yb > 0 : x, y \text{ integers}\}$$

of all positive integer combinations of a and b. Each member of S is an integer and S is non-empty because a, $-a$, b and $-b$ are all integer combinations of a and b of which at least one is positive. Hence the Well-Ordering Principle guarantees the existence of a smallest element d in S. As d is a member of S we know that $d = ma + nb$ for some integers m and n. It is our contention that $d = \text{hcf}(a, b)$.

We first show that d is a common factor of a and b. The Division Algorithm tells us that, from division of a by d, we have $a = dq + r$ for some integers q and r with $0 \leq r < d$. But then

$$\begin{aligned} r &= a - dq \\ &= a - (ma + nb)q \\ &= (1 - mq)a + (-nq)b, \end{aligned}$$

showing that r is an integer combination of a and b. Now if it were the case that $r > 0$ then r would be a member of S and the condition $r < d$ would contradict d being the least member of S. Hence $r = 0$ and d is a factor of a. Similarly d is also a factor of b.

It remains to show that d is the *greatest* of the common factors. Suppose then that d' is any common factor of a and b. Then, by property (e) of Theorem 4.4, d' is a factor of the integer combination $ma + nb$; that is, d' is a factor of d. Property (b) of Theorem 4.4 then gives $d' \leq d$, as required. ■

The ability to write the highest common factor of two integers as an integer combination of them will prove to be very useful. Unfortunately, though our result guarantees that an integer combination exists, it gives us no clue as to how to find that integer combination. We will return to that problem shortly, but first we establish a number of useful consequences of the proposition. The first two concern pairs of integers for which the highest common factor is 1. We have a term to describe such pairs.

Definition 4.8 *Coprime*

Two integers a and b, not both zero, are said to be **coprime**, or **relatively prime**, whenever $\text{hcf}(a, b) = 1$.

For example, 10 and 21 are coprime, since 10 and 21 do not have any common factors other than 1.

We exclude the case where a and b are both zero because $\text{hcf}(0,0)$ does not exist. For any non-zero integer a, $\text{hcf}(0,a) = |a|$ and hence the only integers that are coprime to 0 are 1 and -1.

Lemma 4.9

Integers a and b are coprime if, and only if, there exist integers m and n such that $1 = ma + nb$.

Proof If a and b are coprime then $\text{hcf}(a,b) = 1$ and Proposition 4.7 confirms the existence of integers m and n with $ma + nb = 1$.

Conversely, suppose there are integers m and n with $ma + nb = 1$, and let d be any common factor of a and b. Then property (e) of Theorem 4.4 tells us that d divides $ma + nb$; that is, d divides 1. Finally property (b) of Theorem 4.4 gives $d = 1$, and $\text{hcf}(a,b) = 1$, as claimed. ■

We will see in Exercise 4.3(b) that the integers m and n in the statement of Lemma 4.9 are also coprime. Indeed, if there exist non-zero integers a, b, m and n such that $ma + nb = 1$ then each of a and m is coprime to each of b and n.

Proposition 4.10

For any integers a and b, not both zero, if $\text{hcf}(a,b) = d$ then $\frac{a}{d}$ and $\frac{b}{d}$ are integers such that $\text{hcf}\left(\frac{a}{d}, \frac{b}{d}\right) = 1$.

That is, the integers $\dfrac{a}{\text{hcf}(a,b)}$ and $\dfrac{b}{\text{hcf}(a,b)}$ are coprime.

We often use this in the following form.

If $\text{hcf}(a,b) = d$ there exist integers s and t such that $a = ds$ and $b = dt$, with $\text{hcf}(s,t) = 1$.

The s and t are respectively $\frac{a}{d}$ and $\frac{b}{d}$.

Proof Since d divides a and d divides b there exist integers s and t such that $a = ds$ and $b = dt$. Moreover, Proposition 4.7 gives the existence of integers m and n such that $d = ma + nb$. Dividing this equation through by d leaves us with $1 = ms + nt$, whereupon Lemma 4.9 informs us that $\text{hcf}(s,t) = 1$. That is,

$$\text{hcf}\left(\tfrac{a}{d}, \tfrac{b}{d}\right) = 1. \quad ■$$

In the statement of the proposition we used $\frac{a}{d}$ and $\frac{b}{d}$, knowing that they were integers. In general it is advisable to avoid the use of fractions when trying to prove results concerning integers, since it is important to maintain distinction between the two. Note how, in the proof, we introduced the integers s and t instead of using $\frac{a}{d}$ and $\frac{b}{d}$ respectively.

We take a break from the string of results at this point and let you try your hand at some simple proofs. Each of the problems in Exercise 4.3 involves little more than an appropriate application of one of the preceding results.

Exercise 4.3

Let a and b be integers, not both of which are zero. Prove the following.

(a) The integer c is a multiple of $\text{hcf}(a, b)$ if, and only if, there exist integers x and y such that $c = xa + yb$.

(b) If $\text{hcf}(a, b) = ma + nb$ then $\text{hcf}(m, n) = 1$.

(c) If $\text{hcf}(a, b) = 1$ and $\text{hcf}(a, c) = 1$ then $\text{hcf}(a, bc) = 1$.

(d) For any natural number k, $\text{hcf}(ka, kb) = k \times \text{hcf}(a, b)$.

Note that the result in part (a) asserts that the set of all integer combinations of a and b is precisely the set of all integer multiples of $\text{hcf}(a, b)$.

The next two results establish further divisibility properties.

Lemma 4.11

If $a \mid c$ and $b \mid c$, with $\text{hcf}(a, b) = 1$, then $ab \mid c$.

Proof As a divides c and b divides c, there are integers r and s such that

$$c = ar \quad \text{and} \quad c = bs.$$

Also, $\text{hcf}(a, b) = 1$ implies the existence of integers m and n such that

$$ma + nb = 1.$$

Multiplying through this equation by c, we obtain

$$cma + cnb = c.$$

Substituting bs for the first c and ar for the second c in the left-hand side of this equation gives $ab(sm + rn) = c$, which confirms that ab divides c. ■

Theorem 4.12 *Euclid's Lemma*

If $a \mid bc$, with $\text{hcf}(a, b) = 1$, then $a \mid c$.

Proof Since a divides bc and $\text{hcf}(a, b) = 1$, we have integers k, m and n such that

$$bc = ka \quad \text{and} \quad ma + nb = 1.$$

Multiplying through the latter by c gives

$$mac + nbc = c$$

and then substitution for bc leaves us with

$$mac + nka = c.$$

That is, $a(mc + nk) = c$, which confirms that a divides c, as required. ■

The following worked exercise incorporates several of the preceding ideas.

Worked Exercise 4.13

(a) Suppose that a divides bc, where $\text{hcf}(b, c) = 1$. Show that a can be written as a product rs, where r divides b and s divides c.

(b) Show that if a is an integer with $\text{hcf}(6, a) = 1$ then 24 divides $a^2 - 1$.

Solution

(a) To illustrate the result, notice that 21 divides 6×35, and $21 = 3 \times 7$, where 3 divides 6 and 7 divides 35, and this is the only way of expressing 21 as a product of two numbers to fit the statement.

Let $r = \text{hcf}(a, b)$. By Proposition 4.10, there exist integers s and t such that

$$a = rs \quad \text{and} \quad b = rt, \quad \text{where } \text{hcf}(s, t) = 1.$$

Hence we have written $a = rs$, where r divides b. It remains to show that s divides c.

Since a divides bc there exists an integer k such that

$$bc = ka.$$

That is, substituting for a and b,

$$rtc = krs.$$

Cancelling the r on both sides gives

$$tc = ks,$$

which implies that s divides tc. However, s and t are coprime and so, by Euclid's Lemma, s divides c, as required.

In fact the only integers r and s to fit the statement are $r = \text{hcf}(a, b)$ and, by symmetry, $s = \text{hcf}(a, c)$.

(b) We show that 3 divides $a^2 - 1$ and 8 divides $a^2 - 1$. The result then follows from Lemma 4.11, since $\text{hcf}(3, 8) = 1$.

Thinking about division by 3, we note that a must be of one of the forms $3n$, $3n + 1$ or $3n + 2$. However, since $\text{hcf}(6, 3n) \geq 3$, the condition that $\text{hcf}(6, a) = 1$ rules out the first of these possibilities. For the case $a = 3n + 1$ we have

$$a^2 - 1 = (3n + 1)^2 - 1 = 9n^2 + 6n + 1 - 1 = 3(3n^2 + 2n),$$

and for the case $a = 3n + 2$ we have

$$a^2 - 1 = (3n + 2)^2 - 1 = 9n^2 + 12n + 4 - 1 = 3(3n^2 + 4n + 1).$$

So in each case $a^2 - 1$ is divisible by 3.

