Contents

Preface

The text of these notes is mainly an edited version of the first half of a subject guide I wrote for the University of London International Programmes.

I thank Michele Harvey for her help with the material on complex numbers. I am grateful to Keith Martin, Jan van den Heuvel and Amol Sasane for carefully reading a draft of the subject guide and for suggesting ways in which to improve it.

These notes also incorporate materials written in previous years by Jan van den Heuvel and Graham Brightwell and I am grateful to them for allowing me to use these.

Thanks also to Mark Baltovic for help with typesetting.

Chapter 1
Introduction

In this very brief introduction, I aim to give you an idea of the nature of this subject and to advise on how best to approach it. I also give general information about these notes and recommended reading.

1.1 This course

In **Introduction to Abstract Mathematics** the emphasis is on theory rather than method: we will want to understand why certain techniques work, and how we might be able to prove that they do, for example. The main central topic in this course is *proof*. This course is an introduction to formal mathematical reasoning, in which proof is central. We will meet the fundamental concepts and constructions of mathematics and see how to formulate mathematical statements in precise terms, and we will see how such statements can be proved or disproved.

In this course, we need to work with *precise definitions* and *statements*, and you will need to know these. Not only will you need to know these, but you will have to understand them, and be able (through the use of them) to demonstrate that you understand them. Simply learning the definitions without understanding what they mean is not going to be adequate. I hope that these words of warning don't discourage you, but I think it's important to make it clear that this is a subject at a higher conceptual level than most of the mathematics you are likely to have studied before.

In this course, you will learn how to *prove* mathematical statements precisely. This is a very different sort of mathematics from that which you will have encountered in many other mathematics courses you have previously taken, where the emphasis is on solving problems through *calculation*. In Abstract Mathematics, one has to be able to produce convincing mathematical arguments as to why a given mathematical statement is true or false. For example, a prime number is a positive integer greater than 1 that is only divisible by itself and the number 1 (so 7 is a prime number, but 8 is not). The statement "There are infinitely many prime numbers" is a mathematical statement, and it is either true (there are infinitely many prime numbers) or false (there are only a finitely many prime numbers). In fact, the statement is true. But why? There's no quick 'calculation' we can do to establish the truth of the statement. What is needed is a proof: a watertight, logical argument. This is the type of problem we consider in this course.

1.1.1 Aims

The course is designed to enable you to:

- develop your ability to think in a critical manner;
- formulate and develop mathematical arguments in a logical manner;
- improve your skill in acquiring new understanding and expertise;
- acquire an understanding of basic abstract mathematics, and the role of logical argument in mathematics.

1.1.2 Learning objectives

Having taken this subject, you should:

- have a knowledge of basic mathematical concepts in discrete mathematics, algebra, and real analysis;
- be able to use formal notation correctly and in connection with precise statements in English;
- be able to solve mathematical problems in discrete mathematics, algebra and real analysis;
- be able to find and formulate simple proofs.

1.1.3 Topics covered (first half of the course)

Descriptions of topics to be covered appear in the relevant chapters. However, it is useful to give a brief overview at this stage. These notes cover the first half of this course. Here, we are concerned primarily with proof, logic, and number systems. We will first investigate how precise mathematical statements can be formulated, and here we will use the language and symbols of mathematical logic. We will then study how one can prove or disprove mathematical statements. Next, we look at some important ideas connected with functions, relations, and numbers. For example, we will look at prime numbers and learn what special properties these important numbers have, and how one may prove such properties.

1.2 Moodle

All information and materials for this course are on Moodle:

`http://moodle.lse.ac.uk/course/view.php?id=1989`

On the course Moodle page, you will find assignments, solutions, lecture notes, and so on.

1.3 Reading

There are many books that would be useful for this subject, since abstract mathematics is a components of almost all university-level mathematics degree programmes.

For the first half of the course (the part covered by these notes), the following two books are recommended.

☞ Biggs, Norman L., *Discrete Mathematics*, Second edition. (Oxford University Press, 2002). [ISBN 0198507178].

☞ Eccles, P.J., *An Introduction to Mathematical Reasoning: numbers, sets and functions*. (Cambridge University Press, 1997). [ISBN 0521597188].

There is one topic that neither of these covers, which is the topic of Complex Numbers. However, this is a topic that is well-covered in a number of other textbooks and I have included a fairly full treatment of it in these notes to compensate for the fact that it is not covered in these two recommended textbooks.

1.4 Activities and sample exercises

Throughout the chapters of these notes, you'll find 'activities'. These are things for you to do or think about as you read, just to reaffirm that you've understood the material.

At the end of each chapter of these notes you will find some sample exercises together with solutions. These are not the exercises that will be assigned for classes, but are *additional* to those. They are a very useful resource. You should try them once you think you have mastered a particular chapter. Really try them: don't just simply read the solutions provided. Make a serious attempt before consulting the solutions. Note that the solutions are often just sketch solutions, to indicate to you how to answer the questions.

Chapter 2

Mathematical statements, proof, logic and sets

☞ Biggs, N.L. *Discrete Mathematics*. Chapters 1–3.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning*. Chapters 1–4 and 6.

2.1 Introduction

In this important chapter, we set the ground for much of what follows in this course. Abstract Mathematics is about making precise mathematical statements and establishing, by proof or disproof, whether these statements are true or false. In this chapter we look at what this means, concentrating on fairly simple types of mathematical statement, in order to emphasise techniques of proof. In later chapters (such as those on numbers, analysis and algebra) we will use these proof techniques extensively. You might think that some of the things we prove in this chapter are very obvious and hardly merit proving, but proving even 'obvious' statements can be quite tricky sometimes, and it is good preparation for proving more complicated things later.

2.2 Mathematical statements and proof

To introduce the topics of mathematical statement and proof, we start by giving some explicit examples. Later in the chapter we give some general theory and principles. Our discussion of the general theory is limited because this is not a course in logic, as such. What we do need is enough logic to understand what mathematical statements mean and how we might prove or disprove them.

2.2.1 Examples of Mathematical Statements

Consider the following statements (in which, you should recall that the natural numbers are the positive integers):

(a) 20 is divisible by 4.

(b) 21 is not divisible by 7.

(c) 21 is divisible by 4.

(d) 21 is divisible by 3 or 5.

(e) 50 is divisible by 2 and 5.

(f) n^2 is even.

(g) For every natural number n, the number $n^2 + n$ is even.

(h) There is a natural number n such that $2n = 2^n$.

(i) If n is even, then n^2 is even.

(j) For all odd numbers n, n^2 is odd.

(k) For natural numbers n, n^2 is even if and only if n is even.

(l) There are no natural numbers m and n such that $\sqrt{2} = m/n$.

These are all mathematical statements, of different sorts (all of which will be discussed in more detail in the remainder of this chapter).

Statements (a) to (e) are straightforward *propositions* about certain numbers, and these are either true or false. Statements (d) and (e) are examples of *compound statements*. Statement (d) is true precisely when *either one (or both)* of the statements '21 is divisible by 3' and '21 is divisible by 5' is true. Statement (e) is true precisely when *both* of the statements '50 is divisible by 2' and '50 is divisible by 5' are true.

Statement (f) is different, because the number n is not specified and whether the statement is true or false will depend on the value of the so-called 'free variable' n. Such a statement is known as a *predicate*.

Statement (g) makes an assertion about *all* natural numbers and is an example of a *universal statement*.

Statement (h) asserts the existence of a particular number and is an example of an *existential* statement.

Statement (i) can be considered as an assertion about all even numbers, and so it is a universal statement, where the 'universe' is all even numbers. But it can also be considered as an *implication*, asserting that *if* n happens to be even, *then* n^2 is even.

Statement (j) is a universal statement about all odd numbers. It can also be thought of (or rephrased) as an implication, for it says precisely the same as 'if n is odd, then n^2 is odd'.

Statement (k) is an 'if and only if' statement: what it says is that n^2 is even, for a natural number n, *precisely when* n is even. But this means two things: namely that n^2 is even if n is even, and n is even if n^2 is even. Equivalently, it means that n^2 is even if n is even and that n^2 is odd if n is odd. So statement (k) will be true precisely if (i) and (j) are true.

Statement (l) asserts the non-existence of a certain pair of numbers (m, n). Another way of thinking about this statement is that it says that for all choices of (m, n), it is *not* the case that $m/n = \sqrt{2}$. (This is an example of the general rule that a non-existence statement can be thought of as a universal statement, something to be discussed later in more detail.)

It's probably worth giving some examples of things that are *not* proper mathematical statements.

'6 is a nice number' is not a mathematical statement. This is because 'nice number' has no mathematical meaning. However, if, beforehand, we had *defined* 'nice number' in some way, then this would not be a problem. For example, suppose we said:

> Let us say that a number is *nice* if it is the sum of all the positive numbers that divide it and are less than it.

Then '6 is a nice number' would be a proper mathematical statement, and it would be true, because 6 has positive divisors $1, 2, 3, 6$ and $6 = 1 + 2 + 3$. But without defining what 'nice' means, it's not a mathematical statement. Definitions are important.

'$n^2 + n$' is not a mathematical statement, because it does not say anything about $n^2 + n$. It is not a mathematical statement in the same way that 'David Cameron' is not a sentence: it makes no assertion about what David Cameron is or does. However, '$n^2 + n > 0$' is an example of a *predicate* with free variable n and, for a particular value of n, this is a mathematical statement. Likewise, 'for all natural numbers n, $n^2 + n > 0$' is a mathematical statement.

2.2.2 Introduction to proving statements

We've seen, above, various types of mathematical statement, and such statements are either true or false. But how would we establish the truth or falsity of these?

We can, even at this early stage, prove (by which we mean establish the truth of) or disprove (by which we mean establish the falsity of) most of the statements given above. Here's how we can do this.

(a) 20 is divisible by 4.

This statement is true. Yes, yes, I know it's 'obvious', but stay with me. To give a proper proof, we need first to understand exactly what the word 'divisible' means. You will probably most likely think that this means that when we divide 20 by 4 we get no remainder. This is correct: in general, for natural numbers n and d, to say that n is divisible by d (or, equivalently, that n is a multiple of d) means precisely that there is some natural number m for which $n = md$. Since $20 = 5 \times 4$, we see that 20 is divisible by 4. And that's a proof! It's utterly convincing, watertight, and not open to debate. Nobody can argue with it, not even a sociologist! Isn't this fun? Well, maybe it's not that impressive in such a simple situation, but we will certainly prove more impressive results later.

(b) 21 is not divisible by 7.

This is false. It's false because 21 *is* divisible by 7, because $21 = 3 \times 7$.

(c) 21 is divisible by 4.

This is false, as can be established in a number of ways. First, we note that if the natural number m satisfies $m \leq 5$, then $m \times 4$ will be no more than 20. And if $m \geq 6$ then $m \times 4$ will be at least 24. Well, any natural number m is either at most 5 or at least 6 so, for all possible m, we do not have $m \times 4 = 21$ and hence there is no natural number m for which $m \times 4 = 21$. In other words, 21 is not divisible by 4. Another argument (which is perhaps more straightforward, but which relies on properties of rational numbers rather than just simple properties of natural numbers) is to note that $21/4 = 5.25$, and this is not a natural number, so 21 is not divisible by 4. (This second approach is the same as showing that 21 has remainder 1, not 0, when we divide by 4.)

(d) 21 is divisible by 3 or 5.

As we noted above, this is a compound statement and it will be true precisely when one (or both) of the following statements is true:

(i) 21 is divisible by 3

(ii) 21 is divisible by 5.

Statement (i) is true, because $21 = 7 \times 3$. Statement (ii) is false. Because at least one of these two statements is true, statement (d) is true.

(e) 50 is divisible by 2 and 5.

This is true. Again, this is a compound statement and it is true precisely if *both* of the following statements are true:

(i) 50 is divisible by 2

(ii) 50 is divisible by 5.

Statements (i) and (ii) are indeed true because $50 = 25 \times 2$ and $50 = 10 \times 5$. So statement (e) is true.

(f) n^2 is even

As mentioned above, whether this is true or false depends on the value of n. For example, if $n = 2$ then $n^2 = 4$ is even, but if $n = 3$ then $n^2 = 9$ is odd. So, unlike the other statements (which are *propositions*), this is a *predicate* $P(n)$. The predicate will become a proposition when we assign a particular value to n to it, and the truth or falsity of the proposition can then be established. Statements (i), (j), (k) below do this comprehensively.

(g) For every natural number n, the number $n^2 + n$ is even

Here's our first non-immediate, non-trivial, proof. How on earth can we prove this, if it is true, or disprove it, if it is false? Suppose it was false. How would you convince someone of that? Well, the statement says that *for every* natural number n, $n^2 + n$ is even. So if you managed (somehow!) to find a particular N for which $N^2 + N$ happened to be odd, you could prove the statement false by simply observing that 'When $n = N$, it is *not* the case that $n^2 + n$ is even.' And that would be the end of it. So, in other words, if a universal statement about natural numbers is false, you can prove it is false by showing that its conclusion is false for *some particular* value of n. But suppose the statement is true. How could you prove it. Well, you could prove it for $n = 1$, then $n = 2$, then $n = 3$, and so on, but at some point you would expire and there would still be numbers n that you hadn't yet proved it for. And that simply wouldn't do, because if you proved it true for the first 9999 numbers, it might be false when $n = 10000$. So what you need is a more sophisticated, *general* argument that shows the statement is true for any *arbitrary* n.

Now, it turns out that this statement is true. So we need a nice general argument to establish this. Well, here's one approach. We can note that $n^2 + n = n(n + 1)$. The numbers n and $n + 1$ are consecutive natural numbers. So one of them is odd and one of them is even. When you multiply any odd number and any even number together, you get an even number, so $n^2 + n$ is even. Are you convinced? Maybe not? We really should be more explicit. Suppose n is even. What that means is that, for some integer k, $n = 2k$. Then $n + 1 = 2k + 1$ and hence

$$n(n + 1) = 2k(2k + 1) = 2\left(k(2k + 1)\right).$$

Because $k(2k + 1)$ is an integer, this shows that $n^2 + n = n(n + 1)$ is divisible by 2; that is, it is even. We supposed here that n was even. But it might be odd, in which case we would have $n = 2k + 1$ for some integer k. Then

$$n(n + 1) = (2k + 1)(2k + 2) = 2\left((2k + 1)(k + 1)\right),$$

which is, again, even, because $(2k + 1)(k + 1)$ is an integer.

Right, we're really proving things now. This is a very general statement, asserting something about *all* natural numbers, and we have managed to prove it. I find that quite satisfying, don't you?

(h) There is a natural number n such that $2n = 2^n$.

This is an *existential statement*, asserting that *there exists* n with $2n = 2^n$. Before diving in, let's pause for a moment and think about how we might deal with such statements. If an existential statement like this is true we would need only to show that its conclusion (which in this case is $2n = 2^n$) holds for some particular n. That is, we need only find an n that works. If the statement is false, we have a lot more work to do in order to prove that it is false. For, to show that it is false, we would need to show that, for *no* value of n does the conclusion holds. Equivalently, for *every* n, the conclusion fails. So we'd need to prove a universal statement and, as we saw in the previous example, that would require us to come up with a suitably general argument.

In fact, this statement is true. This is because when $n = 1$ we have $2n = 2 = 2^1 = 2^n$.

(i) If n is even, then n^2 is even

This is true. The most straightforward way to prove this is to assume that n is some (that is, *any*) even number and then show that n^2 is even. So suppose n is even. Then $n = 2k$ for some integer k and hence $n^2 = (2k)^2 = 4k^2$. This is even because it is $2(2k^2)$ and $2k^2$ is an integer.