Turning to division by 8, there are eight possible forms for a, namely $8n + r$, for $r = 0, 1, \ldots, 7$. However, we notice that $\text{hcf}(a, 6) = 1$ tells us that a cannot be even. Thus we are left with the odd cases, which are $8n + 1$, $8n + 3$, $8n + 5$ and $8n + 7$. Taking these four cases in turn, we have

$$(8n + 1)^2 - 1 = 64n^2 + 16n + 1 - 1 = 8(8n^2 + 2n),$$
$$(8n + 3)^2 - 1 = 64n^2 + 48n + 9 - 1 = 8(8n^2 + 6n + 1),$$
$$(8n + 5)^2 - 1 = 64n^2 + 80n + 25 - 1 = 8(8n^2 + 10n + 3),$$
$$(8n + 7)^2 - 1 = 64n^2 + 112n + 49 - 1 = 8(8n^2 + 14n + 6).$$

In each case $a^2 - 1$ is divisible by 8.

This completes the proof.

Next an exercise for you to try, which should give you practice at applying some of the preceding results.

Exercise 4.4

(a) Suppose that $\text{hcf}(a, b) = 1$. Use Lemma 4.9 to prove the following.

(i) If $c \mid a$ then $\text{hcf}(b, c) = 1$.

(ii) $\text{hcf}(a, a + b) = 1$.

(b) Prove that the following statements are true for any integer n.

(i) $\text{hcf}(3n + 2, 5n + 3) = 1$.

Hint: use property (e) of Theorem 4.4 in the following way. Suppose that d is any common factor of $3n + 2$ and $5n + 3$. Then d divides $(r(3n + 2) + s(5n + 3))$ for any integers r and s. Try to choose r and s so that d divides 1.

(ii) $\text{hcf}(n^2 + 3n + 1, 2n + 1) = 1$.

Hint: this time you may have to apply property (e) of Theorem 4.4 more than once.

4.3 The least common multiple

In contrast to common factors we also have the notion of common multiples. For example, since $6 \mid 60$ and $10 \mid 60$ we say that 60 is a common multiple of 6 and 10. Of course there are infinitely many common multiples of 6 and 10, the positive ones including 30, 60, 90, 120, 150, 180, and so on. Indeed, it would appear that the common multiples of 6 and 10 are precisely the multiples of 30, 30 being the *least common multiple* of 6 and 10.

Definition 4.14 *Least common multiple*

The **least common multiple** of the non-zero integers a and b is the natural number n satisfying the following:

(a) $a \mid n$ and $b \mid n$

(b) if $a \mid m$ and $b \mid m$ then $n \leq |m|$.

The least common multiple of a and b is denoted by $\operatorname{lcm}(a, b)$.

Property (a) of this definition says that $\operatorname{lcm}(a, b)$ is a common multiple of a and b, while property (b) says that it is the smallest natural number with this property. The fact that $|ab|$ is a common multiple of a and b, coupled with the Well-Ordering Principle, ensures that $\operatorname{lcm}(a, b)$ exists.

The example preceding the definition suggested something stronger than property (b), namely that $\operatorname{lcm}(a, b)$ divides any common multiple of a and b. As a first exercise in handling common multiples, we invite you to prove that this property does indeed follow from the definition.

Exercise 4.5

Show that if m is a common multiple of the non-zero integers a and b then $\operatorname{lcm}(a, b)$ divides m.

Hint: use the Division Algorithm and show that the remainder must be zero.

Not surprisingly, the twin concepts of highest common factor and least common multiple are interrelated. Recall from the earlier numerical example that $\operatorname{lcm}(6, 10) = 30$ and $\operatorname{hcf}(6, 10) = 2$. It is no accident that $\operatorname{lcm}(6, 10) \times \operatorname{hcf}(6, 10) = 6 \times 10$.

Proposition 4.15

For any pair of natural numbers a and b,

$$\operatorname{lcm}(a, b) \times \operatorname{hcf}(a, b) = ab.$$

The corresponding result in which either a or b is negative is $\operatorname{lcm}(a, b) \times \operatorname{hcf}(a, b) = |ab|$. It is readily deduced from this one.

Proof Let $d = \text{hcf}(a, b)$. Then since d divides a, it also divides ab. So there exists a natural number n such that

$$ab = dn.$$

Our task is to show that $n = \text{lcm}(a, b)$.

Since d divides both a and b, there exist integers u and v such that

$$a = du \quad \text{and} \quad b = dv.$$

Substituting for a and then b in $ab = dn$, and cancelling d in each case, gives

$$n = bu \quad \text{and} \quad n = av.$$

This shows that n is a common multiple of a and b, which leaves us to show it is the least such multiple.

Let m be any other common multiple of a and b. There exist integers r and s such that

$$m = ar \quad \text{and} \quad m = bs.$$

We also know that $d = \text{hcf}(a, b)$ is an integer combination of a and b, say

$$d = ax + by.$$

Multiplying this equation by m, and substituting first bs and then ar for m on the right-hand side, we get the following.

$$\begin{aligned} md &= max + mby \\ &= bsax + arby \\ &= ab(sx + ry) \\ &= dn(sx + ry), \quad \text{from the third line of the proof.} \end{aligned}$$

Lastly, cancellation of the d gives

$$m = n(sx + ry).$$

This shows that n is a factor of m and so $n \leq |m|$ as required. ■

Exercise 4.6

Use Proposition 4.15 to determine

(a) $\text{lcm}(18, 24)$; (b) $\text{lcm}(27, 45)$; (c) $\text{lcm}(39, 40)$.

The following exercise revises the concepts of divisibility and mathematical induction.

Exercise 4.7

(a) Show that an integer of the form $12k + 8$ is necessarily of the form $3m + 2$, but not conversely.

For which values of r in the range $0 \leq r \leq 11$ is $12k + r$ also of the form $3m + 2$?

(b) The number 64 is both a square and a cube:

$$64 = 8^2 = 4^3.$$

Show that any number that is both a square and a cube must be of one of the forms $7k$ or $7k + 1$.

(c) Use mathematical induction to prove that 9 divides $4^n + 6n - 1$ for all integers $n \geq 0$.

(d) Note that

$$36 \mid 108, \quad 36 \mid 1008, \quad 36 \mid 10\,008, \quad 36 \mid 100\,008, \quad \ldots.$$

Formulate the general result illustrated by this and prove your result using mathematical induction.

(e) Prove that if $\text{hcf}(a, b) = 1$ then, for any integer c,

$$\text{hcf}(ac, b) = \text{hcf}(c, b).$$

(f) Suppose that $\text{hcf}(a, b) = 1$. Show that:

(i) $\text{hcf}(2a + b, a + b) = 1$

(ii) $\text{hcf}(a + b, a - b) = 1$ or 2.

5 Linear Diophantine equations

Pythagoras and Euclid developed much of the material discussed in this section. You may wish to learn more about these two mathematicians in the *History Reader* for this topic.

5.1 The Euclidean Algorithm

We have seen that the problem of finding least common multiples can readily be solved by first finding the corresponding highest common factor. However, as yet we have no systematic method of finding highest common factors. Given two positive numbers, a and b, we could determine a list of all factors for each and select the largest one in common. That is fine for reasonably small integers, but would you want to tackle finding $\text{hcf}(3108, 5291)$ by such an approach? There is a more efficient method that uses nothing more than repeated division, as in the Division Algorithm; it is called the *Euclidean Algorithm*. First we ask you to study a numerical example illustrating the procedure before returning to see why it works.

Worked Exercise 5.1

Determine $\text{hcf}(3108, 5291)$.

Solution

We begin by dividing the larger number by the smaller.

$$5291 = 1 \times 3108 + 2183$$

Next we divide 3108 by the remainder in the above equation.

$$3108 = 1 \times 2183 + 925$$

We continue in this fashion, at each stage dividing the previous factor by the resulting remainder.

$$\begin{aligned} 2183 &= 2 \times 925 + 333 \\ 925 &= 2 \times 333 + 259 \\ 333 &= 1 \times 259 + 74 \\ 259 &= 3 \times 74 + 37 \\ 74 &= 2 \times 37 + 0 \end{aligned}$$

The process stops when a remainder of 0 is encountered and the last non-zero remainder, namely 37, is the required $\text{hcf}(3108, 5291)$.

Let us spell out generally the system of equations that arise in applying the **Euclidean Algorithm** to determine $\text{hcf}(a, b)$. Notice that since $\text{hcf}(-a, b) = \text{hcf}(a, b)$ we can safely assume that the numbers in question are both positive. Repeated application of the Division Algorithm gives the following.

$$\begin{array}{lll} a & = q_1 \times b + r_1 & 0 < r_1 < b \\ b & = q_2 \times r_1 + r_2 & 0 < r_2 < r_1 \\ r_1 & = q_3 \times r_2 + r_3 & 0 < r_3 < r_2 \\ r_2 & = q_4 \times r_3 + r_4 & 0 < r_4 < r_3 \\ & \dots & \\ r_{n-2} & = q_n \times r_{n-1} + r_n & 0 < r_n < r_{n-1} \\ r_{n-1} & = q_{n+1} \times r_n + 0 & \end{array}$$

The first fact to observe is that b and the remainders form a strictly decreasing sequence of natural numbers:

$$b > r_1 > r_2 > \cdots > r_{n-2} > r_{n-1} > r_n > 0.$$

This establishes that the algorithm must terminate; a zero remainder must be reached in a finite number of steps.

The outstanding claim to be justified is that the final non-zero remainder, namely r_n, is equal to the highest common factor of the two original numbers, a and b. This is an immediate consequence of the following simple property of highest common factors.

If $a = qb + r$, then $\text{hcf}(a, b) = \text{hcf}(b, r)$.

Why is this? As $\text{hcf}(a, b)$ divides both a and b, it also divides $a - qb$. That is, $\text{hcf}(a, b)$ divides r and, being a common factor of b and r,

$$\text{hcf}(a, b) \leq \text{hcf}(b, r).$$

But the argument works conversely as well: $\text{hcf}(b, r)$ divides both b and $qb + r$ and so is a common factor of a and b. Hence

$$\text{hcf}(b, r) \leq \text{hcf}(a, b).$$

From the two inequalities, it follows that

$$\text{hcf}(b, r) = \text{hcf}(a, b).$$

If we apply this fact to the equations in the Euclidean Algorithm system, we get

$$\text{hcf}(a, b) = \text{hcf}(b, r_1) = \text{hcf}(r_1, r_2) = \cdots = \text{hcf}(r_{n-1}, r_n) = \text{hcf}(r_n, 0).$$

Therefore,

$$\text{hcf}(a, b) = \text{hcf}(r_n, 0) = r_n.$$

The system of equations that arises in the Euclidean Algorithm does more for us than just lead to the highest common factor of numbers a and b. Proposition 4.7 tells us that $\text{hcf}(a, b)$ can be expressed as an integer combination of a and b, and we are at last in a position to see how to find this expression.

Worked Exercise 5.2

Find integers m and n such that

$$\text{hcf}(3108, 5291) = m \times 3108 + n \times 5291.$$

Solution

Consider the system of equations in Worked Exercise 5.1 that led to $\text{hcf}(3108, 5291) = 37$. We take these equations in reverse order, starting at the penultimate one, which gives the following expression for 37.

$$37 = 259 - 3 \times 74$$

The third-to-last equation gives an expression for the previous remainder, $74 = 333 - 259$, which when substituted into our expression for 37 gives the following.

$$37 = 259 - 3 \times (333 - 259) = -3 \times 333 + 4 \times 259$$

We continue to step back through the system of equations, at each stage substituting for the remainder in the developing expression for 37.

$$\begin{aligned}
37 &= -3 \times 333 + 4 \times (925 - 2 \times 333) \\
&= 4 \times 925 - 11 \times 333 \\
&= 4 \times 925 - 11 \times (2183 - 2 \times 925) \\
&= -11 \times 2183 + 26 \times 925 \\
&= -11 \times 2183 + 26 \times (3108 - 2183) \\
&= 26 \times 3108 - 37 \times 2183 \\
&= 26 \times 3108 - 37 \times (5291 - 3108) \\
&= -37 \times 5291 + 63 \times 3108
\end{aligned}$$

For ease of working with the original equations we will always write the integer combination at each stage with the larger number first.

Hence we have

$$37 = 63 \times 3108 + (-37) \times 5291,$$

and so $m = 63$ and $n = -37$ are suitable values.

It must be remarked that this is not the only solution. For example, by adding and subtracting 3108×5291 we get

$$\begin{aligned}
37 &= 63 \times 3108 + (-37) \times 5291 + 3108 \times 5291 - 3108 \times 5291 \\
&= 5354 \times 3108 + (-3145) \times 5291,
\end{aligned}$$

giving $m = 5354$ and $n = -3145$ as another of infinitely many possibilities.

Exercise 5.1

Use the Euclidean Algorithm to determine $\text{hcf}(1092, 777)$ and hence find integers m and n such that

$$\text{hcf}(1092, 777) = 1092m + 777n.$$

5.2 Linear Diophantine equations

Consider the following problem:

> I spent exactly £30 buying some 50p stamps and some 60p stamps so that the difference in number between the two kinds of stamp is as small as possible. How many stamps did I buy?

Translating the information in this problem into an equation to be solved presents little difficulty. Suppose that I bought x stamps at 50p and y stamps at 60p. Then, in all, I would have spent $50x + 60y$ pence and so

$$50x + 60y = 3000.$$

This is a linear equation in two variables and the problem is asking us to find one particular solution of it. Now in principle, for any value of x, we can solve the equation for a corresponding value of y. However, there is a major constraint in our problem: we are interested only in solutions of this equation in which the x and y values are non-negative integers, that is, for which $x \geq 0$ and $y \geq 0$. We are not going to be buying either parts of stamps or negative numbers of stamps! It is certainly not the case that for each non-negative integer x there is a corresponding non-negative integer y solving the equation, so there is more to this problem than elementary algebra.

The term **Diophantine equation** is reserved for algebraic equations for which we seek integer solutions. The Diophantine equation arising from the posed problem is one of the form $ax + by = c$ where a, b and c are integers, a and b not both zero. This latter equation is the general **linear Diophantine equation**. In Worked Exercise 5.2 and Exercise 5.1 we have witnessed some instances of the linear Diophantine equation and made use of the Euclidean Algorithm in solving them. We now build on the techniques of those examples to discover how to solve linear Diophantine equations in general.

The posed problem involves slightly more than a linear Diophantine equation since it requests solutions that are non-negative.

We will return to the posed question at the end of this section, but let us start with a linear Diophantine equation involving smaller coefficients. Consider the equation $6x + 15y = 4$. Allowing x and y to take all possible values, we see that the left-hand side of this equation takes all integer combinations of 6 and 15, which we know to be the set of all multiples of $\text{hcf}(6, 15)$. So the left-hand side takes values in the set

$$\{\ldots, -12, -9, -6, -3, 0, 3, 6, 9, 12, \ldots\}.$$

As 4 does not belong to this set, the Diophantine equation $6x + 15y = 4$ has no solutions. On the other hand, the Diophantine equation $6x + 15y = 9$ does have solutions, because 9 is in the set. But notice that this latter equation can be simplified by cancelling the common factor 3 through the equation to obtain $2x + 5y = 3$. Looked at in this simpler form, $\text{hcf}(2, 5) = 1$ and so the values taken by the left-hand side are all the multiples of 1, which certainly includes 3.

Consider now the general linear Diophantine equation $ax + by = c$. As x and y range through all integers, the values taken by the left-hand side are all the multiples of $\text{hcf}(a, b)$. So a necessary and sufficient condition for the equation to have solutions is that $\text{hcf}(a, b)$ is a factor of c.

Suppose that this condition holds. Then each of a, b and c is divisible by $\text{hcf}(a, b)$ and so, if $\text{hcf}(a, b) > 1$, we divide each coefficient by $d = \text{hcf}(a, b)$ to get the 'same' linear Diophantine equation

$$\tfrac{a}{d}x + \tfrac{b}{d}y = \tfrac{c}{d},$$

where

$$\text{hcf}\left(\tfrac{a}{d}, \tfrac{b}{d}\right) = 1.$$

By 'same' linear Diophantine equation, we mean that the two equations have the same set of solutions.

Hence, under the assumption that the linear Diophantine equation is in its lowest terms (with any common factor of a, b and c already having been cancelled out), the equation $ax + by = c$ has solutions if, and only if, $\text{hcf}(a, b) = 1$.

Let us look at an example.

Worked Exercise 5.3

Solve the linear Diophantine equation

$$66x + 40y = 6.$$

Solution

As $\text{hcf}(66, 40) = 2$, and 2 divides 6, we first cancel this common factor to get the 'same' equation

$$33x + 20y = 3.$$

As $\text{hcf}(33, 20) = 1$ this Diophantine equation is known to have solutions.

We apply the Euclidean Algorithm to the pair of coefficients 33 and 20.

$$\begin{aligned} 33 &= 1 \times 20 + 13 \\ 20 &= 1 \times 13 + 7 \\ 13 &= 1 \times 7 + 6 \\ 7 &= 1 \times 6 + 1 \\ 6 &= 6 \times 1 + 0 \end{aligned}$$

This confirms that $\text{hcf}(33, 20) = 1$.