(j) For all odd numbers n, n^2 is odd.

This is true. The most straightforward way to prove this is to assume that n is *any* odd number and then show that n^2 is also odd. So suppose n is odd. Then $n = 2k + 1$ for some integer k and hence $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$. To establish that this is odd, we need to show that it can be written in the form $2K + 1$ for some integer K. Well, $4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. This is indeed of the form $2K + 1$, where K is the integer $2k^2 + 2k$. Hence n^2 is odd.

Another way to prove this result is to prove that if n^2 is even then n must be even. We won't do that right now, because to do it properly requires a result we meet later concerning the factorisation of numbers into prime numbers. But think about the strategy for a moment. Suppose we were able to prove the following statement, which we'll call Q:

Q: *if* n^2 is even *then* n is even.

Why would that establish what we want (namely that if n is odd then n^2 is odd). Well, one way is to observe that Q is what's called the *contrapositive* of statement (j) that we're trying to prove, and the contrapositive is *logically equivalent* to the initial statement. (This is a bit of formal logic, and we will discuss this more later). But there's another way of thinking about it, which is perhaps easier to understand at this stage. Suppose we have proved statement Q and suppose that n is odd. Then it must be the case that n^2 is odd. For, if n^2 was not odd, it would be even and then Q would tell us that this means n is even. But we have assumed n is odd. It cannot be both even and odd, so we have reached a contradiction. By assuming that the opposite conclusion holds (n^2 even) we have shown that something impossible happens. This type of argument is known as a *proof by contradiction* and it is often very powerful. We will see more about this later.

(k) For natural numbers n, n^2 is even if and only if n is even.

This is true. What we have shown in proving (i) and (j) is that if n is even then n^2 is even, and if n is odd then n^2 is odd. The first, (statement (i)) establishes that *if* n is even, then n^2 is even. The second of these (statement (j)) establishes that n^2 is even *only if* n is even. This is because it shows that n^2 is odd if n is odd, from which it follows that if n^2 is even, n must not have been odd, and therefore must have been even. 'If and only if' statements if this type are very important. As we see here, the proof of such statements breaks down into the proof of two 'If-then' statements.

(l) There are no natural numbers m and n such that $\sqrt{2} = m/n$.

This is, in fact, true, though we defer the proof for now, until we know more about factorisation of numbers into prime numbers. We merely comment that the easiest way to prove the statement is to use a proof by contradiction.

These examples hopefully demonstrate that there are a wide range of statements and proof techniques, and in the rest of this chapter we will explore these further.

Right now, one thing I hope comes out very clearly from these examples is that to prove a mathematical statement, you need to know precisely what it means. Well, that sounds obvious, but you can see how detailed we had to be about the meanings (that is, the *definitions*) of the terms 'divisible', 'even' and 'odd'. Definitions are very important.

2.3 Some basic logic

Mathematical statements can be true or false Let's denote 'true' by T and 'false' by F. Given a statement, or a number of statements, it is possible to form other statements. This was indicated in some of the examples above (such as the compound statements).

A technique known as the use of 'truth tables' enables us to define 'logical operations' on statements, and to determine when such statements are true. This is all a bit vague, so let's get down to some concrete examples.

2.3.1 Negation

The simplest way to take a statement and form another statement is to *negate* the statement. The *negation* of a statement P is the statement $\neg P$ (sometimes just denoted 'not P'), which is defined to be true exactly when P is false. This can be described in the very simple truth table, Table 2.1:

P	$\neg P$
T	F
F	T

Table 2.1: The truth table for 'negation' or 'not'

What does the table signify? Quite simply, it tells us that if P is true then $\neg P$ is false and if P is false then $\neg P$ is true.

> **Example 2.1** If P is '20 is divisible by 3' then $\neg P$ is '20 is not divisible by 3'. Here, P is false and $\neg P$ is true.

It has, I hope, been indicated in the examples earlier in this chapter, that to disprove a universal statement about natural numbers amounts to proving an existential statement. That is, if we want to disprove a statement of the form 'for all natural numbers n, property $p(n)$ holds' (where $p(n)$ is some predicate, such as 'n^2 is even') we need only produce some N for which $p(N)$ fails. Such an N is called a *counterexample*. Equally, to disprove an existential statement of the form 'there is some n such that property $p(n)$ holds', one would have to show that for *every* n, $p(n)$ fails. That is, to disprove an existential statement amounts to proving a universal one. But, now that we have the notion of the negation of a statement we can phrase this a little more formally. Proving that a statement P is false is equivalent to proving that the negation $\neg P$ is true. In the language of logic, therefore, we have the following:

- The negation of a universal statement is an existential statement.
- The negation of an existential statement is a universal statement.

More precisely,

- The negation of the universal statement 'for all n, property $p(n)$ holds' is the existential statement 'there is n such that property $p(n)$ does not hold'.
- The negation of the existential statement 'there is n such that property $p(n)$ holds' is the universal statement 'for all n, property $p(n)$ does not hold'.

We could be a little more formal about this, by defining the negation of a predicate $p(n)$ (which, recall, only has a definitive true or false value once n is specified) to be the predicate $\neg p(n)$ which is true (for any particular n) precisely when $p(n)$ is false. Then we might say that

- The negation of the universal statement 'for all n, $p(n)$ is true' is the existential statement 'there is n such that $\neg p(n)$ is true'.
- The negation of the existential statement 'there is n such that $p(n)$ is true' is the universal statement 'for all n, $\neg p(n)$ is true'.

Now, let's not get confused here. None of this is really difficult or new. We meet such logic in everyday life. If I say 'It rains every day in London' then either this statement is true or it is false. If it is false, it is because on (at least) one day it does not rain. The negation (or disproof) of the statement 'On every, it rains in London' is simply 'There is a day on which it does not rain in London'. The former is a universal statement ('On every day, ...') and the latter is an existential statement ('there is ...'). Or, consider the statement 'There is a student who enjoys reading these lecture notes'. This is an existential statement ('There is ...'). This is false if 'No student enjoys reading these lecture notes'. Another way of phrasing this last statement is 'Every student reading these lecture notes does not enjoy it'. This is a more awkward expression, but it emphasises that the negation of the initial, existential statement, is a universal one ('Every student ...').

The former is an existential statement ('there is something I will write that ...') and the latter is a universal statement ('everything I write will ...). This second example is a little more complicated, but it serves to illustrate the point that much of logic is simple common sense.

2.3.2 Conjunction and disjunction

There are two very basic ways of combining propositions: through the use of 'and' (known as conjunction) and the use of 'or' (known as disjunction).

Suppose that P and Q are two mathematical statements. Then 'P and Q', also denoted $P \wedge Q$, and called the *conjunction* of P and Q, is the statement that is true precisely when *both* P and Q are true. For example, statement (e) above, which is

'50 is divisible by 2 and 5'

is the conjunction of the two statements

- 50 is divisible by 2
- 50 is divisible by 5.

Statement (e) is true because *both* of these two statements are true.

Table 2.2 gives the truth table for the conjunction P and Q:

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Table 2.2: The truth table for 'and'

What Table 2.2 says is simply that $P \wedge Q$ is true precisely when *both* P and Q are true (and in no other circumstances).

Suppose that P and Q are two mathematical statements. Then 'P or Q', also denoted $P \vee Q$, and called the *disjunction* of P and Q, is the statement that is true precisely when P, or Q, or both, are true. For example, statement (d) above, which is

'21 is divisible by 3 or 5'

is the disjunction of the two statements

- 21 is divisible by 3
- 21 is divisible by 5.

Statement (d) is true because at least one (namely the first) of these two statements is true.

Note one important thing about the mathematical interpretation of the word 'or'. It is *always* used in the 'inclusive-or' sense. So $P \vee Q$ is true in the case when P is true, or Q is true, or *both*. In some ways, this use of the word 'or' contrasts with its use in normal everyday language, where it is often used to specify a choice between mutually exclusive alternatives. (For example 'You're either with us or against us'.) But if I say 'Tomorrow I will wear brown trousers or I will wear a yellow shirt' then, in the mathematical way in which the word 'or' is used, the statement would be true if I wore brown trousers and any shirt, any trousers and a yellow shirt, and also if I wore brown trousers and a yellow shirt. You might have your doubts about my dress sense in this last case, but, logically, it makes my statement true.

Table 2.2 gives the truth table for the disjunction P and Q:

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Table 2.3: The truth table for 'or'

What Table 2.3 says is simply that $P \vee Q$ is true precisely when *at least one of* P and Q is true.

2.4 If-then statements

It is very important to understand the formal meaning of the word ‘if’ in mathematics. The word is often used rather sloppily in everyday life, but has a very precise mathematical meaning. Let me give you an example. Suppose I tell you ‘If it rains, then I wear a raincoat’, and suppose that this is a true statement. Well, then, suppose it rains. You can certainly conclude I will wear a raincoat. But what if it does not rain? Well, you can’t conclude anything. My statement only tells you about what happens *if* it rains. If it does not, then I might, or I might not, wear a raincoat: and whether I do or not does not affect the truth of the statement I made. You have to be clear about this: an ‘if-then’ statement only tells you about what follows *if* something particular happens.

More formally, suppose P and Q are mathematical statements (each of which can therefore be either true or false). Then we can form the statement denoted $P \Rightarrow Q$ (‘P implies Q’ or, equivalently, ‘if P, then Q’), which has as its truth table Table 2.4. (This type of statement is known as an *if-then* statement or an *implication*.)

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Table 2.4: The truth table for ‘$P \Rightarrow Q$’

Note that the statement $P \Rightarrow Q$ is false only when P is true but Q is false. (To go back to the previous example, the statement ‘If it rains, I wear a raincoat’ is false precisely if it does rain but I do not wear a raincoat.) This is tricky, so you may have to spend a little time understanding it. As I’ve suggested, perhaps the easiest way is to think about when a statement ‘if P, then Q’ is false.

The statement $P \Rightarrow Q$ can also be written as $Q \Leftarrow P$. There are different ways of describing $P \Rightarrow Q$, such as:

- if P then Q
- P implies Q
- P is sufficient for Q
- Q if P
- P only if Q
- Q whenever P
- Q is necessary for P.

All these mean the same thing. The first two are the ones I will use most frequently.

If $P \Rightarrow Q$ and $Q \Rightarrow P$ then this means that Q will be true precisely when P is. That is Q is true *if and only if* P is. We use the single piece of notation $P \iff Q$ instead of the two separate $P \Rightarrow Q$ and $Q \Leftarrow P$. There are several phrases for describing what $P \iff Q$ means, such as:

- P if and only if Q (sometimes abbreviated to ‘P iff Q’)
- P is equivalent to Q
- P is necessary and sufficient for Q
- Q is necessary and sufficient for P.

The truth table is shown in Table 2.5, where we have also indicated the truth or falsity of $P \Rightarrow Q$ and $Q \Rightarrow P$ to emphasise that $P \iff Q$ is the same as the conjunction $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$P \iff Q$
T	**T**	T	T	**T**
T	**F**	F	T	**F**
F	**T**	T	F	**F**
F	**F**	T	T	**T**

Table 2.5: The truth table for ‘$P \iff Q$’

What the table shows is that $P \iff Q$ is true precisely when P and Q are either both true or both false.

> **Activity 2.1** Look carefully at the truth table and understand why the values for $P \iff Q$ are as they are. In particular, try to explain in words why the truth table is the way it is.

2.5 Logical equivalence

Two statements are *logically equivalent* if when either one is true, so is the other, and if either one is false, so is the other. For example, for statements P and Q, the statements $\neg(P \vee Q)$ and $\neg P \wedge \neg Q$ are logically equivalent. We can see this from the truth table, Table 2.6, which shows that, in all cases, the two statements take the same logical value T or F). (This value is highlighted in bold.)

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$\neg P \wedge \neg Q$
T	T	T	**F**	F	F	**F**
T	F	T	**F**	F	T	**F**
F	T	T	**F**	T	F	**F**
F	F	F	**T**	T	T	**T**

Table 2.6: The truth tables for $\neg(P \vee Q)$ and $\neg P \wedge \neg Q$

The fact that $\neg(P \vee Q)$ and $\neg P \wedge \neg Q$ are logically equivalent is quite easy to understand. The statement $P \vee Q$ is true if and only if at least one of P, Q is true. The statement is therefore false precisely when *both* P and Q are false, which means $\neg P$ and $\neg Q$ are both true, which means $\neg P \wedge \neg Q$ is true. Again, we can understand these things fairly easily with some common sense. If I tell you 'I will wear brown trousers or I will wear a yellow shirt' then this is a false statement only if I *do not* wear brown trousers *and* I *do not* wear a yellow shirt.

Now that we know the meaning of $\iff$, we can see that to say that $\neg(P \vee Q)$ and $\neg P \wedge \neg Q$ are logically equivalent is to say that $\neg(P \vee Q) \iff \neg P \wedge \neg Q$.

> **Activity 2.2** Show that the statements $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are logically equivalent. [This shows that the negation of $P \wedge Q$ is $\neg P \vee \neg Q$. That is, $\neg(P \wedge Q)$ is equivalent to $\neg P \vee \neg Q$.]

2.6 Converse statements

Given an implication $P \Rightarrow Q$, the 'reverse' implication $Q \Rightarrow P$ is known as its *converse*. Generally, there is no reason why the converse should be true just because the implication is. For example, consider the statement 'If it is Tuesday, then I buy the Guardian newspaper.' The converse is 'If I buy the Guardian newspaper, then it is Tuesday'. Well, I might buy that newspaper on other days too, in which case the implication can be true but the converse false. We've seen, in fact, that if both $P \Rightarrow Q$ and $Q \Rightarrow P$ then we have a special notation, $P \iff Q$, for this situation. Generally, then, the truth or falsity of the converse $Q \Rightarrow P$ has to be determined separately from that of the implication $P \Rightarrow Q$.

> **Activity 2.3** What is the converse of the statement 'if the natural number n divides 4 then n divides 12'? Is the converse true? Is the original statement true?

2.7 Contrapositive statements

The *contrapositive* of an implication $P \Rightarrow Q$ is the statement $\neg Q \Rightarrow \neg P$. The contrapositive is logically equivalent to the implication, as Table 2.7 shows. (The columns highlighted in bold are identical.)

P	Q	$P \Rightarrow Q$	$\neg P$	$\neg Q$	$\neg Q \Rightarrow \neg P$
T	T	**T**	F	F	**T**
T	F	**F**	F	T	**F**
F	T	**T**	T	F	**T**
F	F	**T**	T	T	**T**

Table 2.7: The truth tables for $P \Rightarrow Q$ and $\neg Q \Rightarrow \neg P$.

If you think about it, the equivalence of the implication and its contrapositive makes sense. For, $\neg Q \Rightarrow \neg P$ says that if Q is false, P is false also. So, it tells us that we cannot have Q false and P true, which is precisely the same information as is given by $P \Rightarrow Q$.

So what's the point of this? Well, sometimes you might want to prove $P \Rightarrow Q$ and it will, in fact, be easier to prove instead the equivalent (contrapositive) statement $\neg Q \Rightarrow \neg P$. See Biggs, section 3.5 for an example.

2.8 Working backwards to obtain a proof

We've already seen, in the examples earlier in this chapter, how some statements may be proved directly. For example, in order to prove a universal statement 'for all n, $P(n)$' about natural numbers, we would need to provide a proof that starts by assuming that n is any given (that is, *arbitrary*) natural number and show the desired conclusion holds. To disprove such a statement (which is the same as proving its negation), we would simply need to find a single value of n for which $P(n)$ is false (and such an n is known as a *counterexample*).

However, some statements are difficult to prove directly. It is sometimes easier to 'work backwards'. Suppose you are asked to prove something, such as an inequality or equation. It might be easier to see how to do so if the end-result (the inequality or equation you are required to prove) is simplified, or expanded, or re-written in some way. Here's an example.

> **Example 2.2** Prove the statement that: 'if a, b are real numbers and $a \neq b$, then $ab < (a^2 + b^2)/2$'.
>
> It's certainly not immediately obvious how to approach this. But let's start with what we want to prove. This is the inequality $ab < (a^2 + b^2)/2$, which can be rewritten as $a^2 + b^2 - 2ab > 0$. Now, this can be simplified as $(a - b)^2 > 0$ and maybe now you can see why it is true: the given fact that $a \neq b$ means that $a - b \neq 0$ and hence $(a - b)^2$ is a positive number. So we see why the statement is true. To write down a nice proof, we can now reverse this argument, as follows:
>
> **Proof** Since $a \neq b$, $a - b \neq 0$ and, hence, $(a - b)^2 > 0$. But $(a - b)^2 = a^2 + b^2 - 2ab$. So we have $a^2 + b^2 > 2ab$ and, therefore, $ab < (a^2 + b^2)/2$, as required. □

There are a few things to note here. First, mathematics is a language and what you write has to make good sense. Often, it is tempting to make too much use of symbols rather than words. But the words used in this proof, and the punctuation, make it easy to read and give it a structure and an argument. You should find yourself using words like 'so, 'hence', 'therefore', 'since', 'because', and so on. *Do* use words and punctuation and, whatever you do, do not replace them by symbols of your own invention! A second thing to note is the use of the symbol '□'. There is nothing particularly special about this symbol: others could be used. What it achieves is that it indicates that the proof is finished. There is no need to use such a symbol, but you will find that textbooks do make much use of symbols to indicate when proofs have ended. It enables the text to be more readable, with proofs not running into the main body of the text. Largely, these are matters of style, and you will develop these as you practice and read the textbooks.

2.9 Sets

2.9.1 Basics

You have probably already met some basic ideas about sets and there is not too much more to add at this stage, but they are such an important idea in abstract mathematics that they are worth discussing here.

Loosely speaking, a set may be thought of as a collection of objects. A set is usually described by listing or describing its *members* inside curly brackets. For example, when we write $A = \{1, 2, 3\}$, we mean that the objects belonging to the set A are the numbers $1, 2, 3$ (or, equivalently, the set A consists of the numbers $1, 2$ and 3). Equally (and this is what we mean by 'describing' its members), this set could have been written as

$$A = \{n \mid n \text{ is a whole number and } 1 \leq n \leq 3\}.$$

Here, the symbol $\mid$ stands for 'such that'. Often, the symbol ':' is used instead, so that we might write

$$A = \{n : n \text{ is a whole number and } 1 \leq n \leq 3\}.$$

When x is an object in a set A, we write $x \in A$ and say 'x belongs to A' or 'x is a member of A'. If x is not in A we write $x \notin A$.

As another example, the set

$$B = \{x \in \mathbb{N} \mid x \text{ is even}\}$$

has as its members the set of positive even integers. Here we are specifying the set by *describing* the defining property of its members.

Sometimes it is useful to give a *constructional* description of a set. For example, $C = \{n^2 \mid n \in \mathbb{N}\}$ is the set of natural numbers known as the 'perfect squares'.

The set which has no members is called the *empty set* and is denoted by $\emptyset$. The empty set may seem like a strange concept, but it has its uses.

2.9.2 Subsets

We say that the set S is a *subset* of the set T, and we write $S \subseteq T$, if every member of S is a member of T. For example, $\{1, 2, 5\} \subseteq \{1, 2, 4, 5, 6, 40\}$. (Be aware that some texts use $\subset$ where we use $\subseteq$.) What this means is that the statement

$$x \in S \Rightarrow x \in T$$

is true.

A rather obvious, but sometimes useful, observation is that, given two sets A and B, $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$. So to prove two sets are equal, we can prove that each of these two 'containments' holds. That might seem clumsy, but it is, in many cases, the best approach.

For any set A, the empty set, $\emptyset$, is a subset of A. You might think this is strange, because what it means is that 'every member of $\emptyset$ is also a member of A'. But $\emptyset$ has no members! The point, however, is that there is no object in $\emptyset$ that is *not* also in A (because there are no objects at all in $\emptyset$).

2.9.3 Unions and intersections

Given two sets A and B, the *union* $A \cup B$ is the set whose members belong to A or B (or both A and B): that is,

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

Equivalently, to use the notation we've learned,

$$x \in A \cup B \iff (x \in A) \vee (x \in B).$$

> **Example 2.3** If $A = \{1, 2, 3, 5\}$ and $B = \{2, 4, 5, 7\}$, then $A \cup B = \{1, 2, 3, 4, 5, 7\}$.

Similarly, we define the *intersection* $A \cap B$ to be the set whose members belong to both A and B:

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

So,

$$x \in A \cap B \iff (x \in A) \wedge (x \in B).$$

2.9.4 Universal sets and complements

We've been a little informal about what the possible 'objects' in a set might be. Officially, we always work with respect to some 'universal set' E. For example, if we are thinking about sets of natural numbers, the universal set (the possible candidates for membership of the sets we might want to consider) is the set $\mathbb{N}$ of all natural numbers. This might seem like an unnecessary complication, but it is essential. Suppose I tell you that the set A is the set of all even natural numbers. What are the objects that do not belong to A? Well, in the context of natural numbers, it is all odd natural numbers. The context is important (and it is this that is encapsulated in the universal set). Without that context (or universal set), then there are many other objects that we could say do not belong to A, such as negative integers, apples, bananas and elephants. (I could go on, but I hope you get the point!)

Given a universal set E and a subset A of E, the *complement* of A (sometimes called the *complement of A in E*) is denoted by $E \setminus A$ and is

$$E \setminus A = \{x \in E \mid x \notin A\}.$$

If the universal set is clear, the complement of A is sometimes denoted by $\bar{A}$ or A^c (with textbooks differing in their notation).

Suppose A is any subset of E. Because each member of E is either a member of A, or is not a member of A, it follows that

$$A \cup (E \setminus A) = E.$$

2.9.5 Sets and logic

There are a great many comparisons and analogies between set theory and logic. Using the shorthand notation for complements, one of the 'De-Morgan' laws of complementation is that

$$\overline{A \cap B} = \bar{A} \cup \bar{B}.$$

This looks a little like the fact (observed in an earlier Learning Activity) that $\neg(P \wedge Q)$ is equivalent to $\neg P \vee \neg Q$. And this is more than a coincidence. The negation operation, the conjunction operation, and the disjunction operation on statements behave entirely in the same way as the complementation, intersection, and union operations (in turn) on sets. In fact, when you start to prove things about sets, you often end up giving arguments that are based in logic.

For example, how would we prove that $\overline{A \cap B} = \bar{A} \cup \bar{B}$? We could argue as follows:

$$\begin{aligned} x \in \overline{A \cap B} &\iff x \notin A \cap B \\ &\iff \neg(x \in A \cap B) \\ &\iff \neg((x \in A) \wedge (x \in B)) \\ &\iff \neg(x \in A) \vee \neg(x \in B) \\ &\iff (x \in \bar{A}) \vee (x \in \bar{B}) \\ &\iff x \in \bar{A} \cup \bar{B}. \end{aligned}$$

What the result says is, in fact, easy to understand: if x is not in *both* A and B, then that's precisely because it fails to be in (at least) one of them.

For two sets A and B (subsets of a universal set E), the *complement of B in A*, denoted by $A \setminus B$, is the set of objects that belong to A but not to B. That is,

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

> **Activity 2.4** Prove that $A \setminus B = A \cap (E \setminus B)$.

2.9.6 Cartesian products

For sets A and B, the *Cartesian product* $A \times B$ is the set of all *ordered pairs* (a, b), where $a \in A$ and $b \in B$. For example, if $A = B = \mathbb{R}$ then $A \times B = \mathbb{R} \times \mathbb{R}$ is the set of all ordered pairs of real numbers, usually denoted by $\mathbb{R}^2$.

2.9.7 Power sets

For a set A, the set of all subsets of A, denoted $\mathcal{P}(A)$, is called the *power set* of A. Note that the power set is a set of sets. For example, if $A = \{1, 2, 3\}$, then

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

> **Activity 2.5** Write down the power set of the set $A = \{1, 2, 3, 4\}$.

> **Activity 2.6** Suppose that A has n members, where $n \in \mathbb{N}$. How many members does $\mathcal{P}(A)$ have?

2.10 Quantifiers

We have already met the ideas of universal and existential statements involving natural numbers. More generally, given any set E, a *universal statement* on E is one of the form 'for all $x \in E$, $P(x)$'. This statement is true if $P(x)$ is true for all x in E, and it is false if there is some x in E (known as a *counterexample*) such that $P(x)$ is false. We have a special symbol that is used in universal statements: the symbol '$\forall$' means 'for all'. So the typical universal statement can be written as

$$\forall x \in E,\ P(x).$$

(The comma is not necessary, but I think it looks better.) An *existential statement* on E is one of the form 'there is $x \in E$ such that $P(x)$', which is true if there is some $x \in E$ for which $P(x)$ is true, and is false if for every $x \in E$, $P(x)$ is false. Again, we have a useful symbol, '$\exists$', meaning 'there exists'. So the typical existential statement can be written as

$$\exists x \in E,\ P(x).$$

Here, we have omitted the phrase 'such that', but this is often included if the statement reads better with it. For instance, we could write

$$\exists n \in \mathbb{N},\ n^2 - 2n + 1 = 0,$$

but it would probably be easier to read

$$\exists n \in \mathbb{N} \text{ such that } n^2 - 2n + 1 = 0.$$

Often 'such that' is abbreviated to 's.t.'. (By the way, this statement is true because $n = 1$ satisfies $n^2 - 2n + 1 = 0$.)

We have seen that the negation of a universal statement is an existential statement and vice versa. In symbols, $\neg(\forall x \in E,\ P(x))$ is logically equivalent to $\exists x \in E,\ \neg P(x)$; and $\neg(\exists x \in E,\ P(x))$ is logically equivalent to $\forall x \in E,\ \neg P(x)$.

With these observations, we can now form the negations of more complex statements. Consider the statement

$$\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, m > n.$$

> **Activity 2.7** What does the statement $\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, m > n$ mean? Is it true?

What would the negation of the statement be? Let's take it gently. First, notice that the statement is

$$\forall n \in \mathbb{N}, (\exists m \in \mathbb{N}, m > n).$$

The parentheses here do not change the meaning. According to the rules for negation of universal statements, the negation of this is

$$\exists n \in \mathbb{N}, \neg(\exists m \in \mathbb{N}, m > n).$$

But what is $\neg(\exists m \in \mathbb{N}, m > n)$? According to the rules for negating existential statements, this is equivalent to $\forall m \in \mathbb{N}, \neg(m > n)$. What is $\neg(m > n)$? Well, it's just $m \leq n$. So what we see is that the negation of the initial statement is

$$\exists n \in \mathbb{N}, \forall m \in \mathbb{N}, m \leq n.$$

We can put this argument more succinctly, as follows:

$$\begin{aligned} \neg\left(\forall n \in \mathbb{N}(\exists m \in \mathbb{N}, m > n)\right) &\iff \exists n \in \mathbb{N}, \neg(\exists m \in \mathbb{N}, m > n) \\ &\iff \exists n \in \mathbb{N}, \forall m \in \mathbb{N}, \neg(m > n) \\ &\iff \exists n \in \mathbb{N}, \forall m \in \mathbb{N}, m \leq n. \end{aligned}$$

2.11 Proof by contradiction

We've seen a small example of proof by contradiction earlier in the chapter. Suppose you want to prove $P \Rightarrow Q$. One way to do this is by contradiction. What this means is that you suppose P is true but Q is false (in other words, that the statement $P \Rightarrow Q$ is false) and you show that, somehow, this leads to a conclusion that you know, definitely, to be false.

Here's an example.

> **Example 2.4** There are no integers m, n such that $6m + 8n = 1099$.

To prove this by contradiction, we can argue as follows:

Suppose that integers m, n *do* exist such that $6m + 8n = 1099$. Then since 6 is even, $6n$ is also even; and, since 8 is even, $8n$ is even. Hence $6m + 8n$, as a sum of two even numbers, is even. But this means $1099 = 6m + 8n$ is an even number. But, in fact, it is not even, so we have a contradiction. It follows that m, n of the type required do *not* exist. □

This sort of argument can be a bit perplexing when you first meet it. What's going on in the example just given? Well, what we show is that if such m, n exist, then something impossible happens: namely the number 1099 is both even and odd. Well, this can't be. If supposing something leads to a conclusion you know to be false, then the initial supposition must be false. So the conclusion is that such integers m, n do not exist.

Probably the most famous proof by contradiction is Euler's proof that there are infinitely many prime numbers. A prime number is a natural number greater than 1 which is only divisible by 1 and itself. Such numbers have been historically of huge importance in mathematics, and they are also very useful in a number of important applications, such as information security. The first few prime numbers are $2, 3, 5, 7, 11, \ldots$. A natural question is: does this list go on forever, or is there a largest prime number? In fact, the list goes on forever: there are infinitely many prime numbers. We'll mention this result again later. A full, detailed, understanding of the proof requires some results we'll meet later, but you should be able to get the flavour of it at this stage. So here it is, a very famous result:

There are infinitely many prime numbers.

Proof (Informally written for the sake of exposition) Suppose *not*. That is, suppose there are only a finite number of primes. Then there's a largest one. Let's call it M. Now consider the number

$$X = (2 \times 3 \times 5 \times 7 \times 11 \times \cdots \times M) + 1,$$

which is the product of *all* the prime numbers (2 up to M), with 1 added. Notice that $X > M$, so X is not a prime (because M is the largest prime). If a number X is not prime, that means that it has a divisor p that is a prime number and which satisfies $1 < p < X$. [*This is the key observation: we haven't seen this yet, but we will later.*] But p must therefore be one of the numbers $2, 3, 5, \ldots, M$. However, X is *not* divisible by any of these numbers, because it has remainder 1 when divided by any of them. So we have reached a contradiction: on the one hand, X must be divisible by one of these primes, and on the other, it is not. So the initial supposition that there were *not* infinitely many primes simply must be wrong. We conclude there are infinitely many primes. □

This proof has been written in a fairly informal and leisurely way to help explain what's happening. It could all be written more succinctly.

2.12 Some terminology

At this point, it's probably worth introducing some important terminology. When, in Mathematics, we prove a true statement, we often say we are proving a *Theorem*, or a *Proposition*. (Usually the word 'Proposition' is used if the statement does not seem quite so significant as to merit the description 'Theorem'.) A theorem that is a preliminary result leading up to a Theorem is often called a *Lemma*, and a minor theorem that is a fairly direct consequence of, or special case of, a theorem is called a *Corollary*, if it is not significant enough itself to merit the title Theorem. For your purposes, it is important just to know that these words all mean true mathematical statements. You should realise that these terms are used subjectively: for instance, the person writing the mathematics has to make a decision about whether a particular result merits the title 'Theorem' or is, instead, merely to be called a 'Proposition'.