Reversing the equations, we obtain the following.

$$\begin{aligned} 1 &= 7 - 6 \\ &= 7 - (13 - 7) \\ &= -1 \times 13 + 2 \times 7 \\ &= -1 \times 13 + 2 \times (20 - 13) \\ &= 2 \times 20 - 3 \times 13 \\ &= 2 \times 20 - 3 \times (33 - 20) \\ &= -3 \times 33 + 5 \times 20 \end{aligned}$$

This expresses 1 as an integer combination of 33 and 20, but our present task is to express 3 in this way. Multiplying through by 3, we get

$$3 = -9 \times 33 + 15 \times 20,$$

so that $x = -9$, $y = 15$ is one solution of $33x + 20y = 3$ (and therefore of $66x + 40y = 6$).

Now, what other solutions are there?

Suppose that $x = -9 + s$, $y = 15 - t$ is another solution, for some integers s and t. For this alternative solution, we could equally well have taken $y = 15 + t$, which might seem more obvious. However, this expression is chosen to conform with the usual general solution given below.

Substituting $x = -9 + s$ and $y = 15 - t$ into the equation gives

$$33(-9 + s) + 20(15 - t) = 3,$$

which simplifies to

$$33s = 20t.$$

This equation carries the implication that 20 divides $33s$, and, as $\text{hcf}(20, 33) = 1$, Euclid's Lemma gives that 20 divides s. So we can put $s = 20k$, for some integer k, and then from

$$33 \times 20k = 20t$$

we deduce that $t = 33k$.

This argument has shown that any solution must take the form

$$x = -9 + 20k, \quad y = 15 - 33k, \quad \text{for some integer } k,$$

and, moreover, substituting back into the original equation confirms that, for any integer k, this is indeed a solution. So the general solution of this linear Diophantine equation is

$$x = -9 + 20k, \quad y = 15 - 33k, \quad k \in \mathbb{Z}.$$

We have illustrated the following.

Theorem 5.4 *Solution of linear Diophantine equations*

The linear Diophantine equation $ax + by = c$ has solutions if, and only if, $\text{hcf}(a, b)$ divides c.

If this condition holds with $\text{hcf}(a, b) > 1$ then division by $\text{hcf}(a, b)$ simplifies the equation to $a'x + b'y = c'$, where $\text{hcf}(a', b') = 1$.

If x_0, y_0 is one solution of this equation then the general solution is

$$x = x_0 + b'k, \quad y = y_0 - a'k, \quad k \in \mathbb{Z}.$$

The equation is trivial when either a or b is zero, so we assume that both a and b are non-zero.

Proof The first assertion in the theorem is equivalent to Exercise 4.3(a). Simplification of a linear Diophantine equation was discussed towards the start of this subsection. So the only assertion that has not already been verified is that concerning the general solution. To that end, suppose that x', y' is any solution of the equation. Then we have

$$a'x_0 + b'y_0 = c' \quad \text{and} \quad a'x' + b'y' = c',$$

and on substituting for c' and rearranging,

$$a'(x' - x_0) = b'(y_0 - y').$$

This equation shows that a' divides $b'(y_0 - y')$ and, since $\text{hcf}(a', b') = 1$, Euclid's Lemma gives that a' divides $(y_0 - y')$. That is, $y_0 - y' = a'k$ for some integer k. Substitution of this into the above equation then gives $x' - x_0 = b'k$, and so

$$x' = x_0 + b'k, \quad y' = y_0 - a'k, \quad \text{for some integer } k,$$

as claimed.

Finally, a simple substitution shows that integers

$$x_0 + b'k \quad \text{and} \quad y_0 - a'k$$

constitute a solution of $a'x + b'y = c'$ for any integer k. ■

The two parts of the following exercise require you to find the general solution of a linear Diophantine equation but each then asks you to find, within the general solution, one particular solution.

Exercise 5.2

(a) Find the general solution of the linear Diophantine equation

$$45x - 14y = 7$$

and also the particular solution in which x takes its least positive value.

Hint: solve the equation for x and $-y$ in the usual way.

This is the question posed at the beginning of this section.

(b) I spent exactly £30 buying some 50p stamps and some 60p stamps so that the difference in number between the two kinds of stamps is as small as possible. How many stamps did I buy?

As the topic progresses you will discover that linear Diophantine equations crop up frequently, in many guises, often as subsidiary tasks in some deeper problem. When we next meet a linear Diophantine equation we will develop some improved techniques for solving them, removing the need to fall back on the Euclidean Algorithm.

Exercise 5.3

(a) Use the Euclidean Algorithm to find integers x and y satisfying

$$\text{hcf}(17, 143) = 17x + 143y.$$

Give the general solution of this equation and the particular solution in which x takes its least positive value.

(b) Find all solutions in natural numbers of the Diophantine equation

$$158x - 57y = 7.$$

(c) Find all solutions in integers of the Diophantine equation

$$39x + 27y = 3.$$

(d) Using only 2p, 10p and 50p coins, in how many ways can exactly 100 coins be used to total £6?

Learning outcomes

We list those topics on which we may set assessment questions to test your understanding of the chapter.

After working through this chapter you should:

(a) Understand arithmetic progressions arising from polygonal and pyramidal numbers, and be able to find the sums of finite such sequences.

(b) Be able to use, and understand the theoretical underpinnings of, mathematical induction.

(c) Understand and be able to apply the Generalised Principle of Mathematical Induction and the Second Principle of Mathematical Induction.

(d) Recognise the importance of the Division Algorithm, and be able to apply it in a variety of scenarios.

(e) Know the meaning of and be able to use the terms 'factor', 'highest common factor', 'integer combination' and 'coprime'; understand elementary results relating these terms.

(f) Understand the term 'least common multiple', and be able to relate it to highest common factor.

(g) Know how to apply the Euclidean Algorithm to find the highest common factor of two integers a and b, and express it as an integer combination of a and b.

(h) Understand the general meaning of the term 'Diophantine equation'.

(i) Recognise when a linear Diophantine equation has solutions, and when it does not.

(j) Be able to find particular and general solutions of given linear Diophantine equations through the application of the Euclidean Algorithm.

Solutions to exercises

Solution to Exercise 2.1

(a) The first term is $a = 3$ and the common difference is $d = 7$. As the last term is $143 = 3 + (20 \times 7)$, we see that the number of terms is $n = 21$. So the formula for an arithmetic series gives the sum as

$$\tfrac{1}{2} \times 21(6 + (20 \times 7)) = 1533.$$

(b) There are 26 terms averaging $100 + \frac{1}{2} \times 25 \times (-2) = 75$ and so the sum is $26 \times 75 = 1950$.

Solution to Exercise 2.2

(a) If k is the triangular number T_n, then $k = \frac{1}{2}n(n+1)$. Therefore

$$\begin{aligned} 9k + 1 &= \tfrac{1}{2} \times 9n(n+1) + 1 \\ &= \tfrac{1}{2}(9n^2 + 9n + 2) \\ &= \tfrac{1}{2}(3n+1)(3n+2) \\ &= \tfrac{1}{2}r(r+1), \quad \text{where } r = 3n+1, \end{aligned}$$

which is the triangular number T_r.

(b) First we show that if k is triangular then $8k + 1$ is square. So let k be the triangular number T_n, namely $k = \frac{1}{2}n(n+1)$. Therefore

$$8k + 1 = \tfrac{1}{2} \times 8n(n+1) + 1 = 4n^2 + 4n + 1 = (2n+1)^2,$$

which is a square.

Conversely, suppose that $8k + 1$ is a square so that $8k + 1 = r^2$ for some integer r. But $8k + 1$ is odd and so r must be odd (because the square of any even integer is even). So $r = 2s + 1$, for some integer s, and

$$8k + 1 = (2s+1)^2 = 4s^2 + 4s + 1 = 4s(s+1) + 1,$$

which on solving for k gives $k = \frac{1}{2}s(s+1)$, the triangular number T_s.

Solution to Exercise 2.3

(a) The nth k-gonal number is the sum of n terms of the arithmetic progression with first term $a = 1$ and common difference $d = k - 2$. So, by the formula for arithmetic series,

$$P(k, n) = \tfrac{1}{2}n(2 + (n-1)(k-2)).$$

To get the right-hand column of the table of polygonal numbers we put $k = 3, 4, 5, \ldots, 8$, giving respectively

$$\tfrac{1}{2}n(n+1),\ n^2,\ \tfrac{1}{2}n(3n-1),\ n(2n-1),\ \tfrac{1}{2}n(5n-3),\ n(3n-2).$$

To get the k-gonal numbers in the bottom row of the table we put $n = 1, 2, 3, \ldots, 8$, giving respectively

$$1,\ k,\ 3(k-1),\ 2(3k-4),\ 5(2k-3),\ 3(5k-8),\ 7(3k-5),\ 4(7k-12).$$

(b) Using the formula from part (a), we obtain the following.

$$\begin{aligned} P(k-1,n)+P(3,n-1) &= \tfrac{1}{2}n(2+(n-1)(k-3))+\tfrac{1}{2}(n-1)n \\ &= \tfrac{1}{2}n(2+(n-1)(k-3)+(n-1)) \\ &= \tfrac{1}{2}n(2+(n-1)(k-3+1)) \\ &= \tfrac{1}{2}n(2+(n-1)(k-2)) \\ &= P(k,n) \end{aligned}$$

Solution to Exercise 2.4

(a) Putting $k=5$, we have

$$Q(5,n)=Q(4,n)+Q(3,n-1).$$

That is, from Table 2.2,

$$\begin{aligned} Q(5,n) &= \tfrac{1}{6}n(n+1)(2n+1)+\tfrac{1}{6}(n-1)(n)(n+1) \\ &= \tfrac{1}{6}n(n+1)(2n+1+n-1) \\ &= \tfrac{1}{6}n(n+1)(3n) \\ &= \tfrac{1}{2}n^2(n+1). \end{aligned}$$

Putting $k=6$, we have

$$Q(6,n)=Q(5,n)+Q(3,n-1),$$

and so

$$\begin{aligned} Q(6,n) &= \tfrac{1}{2}n^2(n+1)+\tfrac{1}{6}(n-1)(n)(n+1) \\ &= \tfrac{1}{6}n(n+1)(3n+n-1) \\ &= \tfrac{1}{6}n(n+1)(4n-1). \end{aligned}$$

(b) If $2n^2 \pm 1 = m^2$ then

$$\begin{aligned} (nm)^2 &= n^2m^2 \\ &= n^2(2n^2 \pm 1) \\ &= \tfrac{1}{2}2n^2(2n^2 \pm 1) \\ &= \tfrac{1}{2}k(k+1), \quad \text{where } k=2n^2 \text{ when } 2n^2+1 \text{ is square,} \\ &\qquad\qquad \text{or } k=2n^2-1 \text{ when } 2n^2-1 \text{ is square,} \\ &= T_k. \end{aligned}$$

The first three positive values of n for which either $2n^2+1$ or $2n^2-1$ is square are:

$n=1$, $2n^2-1=1$, giving $T_1=\frac{1}{2}\times 1\times 2=1$;

$n=2$, $2n^2+1=9$, giving $T_8=\frac{1}{2}\times 8\times 9=36$;

$n=5$, $2n^2-1=49$, giving $T_{49}=\frac{1}{2}\times 49\times 50=1225$.

If you overlooked $n=1$, the next occurs at $n=12$ giving $T_{288}=41\,616$ as a triangular square.

(c) Any triangular number is of the form $\frac{1}{2}k(k+1)$. If the final digit of such a number was any of 2, 4, 7 or 9 then, by doubling, the final digit of $k(k+1)$ would be 4, 8, 4 or 8 respectively. Now k can take any positive value so its final digit can be any of the ten digits. In each case we can determine the final digit of $k(k+1)$, as follows.

Final digit of k	Final digit of $k+1$	Final digit of $k(k+1)$
0	1	0
1	2	2
2	3	6
3	4	2
4	5	0
5	6	0
6	7	2
7	8	6
8	9	2
9	0	0

This confirms that the final digit of $k(k+1)$ cannot be 4 or 8 and so the final digit of $\frac{1}{2}k(k+1)$ cannot be 2, 4, 7 or 9.

Solution to Exercise 3.1

(a) (i) The set $\{n \in \mathbb{Z} : |n| > 20\}$ consists of all the integers except those from -20 to 20 inclusive. That is,

$$\{\ldots, -23, -22, -21, 21, 22, 23, \ldots\}.$$

(ii) The set $\{n : n = 2m^2, \text{ for some } m \in \mathbb{N}\}$ is the set consisting of those integers that are equal to twice the square of a positive integer, namely

$$\{2, 8, 18, 32, \ldots\}.$$

(b) (i) The set is $\{n : n = 2m - 1, \text{ for some } m \in \mathbb{N}\}$.

We could equally well write this set as $\{n : n = 2m + 1, \text{ for some } m \in \mathbb{Z} \text{ with } m \geq 0\}$.

(ii) The set is $\{n \in \mathbb{Z} : |n| \leq 100\}$.

Solution to Exercise 3.2

(a) Let $P(n)$ be the proposition that the formula is true for the natural number n.

Then $P(1)$ is true since

$$1^2 = \tfrac{1}{6} \times 1(1+1)(2+1).$$

Hence we have the basis for the induction.

For the induction step, assume that $P(k)$ is true, that is,

$$1^2 + 2^2 + 3^2 + \cdots + k^2 = \tfrac{1}{6}k(k+1)(2k+1).$$

Then

$$\begin{aligned}
&1^2 + 2^2 + 3^2 + \cdots + k^2 + (k+1)^2, \quad \text{the LHS of } P(k+1),\\
&\quad = \tfrac{1}{6}k(k+1)(2k+1) + (k+1)^2, \quad \text{by the induction hypothesis,}\\
&\quad = \tfrac{1}{6}(k+1)\left(2k^2 + k + 6(k+1)\right)\\
&\quad = \tfrac{1}{6}(k+1)\left(2k^2 + 7k + 6\right)\\
&\quad = \tfrac{1}{6}(k+1)(k+2)(2k+3)\\
&\quad = \tfrac{1}{6}(k+1)((k+1)+1)(2(k+1)+1), \quad \text{the RHS of } P(k+1).
\end{aligned}$$

This shows that $P(k+1)$ is true and completes the induction step.

Hence, by mathematical induction, the formula is true for all natural numbers n.

(b) Let $P(n)$ be the proposition that the formula is true for n.

Then $P(1)$ is true since

$$a = \frac{a(r-1)}{r-1}.$$

So we have the basis for the induction.

Now for the induction step. Assume that $P(k)$ is true; that is,

$$a + ar + ar^2 + ar^3 + \cdots + ar^{k-1} = \frac{a(r^k - 1)}{r-1}.$$

Then

$$\begin{aligned}
&a + ar + ar^2 + ar^3 + \cdots + ar^{k-1} + ar^k\\
&\quad = \frac{a(r^k-1)}{r-1} + ar^k, \quad \text{by the induction hypothesis,}\\
&\quad = \frac{a(r^k-1) + ar^k(r-1)}{r-1}\\
&\quad = \frac{a(r^{k+1}-1)}{r-1},
\end{aligned}$$

which is $P(k+1)$, completing the induction step.

Hence, by mathematical induction, the formula is true for all natural numbers n.

Solution to Exercise 3.3

Let $P(n)$ be the proposition that $L_n < \left(\frac{7}{4}\right)^n$.

We note that $P(1)$ and $P(2)$ are true since

$$L_1 = 1 < \tfrac{7}{4} \quad \text{and} \quad L_2 = 3 < \left(\tfrac{7}{4}\right)^2 = \tfrac{49}{16}.$$

So we have the basis for the induction.

For the induction step (using the Second Principle) suppose that $P(1)$, $P(2)$, …, $P(k)$ are all true and consider $P(k+1)$ for $k \geq 2$.

$$\begin{aligned} L_{k+1} &= L_k + L_{k-1} \\ &< \left(\tfrac{7}{4}\right)^k + \left(\tfrac{7}{4}\right)^{k-1}, \quad \text{by the induction hypothesis,} \\ &= \left(\tfrac{7}{4}\right)^{k-1}\left(\tfrac{7}{4}+1\right) \\ &= \left(\tfrac{7}{4}\right)^{k-1}\left(\tfrac{11}{4}\right) \\ &< \left(\tfrac{7}{4}\right)^{k-1}\left(\tfrac{7}{4}\right)^2, \quad \text{since } \tfrac{11}{4} < \tfrac{49}{16}, \\ &= \left(\tfrac{7}{4}\right)^{k+1}, \end{aligned}$$

which shows that $P(k+1)$ is true.

Hence, by the Second Principle of Mathematical Induction, the proposition is true.

Note that, since the proof of the induction step uses the formula $L_{k+1} = L_k + L_{k-1}$, which holds only for $k \geq 2$, our induction step first deduces the truth of $P(3)$ from that of $P(1)$ and $P(2)$, and then goes on to deduce the truth of $P(4)$, $P(5)$, etc. It does not deduce the truth of $P(2)$ from $P(1)$. Hence we have to show the truth of $P(2)$ separately, which we have done by including it in the basis for the induction.