2.13 General advice

2.13.1 Introduction

Proving things is difficult. Inevitably, when you read a proof, in the textbooks or in these notes, you will ask 'How did the writer know to do that?' and you will often find you asking yourself 'How can I even begin to prove this?'. This is perfectly normal. This is where the key difference between abstract mathematics and more 'methods-based' mathematics lies. If you are asked to differentiate a function, you just go ahead and do it. It might be technically difficult in some cases, but there is no doubt about what approaches you should use. But proving something is more difficult. You might try to prove it, and fail. That's fine: what you should do in that case is try another attack. Keep trying until you crack it. (I suppose this is a little bit like integration. You'll know that there are various methods, but you don't necessarily know which will work on a particular integral, so you should try one, and keep trying until you manage to find the integral.) Abstract mathematics should always be done with a large pile of scrap paper at your disposal. You are unlikely to be able to write down a perfect solution to a problem straight away: some 'scratching around' to get a feel for what's going on might well be needed, and some false starts might be pursued first. If you expect to be able to envisage a perfect solution in your head and then write it down perfectly, you are placing too much pressure on yourself. Abstract mathematics is simply not done like that.

In this chapter I have tried to indicate that there are methodical approaches to proof (such as proof by contradiction, for example). What you have to always be able to do is to understand *precisely* what it is that you have to prove. That sounds obvious, but it is something the importance of which is often underestimated. Once you understand what you need to show (and, here, working backwards a little from that end-point might be helpful, as we've seen), then you have to try to show it. And you must know when you have done so! So it is inevitable that you will have to take a little time to think about what is required: you cannot simply 'dive in' like you might to a differentiation question.

All this becomes much easier as you practice it. You should attempt problems from the textbooks (and also the problems below). Problems are a valuable resource and you are squandering this resource if you simply turn to the answers (should these be available). It is one thing to 'agree' with an answer, or to understand a proof, but it is quite a different thing to come up with a proof yourself. There is no point in looking at the answer before you have tried hard yourself to answer the problem. By trying (and possibly failing), you will learn more than simply by reading answers. Exam questions will be different from problems you have seen, so there is no point at all in 'learning' answers. You need to understand how to approach problems and how to answer them for yourself.

2.13.2 How to write mathematics

You should write mathematics **in English**!! You shouldn't think that writing mathematics is just using formulae. A good way to see if your writing makes sense is by reading it aloud (where you should only read what you really have written, not adding extra words). If it sounds like nonsense, a sequence of loose statements with no obvious relations, then you probably need to write it again.

Don't use more symbols than necessary.
Since many people seem to think that mathematics involves writing formulae, they often use symbols to replace normal English words. An eternal favourite is the double arrow "⟹" to indicate that one thing follows from the other. As in:

$$x^2 = 1 \quad \Longrightarrow \quad x = 1 \text{ or } x = -1.$$

This is not only pure laziness, since it's just as easy to write:

$$x^2 = 1, \text{ hence } x = 1 \text{ or } x = -1.$$

But it is even probably not what was meant! The implication arrow "⟹" has a logical meaning "if . . . , then . . . ". So if you write "$x^2 = 1 \implies x = 1$ or $x = -1$", then that really means "**if** $x^2 = 1$, then $x = 1$ or $x = -1$". And hence this gives no real information about what x is. On the other hand, writing

$$\text{I know } x^2 = 1, \text{ hence } x = 1 \text{ or } x = -1,$$

means that now we know $x = 1$ or $x = -1$ and can use that knowledge in what follows.

Some other unnecessary symbols that are sometimes used are "∴" and "∵". They mean something like "therefore/hence" and "since/because". It is best not to use them, but to write the word instead. It makes things so much easier to read.

Provide all information required.

A good habit is to start by writing what information is given and what question needs to be answered. For instance, suppose you are asked to prove the following:

For any natural numbers a, b, c with $c \geq 2$, there is a natural number n such that $an^2 + bn + c$ is not a prime.

A good start to an answer would be:

Given: natural numbers a, b, c, with $c \geq 2$.
To prove: there is a natural number n such that $an^2 + bn + c$ is not a prime.

At this point you (and any future reader) has all the information required, and you can start thinking what really needs to be done.

2.13.3 How to do mathematics

In a few words: **by trying** and **by doing it yourself**!!

Try hard
The kind of questions you will be dealing with in this subject often have no obvious answers. There is no standard method to come to an answer. That means that you have to find out what to do yourself. And the only way of doing that is by trial and error.

So once you know what you are asked to do (plus all the information you were given), the next thing is to take a piece of paper and start writing down some possible next steps. Some of them may look promising, so have a better look at those and see if they will help you. Hopefully, after some (or a lot) of trying, you see how to answer the question. Then you can go back to writing down the answer. This rough working is a vital part of the process of answering a question (and, in an examination, you should make sure your working is shown). Once you have completed this part of the process, you will then be in a position to write the final answer in a concise form indicating the flow of the reasoning and the arguments used.

Keep trying
You must get used to the situation that not every question can be answered immediately. Sometimes you immediately see what to do and how to do it. But other times you will realise that after a long time you haven't got any further.

Don't get frustrated when that happens. Put the problem aside, and try to do another question (or do something else). Look back at the question later or another day, and see if it makes more sense then. Often the answer will come to you as some kind of "ah-ha" flash. But you can't force these flashes. Spending more time improves the chances they happen, though.

Finally, if you need a long time to answer certain questions, you can consider yourself in good company. For the problem known as "Fermat's Last Theorem", the time between when the problem was first formulated and when the answer was found was about 250 years!

Do it yourself
Here is a possible answer to the previous example:

Given: natural numbers a, b, c, with $c \geq 2$.

To prove: there is a natural number n such that $an^2 + bn + c$ is not a prime.
By definition (see page 70 of Biggs' book, or the footnote on page 4 of Eccles' book), a number p is **prime** if $p \geq 2$ and the only divisors of p are 1 and p itself.
Hence to prove: there is a natural number n for which $an^2 + bn + c$ is smaller than 2 or it has divisors other than 1 or itself.
Let's take $n = c$. Then we have $an^2 + bn + c = ac^2 + bc + c$.
But we can write $ac^2 + bc + c = c\,(ac + b + 1)$, which shows that $ac^2 + bc + c$ has c and $ac + b + 1$ as divisors.
Moreover, it's easy to see that neither c nor $ac+b+1$ can be equal to 1 or to ac^2+bc+c.
We've found a value of n for which $an^2 + bn + c$ has divisors other than 1 or itself. □

The crucial step in the answer above is the one in which I choose to take $n = c$. Why did I choose that? Because it works. How did I get the idea to take $n = c$? Ah, that's far less obvious. Probably some rough paper and lots of trying was involved. In the final answer, no information about how this clever idea was found needs to be given.

You probably have no problems following the reasoning given above, and hence you may think that you understand this problem. But being able to understand the answer, and **being able to find the answer yourself** are two completely different matters. And it is the second skill you are suppose to acquire in this course. (And hence the skill that will be tested in the examination.) Once you have learnt how to approach questions such as the above and come up with the clever trick yourself, you have some hope of being able to answer other questions of a similar type.

But if you only study answers, you will probably never be able to find new arguments for yourself. And hence when you are given a question you've never seen before, how can you trust yourself that you have the ability to see the "trick" that that particular question requires?

For many, abstract mathematics seems full of clever "tricks". But these tricks have always been found by people working very hard to get such a clever idea, not by people just studying other problems and the tricks found by other people.

2.13.4 How to become better in mathematics

One thing you might consider is doing more questions. The books are a good source of exercises. Trying some of these will give you extra practice.

But if you want to go beyond just being able to do what somebody else has written down, you must try to explore the material even further. Try to understand the reason for things that are maybe not explicitly asked.

As an illustration of thinking that way, look again at the formulation of the example we looked at before:

For any natural numbers a, b, c with $c \geq 2$, there is a natural number n such that $an^2 + bn + c$ is not a prime.

Why is it so important that $c \geq 2$? If you look at the proof in the previous section, you see that that proof goes wrong if $c = 1$. (Since we want to use that c is a divisor different from 1.) Does that mean the statement is wrong if $c = 1$? (No, but a different proof is required.)

And what happens if we allow one or more of a, b, c to be zero or negative?

And what about more complicated expression such as $an^3 + bn^2 + cn + d$ for some numbers a, b, c, d with $d \geq 2$? Could it be possible that there is an expression like this for which all n give prime numbers ? If you found the answer to the original question yourself, then you probably immediately see that the answer has to be "no", since similar arguments as before work. But if you didn't try the original question yourself, and just studied the ready-made answer, you'll be less well equipped to answer more general or slightly altered versions.

Once you start thinking like this, you are developing the skills required to be good in mathematics. Trying to see beyond what is asked, asking yourself new questions and seeing which you can answer, is the best way to train yourself to become a mathematician.

2.14 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- demonstrate an understanding of what mathematical statements are
- prove whether mathematical statements are true or false
- negate statements, including universal statements and existential statements
- construct truth tables for logical statements
- use truth tables to determine whether logical statements are logically equivalent or not
- demonstrate knowledge of what is meant by conjunction and disjunction
- demonstrate understanding of the meaning of 'if-then' statements and be able to prove or disprove such statements
- demonstrate understanding of the meaning of 'if and only if' statements and be able to prove or disprove such statements
- find the converse and contrapositive of statements
- prove statements by proving their contrapositive
- prove results by various methods, including directly, by the the method of proof by contradiction, and by working backwards
- demonstrate understanding of the key ideas and notations concerning sets
- prove results about sets
- use existential and universal quantifiers
- be able to negate statements involving several different quantifiers

2.15 Sample exercises

Exercise 2.1

Is the following statement about natural numbers n true or false? Justify your answer by giving a proof or a counterexample:

> If n is divisible by 6 then n is divisible by 3.

What are the converse and contrapositive of this statement? Is the converse true? Is the contrapositive true?

Exercise 2.2

Is the following statement about natural numbers n true or false? Justify your answer by giving a proof or a counterexample:

> If n is divisible by 2 then n is divisible by 4.

What are the converse and contrapositive of this statement? Is the converse true? Is the contrapositive true?

Exercise 2.3

Prove that $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are logically equivalent. □

Exercise 2.4

Prove that the negation of $P \vee Q$ is $\neg P \wedge \neg Q$.

Exercise 2.5

Prove that for all real numbers a, b, c, $ab + ac + bc \leq a^2 + b^2 + c^2$.

Exercise 2.6

Prove by contradiction that there is no largest natural number.

Exercise 2.7

Prove that there is no smallest positive real number.

Exercise 2.8

Suppose A and B are subsets of a universal set E. Prove that

$$(E \times E) \setminus (A \times B) = ((E \setminus A) \times E) \cup (E \times (E \setminus B)).$$

Exercise 2.9

Suppose that $P(x, y)$ is a predicate involving two free variables x, y from a set E. (So, for given x and y, $P(x, y)$ is either true or false.) Find the negation of the statement

$$\exists x \in E, \forall y \in E, P(x, y)$$

2.16 Comments on selected activities

Learning activity 2.2 We can do this by constructing a truth table. Consider Table 2.8. This proves that $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are equivalent.

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$\neg P \vee \neg Q$
T	T	T	**F**	F	F	**F**
T	F	F	**T**	F	T	**T**
F	T	F	**T**	T	F	**T**
F	F	F	**T**	T	T	**T**

Table 2.8: The truth tables for $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$

Learning activity 2.3 The converse is 'if n divides 12 then n divides 4'. This is false. For instance, $n = 12$ is a counterexample. This is because 12 divides 12, but it does not divide 4. The original statement is true, however. For, if n divides 4, then for some $m \in \mathbb{Z}$, $4 = nm$ and hence $12 = 3 \times 4 = 3nm = n(3m)$, which shows that n divides 12.

Learning activity 2.4 We have

$$\begin{aligned} x \in A \setminus B &\iff (x \in A) \wedge (x \notin B) \\ &\iff (x \in A) \wedge (x \in E \setminus B) \\ &\iff x \in A \cap (E \setminus B). \end{aligned}$$

Learning activity 2.5 $\mathcal{P}(A)$ is the set consisting of the following sets:

$$\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\},$$

$$\{1, 2, 3\}, \{2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 4\}, \{1, 2, 3, 4\}.$$

Learning activity 2.6 The members of $\mathcal{P}(A)$ are all the subsets of A. A subset S is determined by which of the n members of A it contains. For each member x of A, either $x \in S$ or $x \notin S$. There are therefore two possibilities, for each $x \in A$. It follows that the number of subsets is $2 \times 2 \times \cdots \times 2$ (where there are n factors, one for each element of A). Therefore $\mathcal{P}(A)$ has 2^n members.

Learning activity 2.7 The statement means that if we take any natural number n there will be some natural number m greater than n. Well, this is true. For example, $m = n + 1$ will do.

2.17 Solutions to exercises

Solution to exercise 2.1

The statement is true. For, suppose n is divisible by 6. Then for some $m \in \mathbb{N}$, $n = 6m$, so $n = 3(2m)$ and since $2m \in \mathbb{N}$, this proves that n is divisible by 3.

The converse is 'If n is divisible by 3 then n is divisible by 6'. This is false. For example, $n = 3$ is a counterexample: it is divisible by 3, but not by 6.

The contrapositive is 'If n is not divisible by 3 then n is not divisible by 6'. This is true, because it is logically equivalent to the initial statement, which we have proved to be true.□

Solution to exercise 2.2

The statement is false. For example, $n = 2$ is a counterexample: it is divisible by 2, but not by 4.

The converse is 'If n is divisible by 4 then n is divisible by 2'. This is true. For, suppose n is divisible by 4. Then for some $m \in \mathbb{N}$, $n = 4m$, so $n = 2(2m)$ and since $2m \in \mathbb{N}$, this proves that n is divisible by 2.

The contrapositive is 'If n is not divisible by 4 then n is not divisible by 2'. This is false, because it is logically equivalent to the initial statement, which we have proved to be false. Alternatively, you can see that it's false because 2 is a counterexample: it is not divisible by 4, but it *is* divisible by 2.

Solution to exercise 2.3

This can be established by using the truth table constructed in Learning activity 2.2. See the solution above.

Solution to exercise 2.4

This is established by Table 2.6. That table shows that $\neg(P \vee Q)$ is logically equivalent to $\neg P \wedge \neg Q$. This is the same as saying that the negation of $P \vee Q$ is $\neg P \wedge \neg Q$.

Solution to exercise 2.5

We work backwards, since it is not immediately obvious how to begin. We note that what we're trying to prove is equivalent to

$$a^2 + b^2 + c^2 - ab - ac - bc \geq 0.$$

This is equivalent to

$$2a^2 + 2b^2 + 2c^2 - 2ab - 2ac - 2bc \geq 0,$$

which is the same as

$$(a^2 - 2ab + b^2) + (b^2 - 2bc + c^2) + (a^2 - 2ac + c^2) \geq 0.$$

You can perhaps now see how this is going to work, for $(a^2 - 2ab + b^2) = (a - b)^2$ and so on. Therefore the given inequality is equivalent to

$$(a - b)^2 + (b - c)^2 + (a - c)^2 \geq 0.$$

We know this to be true because squares are always non-negative. If we wanted to write this proof 'forwards' we might argue as follows. For any a, b, c, $(a - b)^2 \geq 0$, $(b - c)^2 \geq 0$ and $(a - c)^2 \geq 0$, so

$$(a - b)^2 + (b - c)^2 + (a - c)^2 \geq 0$$

and hence

$$2a^2 + 2b^2 + 2c^2 - 2ab - 2ac - 2bc \geq 0,$$

from which we obtain

$$a^2 + b^2 + c^2 \geq ab + ac + bc,$$

as required. □

Solution to exercise 2.6

Let's prove by contradiction that there is no largest natural number. So suppose there is a largest natural number. Let us call it N. (What we want to do now is somehow show that a conclusion, or something we know for sure must be false, follows.) Well, consider the number $N + 1$. This is a natural number. But since N is the largest natural number, we must have $N + 1 \leq N$, which means that $1 \leq 0$, and that's nonsense. So it follows that we must have been wrong in supposing there is a largest natural number. (That's the only place in this argument where we could have gone wrong.) So there is *no* largest natural number. We could have argued the contradiction slightly differently. Instead of using the fact that $N + 1 \leq N$ to obtain the absurd statement that $1 \leq 0$, we could have argued as follows: $N + 1$ is a natural number. But $N + 1 > N$ and this contradicts the fact that N is the largest natural number.

Solution to exercise 2.7

We use a proof by contradiction. Suppose that there is a smallest positive real number and let's call this r. Then $r/2$ is also a real number and $r/2 > 0$ because $r > 0$. But $r/2 < r$, contradicting the fact that r is the smallest positive real number. (Or, we could argue: because $r/2$ is a positive real number and r is the smallest such number, then we must have $r/2 \geq r$, from which it follows that $1 \geq 2$, a contradiction.)

Solution to exercise 2.8

We need to prove that

$$(E \times E) \setminus (A \times B) = ((E \setminus A) \times E) \cup (E \times (E \setminus B)).$$

Now,

$$\begin{aligned}
(x, y) \in (E \times E) \setminus (A \times B) &\iff \neg((x, y) \in A \times B) \\
&\iff \neg((x \in A) \wedge (y \in B)) \\
&\iff \neg(x \in A) \vee \neg(y \in B) \\
&\iff (x \in E \setminus A) \vee (y \in E \setminus B) \\
&\iff ((x, y) \in (E \setminus A) \times E) \vee ((x, y) \in E \times (E \setminus B)) \\
&\iff (x, y) \in ((E \setminus A) \times E) \cup (E \times (E \setminus B)).
\end{aligned}$$

Solution to exercise 2.9

We deal first with the existential quantifier at the beginning of the statement. So, the negation of the statement is

$$\forall x \in E, \neg(\forall y \in E, P(x, y))$$

which is the same as

$$\forall x \in E, \exists y \in E, \ \neg P(x, y).$$

Chapter 3

Natural numbers and proof by induction

☞ Biggs, N. L. *Discrete Mathematics*. Chapter 4.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning*. Chapters 1–4 and 6.

3.1 Introduction

This chapter explores some of the properties of the natural numbers. These will not be new to you, but they shall be explained a little more formally. The chapter also studies a very powerful proof method, known as *proof by induction*. This enables us to prove many universal statements about natural numbers that would be extremely difficult to prove by other means.

3.2 Natural numbers: an axiomatic approach

You know that the natural numbers are the positive integers, and you are certainly very comfortable with them. But suppose an alien visited you and you had to explain to him or her (or 'it', I suppose) what the natural numbers were. How would you do this? Well, you could describe them by the properties they have. (To paraphrase Biggs, we describe what they 'do' rather than what they 'are'.) We have the following *axioms* for the natural numbers. (An axiom is a statement that is assumed to be true.)

1. For all $a, b \in \mathbb{N}$ we have $a + b \in \mathbb{N}$. [Closure under Addition]

2. For all $a, b \in \mathbb{N}$ we have $a \times b \in \mathbb{N}$. [Closure under Multiplication]

3. For all $a, b \in \mathbb{N}$ we have $a + b = b + a$. [Commutative Law for addition]

4. For all $a, b, c \in \mathbb{N}$ we have $(a + b) + c = a + (b + c)$. [Associative Law for addition]

5. For all $a, b \in \mathbb{N}$ we have $a \times b = b \times a$. [Commutative Law for Multiplication]

6. For all $a, b, c \in \mathbb{N}$ we have $(a \times b) \times c = a \times (b \times c)$. [Associative Law for Multiplication]

7. There is a special element of $\mathbb{N}$, denoted by 1, which has the property that for all $n \in \mathbb{N}$ we have $n \times 1 = n$.

8. For all $a, b, c \in \mathbb{N}$, if $a + c = b + c$, then $a = b$. [Additive cancellation]

9. For all $a, b, c \in \mathbb{N}$, if $a \times c = b \times c$, then $a = b$. [Multiplicative Cancellation]

10. For all $a, b, c \in \mathbb{N}$ we have $a \times (b + c) = (a \times b) + (a \times c)$. [Distributive Law]

11. For all $a, b \in \mathbb{N}$, $a < b$ if and only if there is some $c \in \mathbb{N}$ with $a + c = b$.

12. For all $a, b \in \mathbb{N}$, exactly one of the following is true: $a = b$, $a < b$, $b < a$.

We also will write $a\,b$ for $a \times b$.

Other properties of the natural numbers follow from these axioms. (That is, they can be proved assuming these axioms.)

For example, we can prove the following.

P1. For all $a, b, c \in \mathbb{N}$, $(a + b) \times c = (a \times c) + (b \times c)$.

P2. If $a, b \in \mathbb{N}$ satisfy $a \times b = a$, then $b = 1$.

P3. For $a, b, c \in \mathbb{N}$, if $a < b$ and $b < c$, then $a < c$. [Transitivity]

P4. For $a, b, c \in \mathbb{N}$, if $a < b$ then $a + c < b + c$ and $a \times c < b \times c$.

P5. 1 is the least element of $\mathbb{N}$.

P6. 1 is not equal to 1+1.

We don't need to add these to the axioms because they follow from the axioms we already have. (We can **prove** them just from the axioms above.) We'll come back to this later in the chapter.

We have $1 + 1 \neq 1$. We will use the symbol 2 to represent $1 + 1$. It can also be shown that $1 + 1 + 1$ is not equal to 1 or to $1 + 1$. We'll denote it by 3. And so on....

You might find this all a bit weird. We already knew, after all, that $1 + 1 = 2$ and have done for some time!

But the point is that we can **define** the natural numbers axiomatically by a set of rules or axioms and everything else about them follows from those axioms.

3.3 Least and greatest members and the well-ordering principle

If S is a subset of $\mathbb{N}$, then l is a *least member* or *least element* of S if $l \in S$ and, for all $s \in S$, $l \leq s$. The natural number g is a *greatest member* of S if $g \in S$ and, for all $s \in S$, $g \geq s$.

It's quite 'obvious' that any non-empty set of natural numbers has a least member, but it does not follow from the axioms given above. We therefore take this as an additional axiom of the set of natural numbers, to be added to those above. It is known as the *well-ordering principle* or *Least Element Axiom*.

Well-ordering principle (or Least Element Axiom): (Axiom **13.** for the natural numbers): Every non-empty subset of $\mathbb{N}$ has a least element.

This is a very important property of natural numbers (not shared by sets of real numbers, for instance, nor even infinite sets of negative integers).

Activity 3.1 Think of a set of real numbers that has no least member.

3.4 The principle of induction

3.4.1 Proof by induction

One particularly useful principle that follows from the axioms of the natural numbers given above is the following one, known as the *Induction principle*. We can, in fact, take the Induction Principle as one of the axioms of the natural numbers, in place of the well-ordering principle: the two are equivalent. (See Eccles, Section 11.2.) This is, in fact, the approach taken by Biggs. But we shall view the Induction Principle as a consequence of the well-ordering principle along with the other axioms for the natural numbers. (See the end of this chapter for more on this.)

The Induction Principle: Suppose $P(n)$ is a statement involving natural numbers n. Then $P(n)$ is true for all $n \in \mathbb{N}$ if the following two statements are true:

(i) $P(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $P(k) \Rightarrow P(k + 1)$.

Suppose you want to prove $\forall n \in \mathbb{N}, P(n)$. Suppose you can prove two things:

$$P(1) \text{ is true}, \quad \text{and} \quad \forall k \in \mathbb{N},\ P(k) \Rightarrow P(k+1).$$

Then: because $P(1) \Rightarrow P(2)$ and $P(1)$ is true, we must have that $P(2)$ is true. Then, because $P(2) \Rightarrow P(3)$ and $P(2)$ is true, we must have $P(3)$ is true, and so on. So what you see is that this establishes the universal statement that for every $n \in \mathbb{N}$, $P(n)$ is true.

3.4.2 An example

Here's an example of how we might prove by induction a result we proved directly earlier, in the previous chapter, namely:

$$\forall n \in \mathbb{N},\ n^2 + n \text{ is even.}$$

Let $P(n)$ be the statement '$n^2 + n$ is even'. Then $P(1)$ is true, because $1^2 + 1 = 2$, and this is even. (Establishing $P(1)$ is known as proving the *base case* or the *induction basis*.) Next we show that $P(k) \Rightarrow P(k + 1)$ for any $k \in \mathbb{N}$. So we show that if $P(k)$ is true, so will be $P(k + 1)$. To do this we assume that $P(k)$ is true and show that $P(k + 1)$ is then also true. (The assumption that $P(k)$ is true is known as the *inductive hypothesis*.) So suppose $P(k)$ is true, which means that $k^2 + k$ is even. What we need to do now is show that this means that $P(k + 1)$ is also true, namely that $(k + 1)^2 + (k + 1)$ is even. So we need somehow to relate the expression $(k + 1)^2 + (k + 1)$ to the one we are assuming we know something about, $k^2 + k$. Well,

$$(k+1)^2 + (k+1) = k^2 + 2k + 1 + k + 1 = (k^2 + k) + (2k + 2).$$

Now, by the 'inductive hypothesis' (the assumption that $P(k)$ is true), $k^2 + k$ is even. But $2k + 2 = 2(k + 1)$ is also even, so $(k + 1)^2 + (k + 1)$ is an even number, in other words $P(k + 1)$ is true. So we have shown that $\forall k, P(k) \Rightarrow P(k + 1)$. It now follows, by the Principle of Induction, that for all $n \in \mathbb{N}$, $P(n)$ is true.

Once we get used to this technique, we can make our proofs more succinct.

The basic way of proving a result $\forall n \in \mathbb{N},\ P(n)$ by induction is as follows:

- [**The Base Case**] Prove $P(1)$ is true.
- [**The Induction Step**] Prove that, for any $k \in \mathbb{N}$, assuming $P(k)$ is true (the 'inductive hypothesis'), then $P(k + 1)$ is also true.

And that's all you need to do! The principle of induction then establishes that $P(n)$ is true for all $n \in \mathbb{N}$.

3.4.3 Variants

Suppose N is some particular natural number and that $P(n)$ is a statement involving natural numbers n. Then $P(n)$ is true for all $n \geq N$ if the following two statements are true:

(i) $P(N)$ is true;

(ii) For all $k \in \mathbb{N}, k \geq N,\ P(k) \Rightarrow P(k + 1)$.

This is a version of the Induction Principle obtained from the standard one by 'changing the base case'. It can be used to prove a result like the following:

$$\forall n \geq 4, n^2 \leq 2^n.$$

(The inequality $n^2 \leq 2^n$ is false when $n = 3$, so it does not hold for all $n \in \mathbb{N}$.)

> **Activity 3.2** Prove that $\forall n \geq 4, n^2 \leq 2^n$.

Another variant of the Induction Principle is the following, known as the Strong Induction Principle:

The Strong Induction Principle: Suppose $P(n)$ is a statement involving natural numbers n. Then $P(n)$ is true for all $n \in \mathbb{N}$ if the following two statements are true:

(i) $P(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $(P(s)$ true $\forall s \leq k) \Rightarrow P(k + 1)$.

The name is misleading, because, in fact, the strong induction principle follows from the standard induction principle.

> **Activity 3.3** Try to understand why the strong induction principle follows from the induction principle. Hint: consider $Q(n)$, the statement '$\forall s \leq n,\ P(s)$ is true'. [This is difficult, so you may want to omit this activity at first.]

The strong induction principle is, as we shall see, often useful when it comes to proving results about sequences that are defined 'recursively'.

3.5 Summation formulae

Suppose $a_1, a_2, a_3, \ldots$ is a sequence (an infinite, ordered, list) of real numbers. Then the sum $\sum_{r=1}^{n} a_r$ is the sum of the first n numbers in the sequence. It is useful to define these sums 'recursively' or 'by induction', as follows:

$$\sum_{r=1}^{1} a_r = a_1 \quad \text{and} \quad \text{for } n \in \mathbb{N}, \sum_{r=1}^{n+1} a_r = \left(\sum_{r=1}^{n} a_r\right) + a_{n+1}.$$

With this observation, we can use proof by induction to prove many results about the values and properties of such sums. Here is a simple, classical, example.

> **Example 3.1** For all $n \in \mathbb{N}$, $\sum_{r=1}^{n} r = \frac{1}{2}n(n + 1)$. This is simply the statement that the sum of the first n natural numbers is $n(n + 1)/2$.
>
> **Proof.** We prove the result by induction. Let $P(n)$ be the statement that $\sum_{r=1}^{n} r = \frac{1}{2}n(n + 1)$. Then $P(1)$ states that $1 = 1$, which is true. So the base case $P(1)$ is true. Now let's do the induction step. Suppose that $k \in \mathbb{N}$ and that (the inductive hypothesis) $\sum_{r=1}^{k} r = \frac{1}{2}k(k + 1)$ holds. Consider $\sum_{r=1}^{k+1} r$. We have
>
> $$\begin{aligned} \sum_{r=1}^{k+1} r &= \sum_{r=1}^{k} r + (k + 1) \\ &= \frac{1}{2}k(k + 1) + (k + 1) \ \text{ by the induction hypothesis} \\ &= \frac{1}{2}(k^2 + k + 2k + 2) \\ &= \frac{1}{2}(k^2 + 3k + 2) \\ &= \frac{1}{2}(k + 1)(k + 2) \\ &= \frac{1}{2}(k + 1)((k + 1) + 1). \end{aligned}$$
>
> This establishes that $P(k + 1)$ is true (for $P(k + 1)$ is precisely the statement that $\sum_{r=1}^{k+1} r = (k + 1)((k + 1) + 1)/2$.) Therefore, by induction, for all $n \in \mathbb{N}$, $\sum_{r=1}^{n} r = \frac{1}{2}n(n + 1)$.

Note how the the induction hypothesis was used. In the induction step, you always prove $P(k + 1)$ to be true assuming $P(k)$ is. (Unless you do so, it isn't a proof by induction.)

> **Activity 3.4** Prove by induction that the sum of the first n terms of an arithmetic progression with first term a and common difference d is $n(2a + (n - 1)d)/2$.

3.6 Recursively defined sequences

Sequences of numbers are often defined ‘recursively’ or ‘by induction’.

For example, suppose that the sequence x_n is given by $x_1 = 9$, $x_2 = 13$ and, for $n \geq 3$, $x_n = 3x_{n-1} - 2x_{n-2}$. We can prove by induction (using the strong induction principle) that, for all $n \in \mathbb{N}$, $x_n = 5 + 2^{n+1}$. Here’s how:

Since the inductive definition for x_n only applies for $n \geq 3$, the case step of our proof is to verify the result for the cases $n = 1$ and $n = 2$. Now, when $n = 1$, $5 + 2^{n+1} = 9$, which is indeed x_1; and when $n = 2$, $5 + 2^{n+1} = 13$, which equals x_2, so these hold. Assume inductively that $k \in \mathbb{N}$ and that, for all $s \leq k$, $x_s = 5 + 2^{s+1}$. (Note that, here, we use strong induction. This is because x_{k+1} depends not only on x_k but on x_{k-1} too.) In particular, therefore, we have $x_k = 5 + 2^{k+1}$ and $x_{k-1} = 5 + 2^k$. So,

$$\begin{aligned}
x_{k+1} &= 3x_k - 2x_{k-1} \\
&= 3(5 + 2^{k+1}) - 2(5 + 2^k) \\
&= 15 - 10 + 3(2^{k+1}) - 10 - 2(2^k) \\
&= 5 + 3(2^{k+1}) - 2(2^k) \\
&= 5 + 6(2^k) - 2(2^k) \\
&= 5 + 4(2^k) \\
&= 5 + 2^{k+2} \\
&= 5 + 2^{(k+1)+1},
\end{aligned}$$

which is exactly what we need. So the formula for x_n holds for all n.