Solution to Exercise 3.4

(a) (i) Let $P(n)$ be the proposition that the formula is true for n.

$P(1)$ is true, since

$$1^2 = \tfrac{1}{3} \times 1 \times (4 \times 1^2 - 1).$$

This gives the basis for the induction.

Assume that $P(k)$ is true for some natural number k; that is,

$$1^2 + 3^2 + 5^2 + \cdots + (2k-1)^2 = \tfrac{1}{3}k(4k^2-1).$$

Then

$$\begin{aligned} 1^2 + 3^2 + 5^2 + \cdots + (2k-1)^2 + (2k+1)^2, &\quad \text{the LHS of } P(k+1), \\ = \tfrac{1}{3}k(4k^2-1) + (2k+1)^2, &\quad \text{by the induction hypothesis,} \\ = \tfrac{1}{3}k(2k-1)(2k+1) + (2k+1)^2 \\ = \tfrac{1}{3}(2k+1)[k(2k-1) + 3(2k+1)] \\ = \tfrac{1}{3}(2k+1)(2k^2+5k+3) \\ = \tfrac{1}{3}(2k+1)(2k+3)(k+1) \\ = \tfrac{1}{3}(k+1)(4k^2+8k+3) \\ = \tfrac{1}{3}(k+1)(4(k+1)^2-1), &\quad \text{the RHS of } P(k+1), \end{aligned}$$

which establishes the truth of $P(k+1)$, and completes the induction step.

Hence, by the Principle of Mathematical Induction, the formula is true for all natural numbers.

(ii) Let $P(n)$ be the proposition that the formula is true for n.

Then $P(1)$ is true since

$$1 = \tfrac{1}{2} \times 1^2 \times (1+1).$$

Assume that $P(k)$ is true; that is,

$$1 + 5 + 12 + 22 + \cdots + \tfrac{1}{2}k(3k-1) = \tfrac{1}{2}k^2(k+1).$$

Then

$$\begin{aligned} 1 + 5 + 12 + 22 + \cdots + \tfrac{1}{2}k(3k-1) + \tfrac{1}{2}(k+1)(3k+2) \\ &= \tfrac{1}{2}k^2(k+1) + \tfrac{1}{2}(k+1)(3k+2) \\ &= \tfrac{1}{2}(k+1)(k^2+3k+2) \\ &= \tfrac{1}{2}(k+1)^2(k+2), \end{aligned}$$

confirming that $P(k+1)$ is true and completing the induction step.

The truth of $P(n)$ for all $n \geq 1$ follows by mathematical induction.

(b) Let $P(n)$ be the proposition that the formula is true for n.

$P(1)$ is certainly true since

$$1 \times (1!) = 2! - 1,$$

both sides being equal to 1.

For the induction step, suppose that $P(k)$ is true for some $k \geq 1$. That is,

$$1 \times (1!) + 2 \times (2!) + \cdots + k \times (k!) = (k+1)! - 1.$$

Then

$$\begin{aligned} 1 \times (1!) + 2 \times (2!) + \cdots + k \times (k!) + (k+1) \times ((k+1)!) \\ &= (k+1)! - 1 + (k+1)((k+1)!) \\ &= (k+1)!(1+k+1) - 1 \\ &= (k+2)! - 1, \end{aligned}$$

which confirms the truth of $P(k+1)$ and completes the induction step.

The truth of $P(n)$ for all $n \geq 1$ follows by mathematical induction.

(c) We observe that $n!$ increases more rapidly than $6n^2$ and $n!$ exceeds $6n^2$ for the first time when $n = 6$. So we make the conjecture that $n! > 6n^2$ for all $n \geq 6$, and prove it by mathematical induction.

Let $P(n)$ be the proposition that $n! > 6n^2$. Now

$$6! = 720 > 6 \times 6^2 = 216,$$

so $P(6)$ is true, giving us the basis for induction.

Suppose that $P(k)$ is true for some integer $k \geq 6$, that is,

$$k! > 6k^2.$$

On multiplying through by $k+1$, we get

$$(k+1)! > 6k^2(k+1).$$

If we can show that $k^2 \geq k+1$ it will follow that $(k+1)! > 6(k+1)^2$, completing the induction step. But this is easily seen since, for $k \geq 6$,

$$k^2 \geq 6k > 2k = k + k > k + 1,$$

which completes the induction step.

So by the Generalised Principle of Mathematical Induction, $P(n)$ is true for all $n \geq 6$.

(d) Suppose that

$$S = \{k \in \mathbb{Z} : k \geq n_0 \text{ and } P(k) \text{ is false}\}$$

is non-empty. We cannot apply the Well-Ordering Principle directly to this set, as it may not be a subset of $\mathbb{N}$.

Now consider the set T defined by

$$T = \{s - n_0 : s \in S\}.$$

Certainly T is a non-empty set of integers, each of whose elements is greater than or equal to zero. Now, by condition (a), $P(n_0)$ is true so n_0 is not an element of S. Hence zero is not an element of T, and T is a non-empty subset of $\mathbb{N}$.

By the Well-Ordering Principle, T has a least element. We may take this least element of T to be $m - n_0$ for some integer m in S and, by the definition of T, it follows that m is the least element of S.

Furthermore, since $n_0 \notin S$ we know that $m > n_0$, which ensures that the list $P(n_0)$, $P(n_0+1)$, $\ldots$, $P(m-1)$ contains at least one element. By the definition of S it follows that $P(n_0), P(n_0+1), \ldots, P(m-1)$ must all be true. But then condition (b′) gives that $P(m)$ is true, which contradicts the fact that $m \in S$.

Hence the only assumption made, namely that S is non-empty, must be false. Hence S is empty and so the proposition $P(k)$ is true for all $k \geq n_0$.

Solution to Exercise 4.1

Cubing each of the three given forms that the number can take, and then writing each cube in the form $9n + r$, where $0 \leq r < 9$, gives

$$(3n)^3 = 27n^3 = 9(3n^3) + 0,$$
$$(3n+1)^3 = 27n^3 + 27n^2 + 9n + 1 = 9(3n^3 + 3n^2 + n) + 1,$$
$$(3n+2)^3 = 27n^3 + 54n^2 + 36n + 8 = 9(3n^3 + 6n^2 + 4n) + 8.$$

Hence the only possible remainders are 0, 1 and 8.

Solution to Exercise 4.2

(a) The factors of 18 are 1, 2, 3, 6, 9 and 18. The factors of -24 are 1, 2, 3, 4, 6, 8, 12 and 24. (These are the same as the factors of 24.)

The common factors of 18 and -24 are 1, 2, 3 and 6.

(b) The Division Algorithm tells us that n takes one of the three forms $3k$, $3k+1$ or $3k+2$. But n is not a multiple of 3 and so the first of these forms cannot occur.

For $n = 3k+1$ we get

$$(3k+1)^2 - 1 = 9k^2 + 6k + 1 - 1 = 3(3k^2 + 2k).$$

For $n = 3k+2$ we get

$$(3k+2)^2 - 1 = 9k^2 + 12k + 4 - 1 = 3(3k^2 + 4k + 1).$$

In each case $n^2 - 1$ has remainder zero on dividing by 3.

Solution to Exercise 4.3

(a) First let m and n be integers such that $\text{hcf}(a,b) = ma + nb$. Then, if c is a multiple of $\text{hcf}(a,b)$, say $c = k \times \text{hcf}(a,b)$, we have

$$c = k(ma + nb) = (km)a + (kn)b,$$

showing that c is an integer combination of a and b.

Remember that 'if, and only if' proofs have two parts.

Conversely, if there are integers x and y such that $c = xa + yb$ then, as $\text{hcf}(a,b)$ divides both a and b, property (e) of Theorem 4.4 confirms that $\text{hcf}(a,b)$ divides $c = xa + yb$.

(b) Let $d = \text{hcf}(a,b)$. Then, from $d = ma + nb$, we have $1 = ma' + nb'$, where $a = da'$ and $b = db'$. It follows from Lemma 4.9 that $\text{hcf}(m,n) = 1$.

(c) From $\text{hcf}(a,b) = 1$ and $\text{hcf}(a,c) = 1$ we infer the existence of integers m, n, r and s such that

$$ma + nb = 1 \quad \text{and} \quad ra + sc = 1.$$

Multiplying these two equations together, we get

$$(ma + nb)(ra + sc) = 1,$$

and collating the terms appropriately we get

$$(mar + nbr + msc)a + (ns)bc = 1,$$

which expresses 1 as an integer combination of a and bc. The result then follows from Lemma 4.9.