3.7 Using the axioms for the natural numbers

Earlier, we said that the following results follow from the axioms for $\mathbb{N}$.

P1. For all $a, b, c \in \mathbb{N}$, $(a + b) \times c = (a \times c) + (b \times c)$.

P2. If $a, b \in \mathbb{N}$ satisfy $a \times b = a$, then $b = 1$.

P3. For $a, b, c \in \mathbb{N}$, if $a < b$ and $b < c$, then $a < c$. [Transitivity]

P4. For $a, b, c \in \mathbb{N}$, if $a < b$ then $a + c < b + c$ and $a \times c < b \times c$.

P5. 1 is the least element of $\mathbb{N}$.

P6. 1 is not equal to 1+1.

Let’s see why.

Proof of P1. For all $a, b, c \in \mathbb{N}$, $(a + b) \times c = (a \times c) + (b \times c)$.

$(a + b) \times c = c \times (a + b)$ by **5** [Commutative]

$= (c \times a) + (c \times b)$ by **10** [Distributive]

$= (a \times c) + (b \times c)$ by **5** [Commutative]

Proof of P2. If $a, b \in \mathbb{N}$ satisfy $a \times b = a$, then $b = 1$.

Suppose $a \times b = a$. Then, since (by **7**), $a = a \times 1$, so

$$a \times b = a \times 1.$$

By **5** [Commutative],

$$b \times a = 1 \times a.$$

Then, by **9** [Cancellation],

$$b = 1.$$

Proof of P3. For $a, b, c \in \mathbb{N}$, if $a < b$ and $b < c$, then $a < c$. [Transitivity]

If $a < b$ and $b < c$ then, by **11**, there are $x, y \in \mathbb{N}$ such that $a + x = b$ and $b + y = c$. Then,

$a + (x + y) = (a + x) + y$ by **4** [Associativity]

$= b + y = c$

and so, by **11**, together with the fact (by Closure, **1**) that $x + y \in \mathbb{N}$, $a < c$.

Proof of P4. For $a, b, c \in \mathbb{N}$, if $a < b$ then $a + c < b + c$ and $a \times c < b \times c$.

If $a < b$ then (by **11**) this means $\exists d \in \mathbb{N}$ with $a + d = b$.

(i) We prove $a + c < b + c$. We have:
$(a + c) + d = d + (a + c)$ by **3**
$= (d + a) + c$ by **4**
$= (a + d) + c$ by **3**
$= b + c$.
This shows $a + c < b + c$.

(ii) We prove $a \times c < b \times c$. We have:
$(a \times c) + (d \times c) = (a + d) \times c$. This is **P1** from above.
So, since $a + d = b$, we have $(a \times c) + (d \times c) = b \times c$.

Since $d \times c \in \mathbb{N}$ (by **2**), we have a natural number z ($z = d \times c$) such that $(a \times c) + z = (b \times c)$. So, by **11**, $a \times c < b \times c$.

Proof of P5. 1 is the least element of $\mathbb{N}$.

Certainly, by **13** [Well-ordering], $\mathbb{N}$ has a least member. Call it a. Suppose $a \neq 1$. Axiom **12** says $a < 1$ or $a = 1$ or $1 < a$. We are assuming $a \neq 1$ and can’t have $1 < a$ by minimality of a. So $a < 1$. By **P4**,

$$a \times a < 1 \times a.$$

But (Commutativity, **5**), $1 \times a = a \times 1 = a$ (by **7**). So:

$$a \times a < a.$$

But $a \times a \in \mathbb{N}$ (by Closure, **2**) and this contradicts the minimality of a.

Proof of P6. 1 is not equal to 1+1.

11 says $a < b$ if and only if there is some $c \in \mathbb{N}$ with $a + c = b$. So $1 < 1 + 1$.

But **12** says that for all $a, b \in \mathbb{N}$, exactly one of the following is true: $a = b$, $a < b$, $b < a$. So we do not have $1 = 1 + 1$.

3.8 Why the Principle works

We can now **Prove** the Principle of Induction from our axioms for the natural numbers (including the least element axiom).

Theorem 3.1 (The Induction Principle) Suppose

(i) $P(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $P(k) \Rightarrow P(k+1)$.

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

Suppose it's not the case that $P(n)$ is true for all $n \in \mathbb{N}$. Then the set S of $n \in \mathbb{N}$ for which it is not true is non-empty and, by the Least Element Axiom (**13**), has a least member a. Now, $a \neq 1$ because $P(1)$ is true. And we can't have $a < 1$ because (**P5**) 1 is the least member of $\mathbb{N}$. So, by **12**, $1 < a$. Consider $a - 1$: this is a natural number less than a and is therefore not in S. So $P(a-1)$ is true. But since $P(k) \Rightarrow P(k+1)$ for all k, it follows that $P(a)$ is true, meaning $a \notin S$, a contradiction.

You might not be entirely satisfied with that proof. It used $a - 1$ but we haven't defined subtraction! Here's another way of explaining the last bit:

Since $1 < a$, there is some $c \in \mathbb{N}$ such that $1 + c = a$ (by Axiom **11**). By Axiom **5** (Commutativity), $c + 1 = a$, which, by **11**, means $c < a$. So, because a is the least element of S, $c \notin S$ and hence $P(c)$ is true. But $P(c) \Rightarrow P(c+1)$ and hence $P(a)$ is true, a contradiction.

3.9 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- understand how the natural numbers can be defined by axioms and understand that other properties of natural numbers can be proved from the axioms
- state what is meant by a greatest and least member of a set of natural numbers and know what is meant by the well-ordering principle (or least element axiom)
- state the Induction Principle and its variants
- use Proof by Induction to prove a range of statements, including those involving summation and recursive sequences

3.10 Sample exercises

Exercise 3.1
Prove by induction that, for all $n \in \mathbb{N}$, $2^n \geq n + 1$.

Exercise 3.2
Prove by induction that the sum $a + ar + ar^2 + \cdots + ar^{n-1}$ of the first n terms of a geometric progression with first term a and common ratio $r \neq 1$ is $a(1 - r^n)/(1 - r)$. □

Exercise 3.3
Prove by induction that for all $n \in \mathbb{N}$,

$$\sum_{r=1}^{n} r^2 = \frac{1}{6} n(n+1)(2n+1).$$

Exercise 3.4
Prove by induction that $\displaystyle\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$.

Exercise 3.5
Suppose the sequence x_n is given by $x_1 = 7$, $x_2 = 23$ and, for $n \geq 3$, $x_n = 5x_{n-1} - 6x_{n-2}$. Prove by induction that, for all $n \in \mathbb{N}$, $x_n = 3^{n+1} - 2^n$.

Exercise 3.6
Prove by induction that, for all $n \in \mathbb{N}$, $2^{n+2} + 3^{2n+1}$ is divisible by 7.

Exercise 3.7
For a sequence of numbers $x_1, x_2, x_3, \ldots$, and for $n \in \mathbb{N}$, the number $\prod_{r=1}^{n} x_r$ is the product of the first r numbers of the sequence. It can be defined inductively as follows:

$$\prod_{r=1}^{1} x_r = x_1, \quad \text{and} \quad \text{for } k \geq 1, \prod_{r=1}^{k+1} x_r = \left(\prod_{r=1}^{k} x_r \right) x_{k+1}.$$

Suppose that $x \neq 1$. Prove that

$$\prod_{r=1}^{n} (1 + x^{2^{r-1}}) = \frac{1 - x^{2^n}}{1 - x}.$$

3.11 Comments on selected activities

Learning activity 3.2 When $n = 4$, $n^2 = 16$ and $2^n = 2^4 = 16$, so in this base case, the statement is true. Suppose we make the inductive hypothesis that for some $k \geq 4$, $k^2 \leq 2^k$. We want to show

$$(k+1)^2 \leq 2^{k+1}.$$

We have

$$(k+1)^2 = k^2 + 2k + 1 \leq 2^k + 2k + 1$$

(by the inductive hypothesis). So we'll be done if we can show that $2k + 1 \leq 2^k$. This will follow from $2k + 1 \leq k^2$ and the assumed fact that $k^2 \leq 2^k$. Now,

$$2k + 1 \leq k^2 \iff k^2 - 2k - 1 \geq 0 \iff (k-1)^2 \geq 2,$$

which is true for $k \geq 4$. So, finally,

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

as required. So the result is true for all $n \geq 4$.

Learning activity 3.3 Let $Q(n)$ be the statement '$\forall s \leq n,\ P(s)$ is true'. Then $Q(1)$ is true if and only if $P(1)$ is true. The statement $Q(k) \Rightarrow Q(k+1)$ is the same as

$$(P(s) \text{ true } \forall s \leq k) \Rightarrow (P(s) \text{ true } \forall s \leq k+1).$$

But if $P(s)$ is true for all $s \leq k$ then its truth for all $s \leq k+1$ follows just from its truth when $s = k+1$. That is, $Q(k) \Rightarrow Q(k+1)$ is the same as $(P(s) \text{ true } \forall s \leq k) \Rightarrow P(k+1)$. The (standard) Induction Principle applied to the statement $Q(n)$ tells us that: $Q(n)$ is true for all $n \in \mathbb{N}$ if the following two statements are true:

(i) $Q(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $Q(k) \Rightarrow Q(k+1)$.

What we've established is that (i) and (ii) can be rewritten as:

(i) $P(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $(P(s) \text{ true } \forall s \leq k) \Rightarrow P(k+1)$.

We deduce that: $P(n)$ is true for all $n \in \mathbb{N}$ if the following two statements are true:

(i) $P(1)$ is true;

(ii) For all $k \in \mathbb{N}$, $(P(s) \text{ true } \forall s \leq k) \Rightarrow P(k+1)$.

This is exactly the Strong Induction Principle. So the Strong Induction Principle follows from the standard one and is, therefore, not really 'stronger'.

Learning activity 3.4 Let $P(n)$ be the statement that the sum of the first n terms is $(n/2)(2a + (n-1)d)$. The base case is straightforward. The first term is a, and the formula $(n/2)(2a + (n-1)d)$ gives a when $n = 1$. Suppose that $P(k)$ holds, so the sum of the first k terms is $(k/2)(2a + (k-1)d)$. Now, the $(k+1)$st term is $a + kd$, so the sum of the first $k+1$ terms is therefore

$$\begin{aligned}
a + kd + \frac{k}{2}(2a + (k-1)d) &= a + kd + ak + \frac{k(k-1)}{2}d \\
&= (k+1)a + \frac{k(k+1)}{2}d \\
&= \frac{(k+1)}{2}(2a + kd) \\
&= \frac{(k+1)}{2}(2a + ((k+1)-1)d),
\end{aligned}$$

so $P(k+1)$ is true. The result follows for all n by induction.

3.12 Solutions to exercises

Solution to exercise 3.1

Let $P(n)$ be the statement '$2^n \geq n+1$'. When $n = 1$, $2^n = 2$ and $n + 1 = 2$, so $P(1)$ is true. Suppose $P(k)$ is true for some $k \in \mathbb{N}$. Then $2^k \geq k+1$. It follows that

$$2^{k+1} = 2.2^k \geq 2(k+1) = 2k + 2 \geq k + 2 = (k+1) + 1,$$

so $P(k+1)$ is also true. Hence, by induction, for all $n \in \mathbb{N}$, $2^n \geq n+1$. □

Solution to exercise 3.2

Let $P(n)$ be the statement that the sum of the first n terms is $a(1 - r^n)/(1-r)$. $P(1)$ states that the first term is $a(1 - r^1)/(1-r) = a$, which is true. Suppose $P(k)$ is true. Then the sum of the first $k+1$ terms is the sum of the first k plus the $(k+1)$st term, which is ar^k, so this sum is

$$\begin{aligned}
\frac{a(1-r^k)}{1-r} + ar^k &= \frac{a(1-r^k) + (1-r)ar^k}{1-r} \\
&= \frac{a - ar^k + ar^k - ar^{k+1}}{1-r} \\
&= \frac{a(1-r^{k+1})}{1-r},
\end{aligned}$$

which shows that $P(k+1)$ is true. Hence, for all $n \in \mathbb{N}$, $P(n)$ is true, by induction. □

Solution to exercise 3.3

Let $P(n)$ be the statement that

$$\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1).$$

Then $P(1)$ states that $1 = 1(2)(3)/6$, which is true. Suppose $P(k)$ is true for $k \in \mathbb{N}$. Then

$$\sum_{r=1}^{k} r^2 = \frac{1}{6}k(k+1)(2k+1)$$

and $P(k+1)$ is the statement that

$$\sum_{r=1}^{k+1} r^2 = \frac{1}{6}(k+1)(k+2)(2(k+1)+1) = \frac{1}{6}(k+1)(k+2)(2k+3).$$

We have

$$\begin{aligned}
\sum_{r=1}^{k+1} r^2 &= (k+1)^2 + \sum_{r=1}^{k} r^2 \\
&= (k+1)^2 + \frac{1}{6}k(k+1)(2k+1) \quad \text{(by the induction hypothesis)} \\
&= \frac{1}{6}(k+1)\left[6(k+1)+k(2k+1)\right] \\
&= \frac{1}{6}(k+1)\left(2k^2+7k+6\right) \\
&= \frac{1}{6}(k+1)(k+2)(2k+3),
\end{aligned}$$

so $P(k+1)$ is true. By induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

Solution to exercise 3.4

Let $P(n)$ be the statement that $\displaystyle\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$. Then $P(1)$ states that $\dfrac{1}{1 \times 2} = \dfrac{1}{1+1}$, which is true. Suppose $P(k)$ is true for $k \in \mathbb{N}$. Then

$$\sum_{i=1}^{k} \frac{1}{i(i+1)} = \frac{k}{k+1}$$

and $P(k+1)$ is the statement that

$$\sum_{i=1}^{k+1} \frac{1}{i(i+1)} = \frac{k+1}{k+2}.$$

Now,

$$\begin{aligned}
\sum_{i=1}^{k+1} \frac{1}{i(i+1)} &= \frac{1}{(k+1)(k+2)} + \sum_{i=1}^{k} \frac{1}{i(i+1)} \\
&= \frac{1}{(k+1)(k+2)} + \frac{k}{k+1} \quad \text{(by the induction hypothesis)} \\
&= \frac{1+k(k+2)}{(k+1)(k+2)} \\
&= \frac{k^2+2k+1}{(k+1)(k+2)} \\
&= \frac{(k+1)^2}{(k+1)(k+2)} \\
&= \frac{k+1}{k+2},
\end{aligned}$$

so $P(k+1)$ is true. By induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

Solution to exercise 3.5

Let $P(n)$ be the statement that $x_n = 3^{n+1} - 2^n$. We use the Strong Induction Principle to prove $P(n)$ is true for all $n \in \mathbb{N}$. The base cases are $n = 1$ and $n = 2$. When $n = 1$, $x_1 = 7$ and $3^{n+1} - 2^n = 9 - 2 = 7$. When $n = 2$, $x_2 = 23$ and $3^{n+1} - 2^n = 27 - 4 = 23$, so these are true. Suppose that $k \geq 2$ and that for all $s \leq k$, $P(s)$ is true. In particular, $P(k)$ and $P(k-1)$ are true and so

$$\begin{aligned}
x_{k+1} &= 5x_k - 6x_{k-1} \\
&= 5(3^{k+1} - 2^k) - 6(3^k - 2^{k-1}) \\
&= 5(3^{k+1}) - 5(2^k) - 6(3^k) + 6(2^{k-1}) \\
&= 15(3^k) - 6(3^k) - 10(2^{k-1}) + 6(2^{k-1}) \\
&= 9(3^k) - 4(2^{k-1}) \\
&= 3^{k+2} - 2^{k+1} \\
&= 3^{(k+1)+1} - 2^{k+1},
\end{aligned}$$

so $P(k+1)$ is true. Therefore, $P(n)$ is true for all $n \in \mathbb{N}$. □

Solution to exercise 3.6

Let $P(n)$ be the statement that $2^{n+2} + 3^{2n+1}$ is divisible by 7. When $n = 1$, $2^{n+2} + 3^{2n+1} = 8 + 27 = 35$ and this is a multiple of 7 because $35 = 5 \times 7$. Suppose $P(k)$ is true, which means that for some $m \in \mathbb{N}$, $2^{k+2} + 3^{2k+1} = 7m$. Now, when we take $n = k+1$,

$$\begin{aligned}
2^{n+2} + 3^{2n+1} &= 2^{k+3} + 3^{2k+3} \\
&= 2(2^{k+2}) + 9(3^{2k+1}) \\
&= 2(2^{k+2} + 3^{2k+1}) + 7(3^{2k+1}) \\
&= 7m + 7(3^{2k+1}) \\
&= 7\left(m + 3^{2k+1}\right),
\end{aligned}$$

which is a multiple of 7. So the statement is true for $P(k+1)$ and hence, by induction, for all $n \in \mathbb{N}$. □

Solution to exercise 3.7

Let $P(n)$ be the statement

$$\prod_{r=1}^{n} (1 + x^{2^{r-1}}) = \frac{1 - x^{2^n}}{1 - x}.$$

When $n = 1$, the left hand side is $1 + x^{2^0} = 1 + x$ and the right hand side is $(1 - x^2)/(1 - x) = 1 + x$, so $P(1)$ is true. Suppose $P(k)$ is true, so that

$$\prod_{r=1}^{k} (1 + x^{2^{r-1}}) = \frac{1 - x^{2^k}}{1 - x}.$$

Then

$$\begin{aligned}
\prod_{r=1}^{k+1} (1 + x^{2^{r-1}}) &= (1 + x^{2^{(k+1)-1}}) \times \prod_{r=1}^{k} (1 + x^{2^{r-1}}) \\
&= (1 + x^{2^k}) \frac{1 - x^{2^k}}{1 - x} \quad \text{(by the induction hypothesis)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1-(x^{2^k})^2}{1-x} \quad \text{(where we've used } (1+y)(1-y)=1-y^2\text{)}\\
&= \frac{1-x^{2^k\times 2}}{1-x}\\
&= \frac{1-x^{2^{k+1}}}{1-x},
\end{aligned}
$$

which shows that $P(k+1)$ is true. So $P(n)$ is true for all $n \in \mathbb{N}$, by induction.

Chapter 4 Functions and counting

☞ Biggs, N. L. *Discrete Mathematics.* Chapters 5 and 6.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning.* Chapter 10, Sections 10.1 and 10.2, and Chapter 11.

4.1 Introduction

In this chapter we look at the theory of functions, and we see how the idea of the 'size' of a set can be formalised.

4.2 Functions

4.2.1 Basic definitions

You have worked extensively with functions in your previous mathematical study. Chiefly, you will have worked with functions from the real numbers to the real numbers, these being the primary objects of interest in calculus.

We'll now take a more abstract approach to functions and their properties.

Here is the definition of what we mean by a function.

Definition 4.1 Suppose that X and Y are sets. Then a *function* (also known as a *mapping*) from X to Y is a rule that associates a unique member of Y to each member of X. We write $f : X \to Y$. The set X is called the *domain* of f and Y is called the *codomain.*

The element of Y that is assigned to $x \in X$ is denoted by $f(x)$ and is called the *image* of x. We can write $x \mapsto f(x)$ to indicate that x maps to $f(x)$.

There are various ways of describing a function. Sometimes, if X has only finitely many members, we can simply list the images of the members of X. More usually, we can give a formula for the function. For instance, $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ is the function that maps each real number a to the real number $2a$.

Sometimes a function can be defined *recursively.* For example, we might define $f : \mathbb{N} \to \mathbb{N}$ by

$$f(1) = 1 \quad f(n) = 2 + 3f(n-1), \;\; (n \geq 2).$$

(You can see that the sequence of numbers $f(1), f(2), f(3), \ldots$ is therefore given by a first order difference equation.)

What does it mean to say that two function f, g are equal? Well, first, they must have the same domain X and codomain Y. Then, for each $x \in X$, we must have $f(x) = g(x)$. For example, if $\mathbb{R}^+$ is the set of positive real numbers, then the function $f : \mathbb{R}^+ \to \mathbb{R}$ given by $f(x) = x^2$ and the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^2$ are *not* equal because their domains are different.

For any set X, the *identity* function $i : X \to X$ is given by $i(x) = x$.

4.2.2 Composition of functions

Suppose that X, Y, Z are sets and that $f : X \to Y$ and $g : Y \to Z$. Then the *composition* gf, also denoted by $g \circ f$, is the function from X to Z given by

$$(gf)(x) = g(f(x)) \quad (x \in X).$$

Note the notation, which can sometimes cause confusion. For example, suppose $X = Y = Z = \mathbb{R}$. Then you might be tempted to think that gf denotes the *product* function $(gf)(x) = g(x)f(x)$. But this would be wrong. It should always be clear from the context whether gf should be interpreted as a composition. Be aware of this. If I need to talk about the product of the functions f and g I will denote this by $f(x)g(x)$. The notation $g \circ f$ leads to less confusion, but it is not used in all textbooks.

Example 4.1 Suppose $f : \mathbb{N} \to \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$ are given by $f(x) = x^2 + 1$ and $g(x) = (x+1)^2$. Then,

$$(fg)(x) = f(g(x)) = f((x+1)^2) = ((x+1)^2)^2 + 1 = (x+1)^4 + 1.$$

And,

$$(gf)(x) = g(f(x)) = g(x^2+1) = ((x^2+1)+1)^2 = (x^2+2)^2.$$

4.3 Bijections, surjections and injections

There are three very important properties that a function might possess:

Definition 4.2 (Surjection) Suppose f is a function with domain X and codomain Y. Then f is said to be a *surjection* (or '*f is surjective*') if every $y \in Y$ is the image of some $x \in X$; that is, f is a surjection if and only if $\forall y \in Y, \exists x \in X, \text{s.t. } f(x) = y$.

Definition 4.3 (Injection) Suppose f is a function with domain X and codomain Y. Then f is said to be an *injection* (or '*f is injective*') if every $y \in Y$ is the image of *at most one* $x \in X$. In other words, the function is an injection if different elements of X have different images under f. Thus, f is an injection if and only if

$$\forall x, x' \in X, x \neq x' \Rightarrow f(x) \neq f(x')$$

or (equivalently, taking the contrapositive), if and only if

$$\forall x, x' \in X, f(x) = f(x') \Rightarrow x = x'.$$

This latter characterisation often provides the easiest way to verify that a function is an injection.

Definition 4.4 (Bijection) Suppose f is a function with domain X and codomain Y. Then f is said to be a *bijection* (or '*f is bijective*) if it is *both* an injection and a surjection. So this means two things: each $y \in Y$ is the image of some $x \in X$, and each $y \in Y$ is the image of no more than one $x \in X$. Well, of course, this is equivalent to: each $y \in Y$ is the image of *precisely one* $x \in X$.

Example 4.2 $f : \mathbb{N} \to \mathbb{N}$ given by $f(x) = 2x$ is not a surjection, because there is no $n \in \mathbb{N}$ such that $f(n) = 1$. (For, $2n = 1$ has no solution where $n \in \mathbb{N}$.) However, it is an injection. To prove this, suppose that $m, n \in \mathbb{N}$ and $f(m) = f(n)$. Then $2m = 2n$, which implies $m = n$.

Example 4.3 $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ is a bijection.

Activity 4.1 Prove that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ is a bijection.

4.3.1 An example

Let $X = \mathbb{R}$, the set of real numbers, and let Y be the interval $(-1, 1)$, the set of real numbers x such that $-1 < x < 1$. Then the function $f : X \to Y$ given by

$$f(x) = \frac{x}{1+|x|}$$

is a bijection from X to Y.

First, we prove f is **injective.** To do this, we prove that $f(x) = f(y)$ implies $x = y$. So, suppose $f(x) = f(y)$. Then

$$\frac{x}{1+|x|} = \frac{y}{1+|y|}.$$

So

$$x + x|y| = y + y|x|.$$

Because $x/(1+|x|) = y/(1+|y|)$, x and y are *both* non-negative or *both* negative. For, otherwise, one of $x/(1+|x|)$ and $y/(1+|y|)$ will be negative and the other one will be non-negative, which cannot be the case since they are equal. So, $x|y| = y|x|$, both being xy if $x, y \geq 0$ and $-xy$ if $x, y < 0$. So, we have $x = y$.

Next, we show f is **surjective.** We need to prove that, for each $y \in (-1, 1)$, there's $x \in \mathbb{R}$ such that $x/(1+|x|) = y$. Consider separately the case in which $y \geq 0$ and the case in which $y < 0$.

$y \geq 0$. Then, to have $x/(1+|x|) = y$, need $x \geq 0$. So $|x| = x$ and we need to solve $x/(1+x) = y$. This has solution $x = y/(1-y)$.

$y < 0$. Then we'll need to have $x < 0$ and the equation to solve is $x/(1-x) = y$, which has solution $x = y/(1+y)$.

4.4 Inverse functions

4.4.1 Definition, and existence

Suppose $f : X \to Y$. Then $g : Y \to X$ is an *inverse function* of f if $(gf)(x) = x$ for all $x \in X$ and $(fg)(y) = y$ for all $y \in Y$. An equivalent characterisation is that $y = f(x) \iff x = g(y)$.

The following theorem tells us precisely when a function has an inverse. It also tells us that if an inverse exists, then there is only one inverse. For this reason we can speak of *the* inverse function, and give it a specific notation, namely f^{-1}.

Theorem 4.1 $f : X \to Y$ has an inverse function if and only if f is a bijection. When f is bijective, there is a unique inverse function.

First, we prove:

$f : X \to Y$ has an inverse $\iff f$ is bijective.

Proof. This is an $\iff$ theorem, so there are two things to prove: the $\Leftarrow$ and the $\Rightarrow$.

First, we show: $f : X \to Y$ has an inverse $\Leftarrow f$ is bijective.

Suppose f is a bijection. For each $y \in Y$ there is exactly one $x \in X$ with $f(x) = y$. Define $g : Y \to X$ by $g(y) = x$. Then this is an inverse of f. Check this!

Next, we show: $f : X \to Y$ has an inverse $\Rightarrow f$ is bijective.

Suppose f has an inverse function g. We know that for any $y \in Y$, $f(g(y)) = (fg)(y) = y$, so there is some $x \in X$ (namely $x = g(y)$) such that $f(x) = y$. So f is surjective.

Now suppose $f(x) = f(x')$. Then $g(f(x)) = g(f(x'))$. But $g(f(x)) = (gf)(x) = x$ and, similarly, $g(f(x')) = x'$. So: $x = x'$ and f is injective.

Now we prove that when f is bijective, the inverse is unique.

Suppose that g and h are inverses of f. Then hf is the identity function on X and fg is the identity function on Y. So, for any $y \in Y$,

$$\begin{aligned} g(y) &= (hf)(g(y)) = ((hf)g)(y) \\ &= (h(fg))(y) = h((fg)(y)) = h(y), \end{aligned}$$

so $g = h$. □

Note that if $f : X \to Y$ is a bijection, then its inverse function (which exists, by Theorem 4.1) is also a bijection.

4.4.2 Examples

Example 4.4 The function $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = 3x + 1$. Find the inverse function.

To find a formula for f^{-1}, we use: $y = f(x) \iff x = f^{-1}(y)$. Now,

$$y = f(x) \iff y = 3x + 1 \iff x = (y - 1)/3,$$

so

$$f^{-1}(y) = \frac{1}{3}(y - 1).$$

Let $\mathbb{Z}$ denote the set of all integers (positive, zero, and negative).

Example 4.5 The function $f : \mathbb{Z} \to \mathbb{N} \cup \{0\}$ is defined as follows:

$$f(n) = \begin{cases} 2n & \text{if } n \geq 0 \\ -2n - 1 & \text{if } n < 0. \end{cases}$$

Prove that f is a bijection and determine a formula for the inverse function f^{-1}.

First, we prove that f is **injective**: Suppose $f(n) = f(m)$. Since $2n$ is even and $-2n - 1$ is odd, either (i) $n, m \geq 0$ or (ii) $n, m < 0$. (For otherwise, one of $f(n), f(m)$ is odd and the other even, and so they cannot be equal.)

In case (i), $f(n) = f(m)$ means $2n = 2m$, so $n = m$.

In case (ii), $f(n) = f(m)$ means $-2n - 1 = -2m - 1$, so $n = m$. Therefore f is injective.

Next, we prove that f is **surjective**: We show that $\forall m \in \mathbb{N} \cup \{0\}, \exists n \in \mathbb{Z}$ such that $f(n) = m$. Consider separately the case m even and the case m odd.

Suppose m is even. Then $n = m/2$ is a non-negative integer and $f(n)$ is $2(m/2) = m$).

If m odd, then $n = -(m + 1)/2$ is a negative integer and

$$f(n) = f(-(m + 1)/2) = -2\left(-\frac{(m + 1)}{2}\right) - 1 = m.$$

The proof that f is surjective reveals to us what the inverse function is. We have

$$f^{-1}(m) = \begin{cases} m/2 & \text{if } m \text{ even} \\ -(m + 1)/2 & \text{if } m \text{ odd.} \end{cases}$$

4.5 Counting as a bijection

What does it mean to say that a set has three objects? Well, it means that I can take an object from the set, and call that ‘Object 1’, then I can take a different object from the set and call that ‘Object 2’, and then I can take a different object from the set and call that ‘Object 3’, and then there will be no objects left in the set without a name. Obvious, I know, but this is the fundamental way in which we can abstractly define what we mean by saying that a set has m members.

For $m \in \mathbb{N}$, let $\mathbb{N}_m$ be the set $\{1, 2, \ldots, m\}$ consisting of the first m natural numbers. Then we can make the following formal definition:

Definition 4.5 A set S has m members if there is a bijection from $\mathbb{N}_m$ to S.

So, the set has m members if to each number from 1 to m, we can assign a corresponding member of the set S, and all members of S are accounted for in this process. This is like the attachment of labels 'Object 1', etc, described above.

Note that an entirely equivalent definition is to say that S has m members if there is a bijection from S to $\mathbb{N}_m$. This is because if $f : \mathbb{N}_m \to S$ is a bijection, then the inverse function $f^{-1} : S \to \mathbb{N}_m$ is a bijection also. In fact, because of this, we can simply say that S has m members if there is a bijection 'between' $\mathbb{N}_m$ and S. (Eccles uses the definition that involves a bijection from $\mathbb{N}_m$ to S and Biggs uses the definition that involves a bijection from S to N_m.)

For $m \in \mathbb{N}$, if S has m members, we say that S has *cardinality* m (or *size* m). The cardinality of S is denoted by $|S|$.

4.6 The pigeonhole principle

4.6.1 The principle

The 'pigeonhole principle' is something that you might find obvious, but it is very useful.