(d) There exist integers m, n, r and s such that

$$\text{hcf}(a,b) = ma + nb \quad \text{and} \quad \text{hcf}(ka,kb) = r(ka) + s(kb).$$

But then

$$\begin{aligned} \text{hcf}(ka,kb) &= k(ra + sb) \\ &= k(t \times \text{hcf}(a,b)), \quad \text{for some integer } t, \text{ by part (a) above,} \\ &= t(k \times \text{hcf}(a,b)). \end{aligned}$$

In addition we have

$$\begin{aligned} k \times \text{hcf}(a,b) &= k(ma + nb) \\ &= m(ka) + n(kb) \\ &= u \times \text{hcf}(ka,kb), \quad \text{for some integer } u, \text{ by part (a).} \end{aligned}$$

We now know that each of $\text{hcf}(ka, kb)$ and $k \times \text{hcf}(a, b)$ is a multiple of the other, and so the equality follows from property (c) of Theorem 4.4.

Solution to Exercise 4.4

(a) As $\text{hcf}(a, b) = 1$ there exist integers m and n such that

$$ma + nb = 1.$$

(i) If c divides a then $a = kc$ for some integer k. Substituting for this value of a gives

$$mkc + nb = 1$$

and, as this expresses 1 as an integer combination of b and c, Lemma 4.9 gives $\text{hcf}(b, c) = 1$.

(ii) Our goal is to express 1 as an integer combination of a and $a + b$. This is achieved from the equation $ma + nb = 1$ by careful rearrangement:

$$ma + nb = (m - n)a + n(a + b) = 1.$$

(b) (i) If d divides $3n + 2$ and d divides $5n + 3$ then d divides $5(3n + 2) + (-3)(5n + 3)$. That is, d divides 1. It follows that the only common factor of $3n + 2$ and $5n + 3$ is 1.

(ii) Let $\text{hcf}(n^2 + 3n + 1, 2n + 1) = d$. Then d divides any integer combination of $n^2 + 3n + 1$ and $2n + 1$. We first choose one that eliminates the term in n^2, namely

$$2(n^2 + 3n + 1) + (-n)(2n + 1) = 5n + 2.$$

So d divides $5n + 2$ and d divides $2n + 1$. Therefore, as before, d divides any integer combination of $5n + 2$ and $2n + 1$. This time we choose one that eliminates the term in n, namely

$$5(2n + 1) + (-2)(5n + 2) = 1.$$

We conclude that $d = 1$.

Solution to Exercise 4.5

Let $\text{lcm}(a, b) = l$ and let m be any common multiple of a and b. Using the Division Algorithm for dividing m by l, there exist integers q and r such that

$$m = lq + r, \quad \text{where } 0 \leq r < l.$$

Rearranging gives

$$r = 1m + (-q)l$$

and, as both m and l are common multiples of a and b, so too is the integer combination r. But l is the least common multiple of a and b, and so we cannot have $0 < r < l$.

Hence $r = 0$, $m = ql$ and so l divides m.

Solution to Exercise 4.6

(a) Since $\text{hcf}(18, 24) = 6$, $\text{lcm}(18, 24) = \dfrac{18 \times 24}{6} = 72$.

(b) The factors of 27 are 1, 3, 9 and 27 and the largest of these that is also a factor of 45 is 9. Therefore $\text{hcf}(27, 45) = 9$ and $\text{lcm}(27, 45) = \dfrac{27 \times 45}{9} = 135$.

(c) Any common factor of 39 and 40 must divide $40 - 39$. It follows that $\text{hcf}(39, 40) = 1$ and $\text{lcm}(39, 40) = 39 \times 40 = 1560$.

Note that this example illustrates the general result that, for natural numbers a and b,

$\text{lcm}(a, b) = ab$ if, and only if, $\text{hcf}(a, b) = 1$.

Solution to Exercise 4.7

(a) We have

$$12k + 8 = 3(4k + 2) + 2 = 3m + 2,$$

where $m = 4k + 2$ is an integer.

The converse is false since, for example, $5 = 3 \times 1 + 2$ is of the form $3m + 2$ but has remainder 5 on division by 12.

Since 3 divides 12, the remainder when $12k + r$ is divided by 3 is the same as the remainder when r is divided by 3. This will be equal to 2 when $r = 2$, 5, 8 or 11.

(b) Any number is necessarily of one of the forms $7k + r$, where $r = 0$, 1, 2, 3, 4, 5 or 6. Consider the resulting forms for $(7k + r)^2$ and $(7k + r)^3$. Note that

$$(7k + r)^2 = 49k^2 + 14kr + r^2 = 7(7k^2 + 2kr) + r^2$$

and

$$(7k + r)^3 = 343k^3 + 147k^2r + 21kr^2 + r^3 = 7(49k^3 + 21k^2r + 3kr^2) + r^3.$$

Therefore the remainder on dividing $(7k + r)^2$ by 7 is the same as the remainder on dividing r^2 by 7. Similarly, the remainder on dividing $(7k + r)^3$ by 7 is the same as the remainder on dividing r^3 by 7.

$0^2 = 0 = 7 \times 0 + 0$	$0^3 = 0 = 7 \times 0 + 0$
$1^2 = 1 = 7 \times 0 + 1$	$1^3 = 1 = 7 \times 0 + 1$
$2^2 = 4 = 7 \times 0 + 4$	$2^3 = 8 = 7 \times 1 + 1$
$3^2 = 9 = 7 \times 1 + 2$	$3^3 = 27 = 7 \times 3 + 6$
$4^2 = 16 = 7 \times 2 + 2$	$4^3 = 64 = 7 \times 9 + 1$
$5^2 = 25 = 7 \times 3 + 4$	$5^3 = 125 = 7 \times 17 + 6$
$6^2 = 36 = 7 \times 5 + 1$	$6^3 = 216 = 7 \times 30 + 6$

From the above, we see that the possible remainders on dividing a square by 7 are 0, 1, 2 and 4 and that the possible remainders on dividing a cube by 7 are 0, 1 and 6. So if a number is both a square and a cube then its remainder on division by 7 must be either 0 or 1.

(c) Let $P(n)$ be the statement that 9 divides $4^n + 6n - 1$.

When $n = 0$, $4^n + 6n - 1 = 0$ and, as this is a multiple of 9, $P(0)$ is true. This gives the basis for induction.

Suppose that $P(k)$ is true for some integer $k \geq 0$; that is, $4^k + 6k - 1 = 9m$ for some integer m. Then

$$\begin{aligned} 4^{k+1} + 6(k+1) - 1 &= 4 \times 4^k + 6(k+1) - 1 \\ &= 4(9m - 6k + 1) + 6(k+1) - 1 \\ &= 36m - 18k + 9 \\ &= 9(4m - 2k + 1), \end{aligned}$$

showing that 9 divides $4^{k+1} + 6(k+1) - 1$, which verifies $P(k+1)$ and completes the induction step.

The result follows by mathematical induction.

(d) The general result illustrated is that 36 divides $10^n + 8$ for all $n \geq 2$.

The result is true for $n = 2$ since $108 = 3 \times 36$, which is the basis for the induction.

For the induction step suppose that 36 divides $10^k + 8$, for some integer $k \geq 2$. To be precise, suppose that $10^k + 8 = 36r$ for some integer r. Then

$$\begin{aligned} 10^{k+1} + 8 &= 10 \times 10^k + 8 \\ &= 10(36r - 8) + 8 \\ &= 36(10r) - 72 \\ &= 36(10r - 2), \end{aligned}$$

showing that $10^{k+1} + 8$ is divisible by 36, as required.

Hence, by mathematical induction, 36 divides $10^n + 8$ for all integers $n \geq 2$.

(e) There exist integers m and n such that $am + bn = 1$. Multiplying by c gives

$$m(ac) + (cn)b = c.$$

This expresses c as an integer combination of ac and b and so, by Exercise 4.3(a), c is a multiple of $\text{hcf}(ac, b)$. But, by definition, b is a multiple of $\text{hcf}(ac, b)$ and so $\text{hcf}(ac, b)$ is a common factor of b and c and, in particular,

$$\text{hcf}(ac, b) \leq \text{hcf}(b, c).$$

On the other hand, $\text{hcf}(b, c)$ divides b and $\text{hcf}(b, c)$ divides ac (since it divides c), and so

$$\text{hcf}(b, c) \leq \text{hcf}(ac, b).$$

The equality follows.

(f) (i) Let $\text{hcf}(2a+b, a+b) = d$. Then, using property (e) of Theorem 4.4,

$$d \mid ((2a+b)-(a+b)), \text{ that is, } d \mid a$$

and

$$d \mid ((-1)(2a+b)+2(a+b)), \text{ that is, } d \mid b.$$

It follows that d divides $\text{hcf}(a,b)$, and so $d = 1$.

(ii) Let $\text{hcf}(a+b, a-b) = d$. Then, using property (e) of Theorem 4.4,

$$d \mid ((a+b)+(a-b)), \text{ that is, } d \mid 2a$$

and

$$d \mid ((a+b)-(a-b)), \text{ that is, } d \mid 2b.$$

Hence

$$\begin{aligned} d \leq \text{hcf}(2a, 2b) &= 2 \times \text{hcf}(a,b), \quad \text{by Exercise 4.3(d)}, \\ &= 2, \end{aligned}$$

so $d = 1$ or 2 as required.

Solution to Exercise 5.1

$$\begin{aligned} 1092 &= 1 \times 777 + 315 \\ 777 &= 2 \times 315 + 147 \\ 315 &= 2 \times 147 + 21 \\ 147 &= 7 \times 21 + 0 \end{aligned}$$

So $\text{hcf}(1092, 777)$ is equal to the last non-zero remainder, namely 21.

Reversing the equations, we get the following.

$$\begin{aligned} \text{hcf}(1092, 777) &= 21 \\ &= 315 - 2 \times 147 \\ &= 315 - 2 \times (777 - 2 \times 315) \\ &= -2 \times 777 + 5 \times 315 \\ &= -2 \times 777 + 5 \times (1092 - 777) \\ &= 5 \times 1092 - 7 \times 777 \end{aligned}$$

The required integers are $m = 5$ and $n = -7$.

Solution to Exercise 5.2

(a) The negative coefficient causes no problem if we think of the equation as being $45x + 14(-y) = 7$ and solve for x and $-y$. The Euclidean Algorithm yields the following equations.

$$\begin{aligned} 45 &= 3 \times 14 + 3 \\ 14 &= 4 \times 3 + 2 \\ 3 &= 1 \times 2 + 1 \\ 2 &= 2 \times 1 + 0 \end{aligned}$$

Reversing the equations, we have

$$\begin{aligned} 1 &= 3 - 2 \\ &= 3 - (14 - 4 \times 3) \\ &= -1 \times 14 + 5 \times 3 \\ &= -1 \times 14 + 5 \times (45 - 3 \times 14) \\ &= 5 \times 45 - 16 \times 14. \end{aligned}$$

Multiplying by 7,

$$7 = 35 \times 45 - 112 \times 14,$$

so that $x = 35$, $-y = -112$ (that is, $y = 112$) is one solution of the equation.

The general solution is

$$x = 35 + 14k, \quad y = 112 + 45k, \quad k \in \mathbb{Z}.$$

The expression for y comes from our previous general solution, which gave $-y = -112 - 45k$. If you substitute back into the original equation, note that these 'added' terms do cancel.

The solution with the least positive value of x is given by $k = -2$, namely $x = 7$, $y = 22$.

Note that the general solution can therefore be written as

$$x = 7 + 14k, \quad y = 22 + 45k, \quad k \in \mathbb{Z}.$$

(b) If x is the number of 50p stamps bought and y is the number of 60p stamps then the total amount spent is $50x + 60y$ pence. Hence $50x + 60y = 3000$, and, as we are not allowed fractions of stamps, our task is to solve this in integers.

We first note that $\text{hcf}(50, 60) = 10$, and, since 10 divides 3000, there are solutions. Dividing the whole equation by 10 gives $5x + 6y = 300$. It is not difficult to 'spot' one solution of this equation. As 5 divides 300 we see that $x = 60$, $y = 0$ is a solution. Theorem 5.4 then gives the general solution:

You can, of course, use the Euclidean Algorithm, but do look out for shortcuts.

$$x = 60 - 6k, \quad y = 0 + 5k, \quad k \in \mathbb{Z}.$$

We seek the particular solution in which $|x - y|$ is as small as possible; that is, we want to minimise $|60 - 11k|$. But $11k$ is nearest to 60 when $k = 5$, giving $x = 30$, $y = 25$ as the required solution.

Solution to Exercise 5.3

(a) Since we can write

$$\begin{aligned} 143 &= 8 \times 17 + 7 \\ 17 &= 2 \times 7 + 3 \\ 7 &= 2 \times 3 + 1 \\ 3 &= 3 \times 1 + 0, \end{aligned}$$

it follows that

$$\begin{aligned} \text{hcf}(143, 17) &= 1 \\ &= 7 - 2 \times 3 \\ &= 7 - 2(17 - 2 \times 7) \\ &= -2 \times 17 + 5 \times 7 \\ &= -2 \times 17 + 5(143 - 8 \times 17) \\ &= 5 \times 143 - 42 \times 17. \end{aligned}$$

Therefore $x = -42$, $y = 5$ is one solution. The general solution is therefore

$$x = -42 + 143k, \quad y = 5 - 17k, \quad k \text{ any integer.}$$

The particular solution in which x takes its least positive value occurs when $k = 1$, and is $x = 101$, $y = -12$.

(b) Since we can write

$$\begin{aligned} 158 &= 2 \times 57 + 44 \\ 57 &= 1 \times 44 + 13 \\ 44 &= 3 \times 13 + 5 \\ 13 &= 2 \times 5 + 3 \\ 5 &= 1 \times 3 + 2 \\ 3 &= 1 \times 2 + 1 \\ 2 &= 2 \times 1 + 0, \end{aligned}$$

it follows that

$$\begin{aligned} \text{hcf}(158, 57) &= 1 \\ &= 3 - 1 \times 2 \\ &= 3 - (5 - 1 \times 3) \\ &= -1 \times 5 + 2 \times 3 \\ &= -1 \times 5 + 2(13 - 2 \times 5) \\ &= 2 \times 13 - 5 \times 5 \\ &= 2 \times 13 - 5(44 - 3 \times 13) \\ &= -5 \times 44 + 17 \times 13 \\ &= -5 \times 44 + 17(57 - 44) \\ &= 17 \times 57 - 22 \times 44 \\ &= 17 \times 57 - 22(158 - 2 \times 57) \\ &= -22 \times 158 + 61 \times 57. \end{aligned}$$

Therefore $x = -22$, $y = -61$ is one solution of the Diophantine equation $158x - 57y = 1$ and, multiplying by 7, $x = -154$, $y = -427$ is one solution of the equation $158x - 57y = 7$.

The general solution of the latter equation is therefore given by

$$x = -154 + 57k, \quad y = -427 + 158k, \quad k \text{ any integer.}$$

When $k \geq 3$, x and y are both positive, the least positive solution being $x = 17$, $y = 47$. The set of all positive solutions is

$$x = 17 + 57s, \quad y = 47 + 158s, \quad s \geq 0.$$

(c) First notice that we can divide the equation $39x + 27y = 3$ through by 3, which gives us

$$13x + 9y = 1.$$

We now use the Euclidean Algorithm on 13 and 9:

$$\begin{aligned} 13 &= 1 \times 9 + 4 \\ 9 &= 2 \times 4 + 1 \\ 4 &= 4 \times 1 + 0. \end{aligned}$$

Thus, we have

$$\begin{aligned} \text{hcf}(9, 13) &= 1 \\ &= 9 - 2 \times 4 \\ &= 9 - 2 \times (13 - 9) \\ &= 3 \times 9 - 2 \times 13. \end{aligned}$$

Therefore $x = -2$ and $y = 3$ is one solution of the Diophantine equation $13x + 9y = 1$, and therefore also one solution of the original Diophantine equation $39x + 27y = 3$.

Applying Theorem 5.4, the general solution to $39x + 27y = 3$ is therefore

$$x = -2 + 9k, \quad y = 3 - 13k, \quad k \text{ any integer.}$$

(d) Suppose that x 2p coins, y 10p coins and z 50p coins are used. Then, since the total value is £6,

$$2x + 10y + 50z = 600.$$

Furthermore, as the total number of coins used is 100,

$$x + y + z = 100.$$

Substituting $2x = 200 - 2y - 2z$ into the former equation gives $200 + 8y + 48z = 600$, which simplifies to

$$y + 6z = 50.$$

One solution of this Diophantine equation may be spotted immediately, namely $y = 50$, $z = 0$. Hence the general solution is

$$y = 50 - 6k, \quad z = k, \quad k \in \mathbb{Z}.$$

Now as $x = 100 - y - z$, we seek non-negative solutions from

$$x = 50 + 5k, \quad y = 50 - 6k, \quad z = k, \quad k \in \mathbb{Z}.$$

Choosing $k \geq 0$ guarantees that x and z are non-negative and $k \leq 8$ keeps y non-negative. Hence there are nine solutions corresponding to $k = 0$ to 8 inclusive. These are given in the following table.

k	x	y	z
0	50	50	0
1	55	44	1
2	60	38	2
3	65	32	3
4	70	26	4
5	75	20	5
6	80	14	6
7	85	8	7
8	90	2	8