Informally, what it says is that says is that if you have n letters and you place them into m pigeonholes in such a way that no pigeonhole contains more than one letter, then $n \leq m$. Equivalently, if $n > m$ (so that you have more letters than pigeonholes), then some pigeonhole will end up containing more than one letter. This is very intuitive. Obvious as it may be, however, can you think about how you would actually prove it? We shall prove it below. But let's state the principle more formally, first. Recall that, for $r \in \mathbb{N}$, $\mathbb{N}_r = \{1, 2, \ldots, r\}$.

Theorem 4.2 (Pigeonhole Principle (PP)) Let m be a natural number. Then the following statement is true for all $n \in \mathbb{N}$: if there is an injection from $\mathbb{N}_n$ to $\mathbb{N}_m$, then $n \leq m$.

Proof. We prove this by induction. The statement we want to prove is the statement $P(n)$: 'if there is an injection from $\mathbb{N}_n$ to $\mathbb{N}_m$, then $n \leq m$.' The base case, $n = 1$, is true because (since $m \in \mathbb{N}$), $1 \leq m$. Suppose $P(k)$ is true. We now want to show that $P(k+1)$ is also true. So suppose there is an injection $f : \mathbb{N}_{k+1} \to \mathbb{N}_m$. (What we want to show is that $k+1 \leq m$.) Since $k \geq 1$, $k+1 \geq 2$. So m must be at least 2. (If m was 1, then, for example, $f(1)$ and $f(2)$ would be equal, and f would not be an injection.) Since $m \geq 2$ we can write m as $m = s+1$ where $s \in \mathbb{N}$. Now, either there is some $x \in \mathbb{N}_k = \{1, 2, \ldots, k\}$ with $f(x) = s+1$, or there is not. Let's examine each case separately.

- Suppose, then, first, that for no $x \in \mathbb{N}_k$ do we have $f(x) = s+1$. Then define $f_* : \mathbb{N}_k \to \mathbb{N}_s$ by $f_*(x) = f(x)$ for $x \in \mathbb{N}_k$. Then, because f is an injection, so too is f_*. So there is an injection (namely, f_*) from $\mathbb{N}_k$ to $\mathbb{N}_s$. By the induction hypothesis, therefore, $k \leq s$ and hence $k+1 \leq s+1 = m$, as required.
- Now suppose that there is some $j \in \mathbb{N}_k$ such that $f(j) = s+1$. Then the value $y = f(k+1)$ must be different from $s+1$ and therefore $y \in \mathbb{N}_s$. Define $f_* : \mathbb{N}_k \to \mathbb{N}_s$ by $f_*(j) = y$ and $f_*(x) = f(x)$ if $x \in \mathbb{N}_k \setminus \{j\}$. Then f_* maps from $\mathbb{N}_k$ to $\mathbb{N}_m$ and, furthermore, it is an injection. So, by the inductive hypothesis, $k \leq s$ and hence $k+1 \leq s+1 = m$.

□

A consequence of this is:

Theorem 4.3 Suppose n, m are two natural numbers. If there is a bijection from $\mathbb{N}_n$ to $\mathbb{N}_m$, then $n = m$.

Proof. Suppose $f : \mathbb{N}_n \to \mathbb{N}_m$ is a bijection. Then f is an injection. So from Theorem PP, $n \leq m$.

But there is in inverse function $f^{-1} : \mathbb{N}_m \to \mathbb{N}_n$ and this is also a bijection. In particular, f^{-1} is an injection from $\mathbb{N}_m$ to $\mathbb{N}_n$, and hence $m \leq n$.

Now we have both $n \leq m$ and $m \leq n$, hence $n = m$. □

A slightly more general form of the pigeonhole principle, easy to prove from that above is:

Theorem 4.4 Suppose that A and B are sets with $|A| = n$ and $|B| = m$, where $m, n \in \mathbb{N}$. If there is an injection from A to B, then $m \leq n$.

Proof. From the definition of counting, there are bijections $g : \mathbb{N}_n \to A$ and $h : \mathbb{N}_m \to B$. We also have an inverse bijection $h^{-1} : B \to \mathbb{N}_m$.

Suppose there is an injection $f : A \to B$. Consider the composite function $h^{-1} f\, g : \mathbb{N}_n \to \mathbb{N}_m$. If we can prove that this is an injection, then from Theorem 4.2 it follows that $n \leq m$.

So, let us prove injectivity. Suppose $a, b \in \mathbb{N}_n$ with $a \neq b$. Since g is a bijection $g(a), g(b) \in A$ with $g(a) \neq g(b)$. Since f is an injection $f(g(a)), f(g(b)) \in B$ with $f(g(a)) \neq f(g(b))$. Since h^{-1} is a bijection $h^{-1}(f(g(a))), h^{-1}(f(g(b))) \in \mathbb{N}_m$ with $h^{-1}(f(g(a))) \neq h^{-1}(f(g(b)))$. This last inequality is what we need. □

The pigeonhole principle is remarkably useful (even in some very advanced areas of mathematics). It has many applications. For most applications, it is the contrapositive form of the principle that is used. This states:

> If $m < n$ then there is no injection $f : \mathbb{N}_n \to \mathbb{N}_m$.

So, if $m < n$, and f is *any* function $f : \mathbb{N}_n \to \mathbb{N}_m$, then there are $x, y \in \mathbb{N}_n$ with $x \neq y$ such that $f(x) = f(y)$.

4.6.2 Some applications of the Pigeonhole Principle

We now prove some theorems using the pigeonhole principle.

We start with an easy example.

Theorem 4.5 In any group of 13 or more people, there are two persons whose birthday is in the same month.

Proof. Consider the function that maps the people to their months of birth. Since $13 > 12$, this cannot be a bijection, so two people are born in the same month. □

This next one is not hard, but perhaps not immediately obvious.

Theorem 4.6 In a room full of people, there will always be at least two people who have the same number of friends in the room.

Proof. Let X be the set of people in the room and suppose $|X| = n \geq 2$. Consider the function $f : X \to \mathbb{N} \cup \{0\}$ where $f(x)$ is the number of friends x has in the room.

Let's assume that a person can't be a friend of themselves. (We could instead assume that a person is always friendly with themselves: we simply need a convention one way or the other.)

Then $f(X) = \{f(x) : x \in X\} \subseteq \{0, 1, \ldots, n-1\}$. But there can't be x, y with $f(x) = n-1$ and $f(y) = 0$. **Why?** Well, such a y would be a friend of all the others, including x, which isn't possible since x has no friends in the room.

So either $f(X) \subseteq \{0, 1, \ldots, n-2\}$ or $f(X) \subseteq \{1, \ldots, n-1\}$. In each case, since $f(x)$ can take at most $n-1$ values, there must, by PP, be at least two $x, y \in X$ with $f(x) = f(y)$. And that's what we needed to prove. □

Here's an interesting geometrical example. For two points (x_1, y_1), (x_2, y_2) in the plane, the **midpoint** of (x_1, y_1) and (x_2, y_2) is the point

$$\left(\frac{1}{2}(x_1 + x_2), \frac{1}{2}(y_1 + y_2)\right)$$

(the point on the middle of the line connecting the two points).

Theorem 4.7 If we have a set A of five or more points in the plane with **integer** coordinates, then there are two points in A whose midpoint has integer coordinates.

Proof. For two integers a, b, $\frac{1}{2}(a+b)$ is an integer if and only if $a+b$ is even, so if and only if a, b or both even or both odd.

So the midpoint of $(x_1, y_1), (x_2, y_2)$ has both coordinates integer if and only if x_1, x_2 are **both** even or **both** odd, **and also** y_1, y_2 are **both** even or **both** odd.

Let's label each of the points (a, b) of A with one of "(even,even)", "(even,odd)", "(odd,even)" or "(odd,odd)".

Since $|A| \geq 5$, there will be at least two points which receive the same label. Hence these two points have the same parity (odd or even) for the first coordinate, and the same parity for the second coordinate. This means the midpoint of these two points must be integer as well. □

By the way, this result would not necessarily hold if we only had four points in the set. Consider $(0, 0)$, $(1, 0)$, $(1, 0)$ and $(1, 1)$.

Here's a very interesting number theory application (with a very sneaky proof). It uses the notion of remainders on division by n, which we'll cover properly soon: for now, all we need is that, for every natural number m, the "remainder, r, upon division by n" is one of the numbers $0, 1, \ldots, n-1$, and that $m - r$ is divisible by n.

Theorem 4.8 Let $a_1, a_2, \ldots, a_n$ be n integers (where $n \geq 2$). Then there exists a non-empty collection of these integers whose sum is divisible by n.

Proof. Consider the numbers $s_0, s_1, \ldots, s_n$ given by

$$s_0 = 0,$$

$$s_1 = a_1,$$
$$s_2 = a_1 + a_2,$$
$$s_3 = a_1 + a_2 + a_3,$$

etc., until

$$s_n = a_1 + a_2 + \cdots + a_n.$$

(It is not obvious, at all, why we should do this, but it will work!)

For each of these s_i, consider the remainder upon division by n. Since there are $n+1$ numbers s_i, but only n possible remainders $(0, 1, \ldots, n-1)$, two of the s_i will have the same remainder upon division by n.

So suppose s_k and s_ℓ have the same remainder, where $k < \ell$. Then $s_\ell - s_k$ is divisible by n. But since $s_\ell - s_k = a_{k+1} + a_{k+2} + \cdots + a_\ell$, this means that the sum $a_{k+1} + a_{k+2} + \cdots + a_\ell$ is divisible by n. Se we have proved the result. □

In fact we proved something even stronger than what we set out to prove:
Let $a_1, a_2, \ldots, a_n$ be a list of n integers (where $n \geq 2$). Then there exists a non-empty collection of **consecutive** numbers from this list $a_{k+1}, a_{k+2}, \ldots, a_\ell$ whose sum is divisible by n.

The theorem isn't true if we have fewer than n integers. For instance, if for any $n \geq 2$ we take the numbers $a_1, \ldots, a_{n-1}$ all equal to 1, then it's impossible to find a sum that adds up to something divisible by n.

4.7 A generalised form of PP

We state without proof the following more general version of the PP. Again, it's rather obvious. Isn't it?

Theorem 4.9 Suppose $f : A \to B$ and that $|A| > k|B|$ where $k \in \mathbb{N}$. Then there is some element of B that is the image of at least $k+1$ elements of A.

Last year, 232 students were registered for this course. I knew, before marking the exams, that at least three of them would get the same exam mark.

Why? Well, apply the theorem, with A being the students, B being the set $\{0, 1, \ldots, 100\}$ of all possible marks (which is of size 101) and $f(x)$ the mark of student x. Since $232 > 2(101)$, there's some mark y such that at least $2 + 1 = 3$ students will have $y = f(x)$, which means they get the same mark.

4.8 Infinite sets

We say that a set A is *finite* when there is some $n \in \mathbb{N}$ such that $|A| = n$. Otherwise, A is said to be *infinite*.

For example, the set of natural numbers is infinite. You might think that's obvious, but how would you prove it? (Remember that the formal definition that a set A has cardinality n is that there is a bijection between $\mathbb{N}_n$ and A.)

One way to show this is to use a proof by contradiction. Suppose (for a contradiction) that $\mathbb{N}$ is finite, of cardinality $n \in \mathbb{N}$, and that $f : \mathbb{N}_n \to \mathbb{N}$ is a bijection. Consider the number $N = f(1) + f(2) + \cdots + f(n)$. Since each $f(i)$ is a natural number, for all $i \in \mathbb{N}_n$, N is also a natural number. But $N > f(i)$ for all $i \in \mathbb{N}_n$. So here is a natural number, N, that is not equal to $f(i)$ for any $i \in \mathbb{N}_n$. But that contradicts the fact that f is a bijection, because if it's a bijection then it's certainly a surjection and there should be some $i \in \mathbb{N}_n$ with $f(i) = N$.

4.9 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- describe precisely what is meant by a function
- describe precisely what it means to say a function is a surjection, an injection and a bijection, and be able to determine whether a given function has these properties
- state the definition of the composite function gf
- establish whether a function has an inverse or not
- demonstrate that you understand the formal definition of the cardinality of a finite set
- state and use the pigeonhole principle
- state what it means to say that a set is infinite; and be able to prove that a set is infinite.

4.10 Sample exercises

Exercise 4.1

Suppose that X, Y, Z are sets and that $f : X \to Y$ and $g : Y \to Z$. Prove that if f and g are injections, so is the composition gf. Prove also that if f and g are surjections, then so if the composition gf.

Exercise 4.2

Let $\mathbb{Z}$ be the set of all integers and suppose that $f : \mathbb{Z} \to \mathbb{Z}$ is given, for $x \in \mathbb{Z}$, by

$$f(x) = \begin{cases} x+1 & \text{if } x \text{ is even} \\ -x+3 & \text{if } x \text{ is odd.} \end{cases}$$

Determine whether f is injective. Determine also whether f is surjective.

Exercise 4.3

Suppose that X, Y, Z are sets, that $f : X \to Y$, $g : Y \to Z$, and $h : Y \to Z$. Suppose that the compositions hf and gf are equal, and also that f is surjective. Prove that $g = h$. □

Exercise 4.4

Suppose that X, Y, Z are sets and that $f : X \to Y$ and $g : Y \to Z$. Prove that if the composition gf in injective, then f is injective. Prove that if gf is surjective, then g is surjective. □

Exercise 4.5

Suppose that A and B are non-empty finite sets and that they are disjoint (meaning that $A \cap B = \emptyset$). Prove, using the formal definition of cardinality, that $|A \cup B| = |A| + |B|$. □

Exercise 4.6

Suppose that X, Y are any two finite sets. By using the fact that

$$X \cup Y = (X \setminus Y) \cup (Y \setminus X) \cup (X \cap Y),$$

together with the result of Exercise 4.5, prove that

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

Exercise 4.7

Suppose $n \in \mathbb{N}$ and that $f : \mathbb{N}_{2n+1} \to \mathbb{N}_{2n+1}$ is a bijection. Prove that there is some odd integer $k \in \mathbb{N}_{2n+1}$ such that $f(k)$ is also odd. (State clearly any results you use.) □

4.11 Comments on selected activities

Learning activity 4.1 Given any $y \in \mathbb{R}$, let $x = y/2$. Then $f(x) = 2(y/2) = y$. This shows that f is surjective. Also, for $x, y \in \mathbb{R}$,

$$f(x) = f(y) \Rightarrow 2x = 2y \Rightarrow x = y,$$

which shows that f is injective. Hence f is a bijection.

4.12 Solutions to exercises

Solution to exercise 4.1

Suppose f and g are injective. Then, for $x, y \in X$,

$$\begin{aligned}(gf)(x) = (gf)(y) \quad &\Rightarrow \quad g(f(x)) = g(f(y)) \\ &\Rightarrow \quad f(x) = f(y) \text{ (because } g \text{ is injective)} \\ &\Rightarrow \quad x = y \text{ (because } f \text{ is injective).}\end{aligned}$$

This shows that gf is injective.

Suppose that f and g are surjective. Let $z \in Z$. Then, because g is surjective, there is some $y \in Y$ with $g(y) = z$. Because f is surjective, there is some $x \in X$ with $f(x) = y$. Then

$$(gf)(x) = g(f(x)) = g(y) = z,$$

so z is the image of some $x \in X$ under the mapping gf. Since z was any element of Z, this shows that gf is surjective. □

Solution to exercise 4.2

Suppose one of x, y is even and the other odd. Without any loss of generality, we may suppose x is even and y odd. ('Without loss of generality' signifies that there is no need to consider also the case in which x is odd and y is even, because the argument we'd use there would just be the same as the one we're about to give, but with x and y interchanged.) So $f(x) = x + 1$ and $f(y) = -y + 3$. But we cannot then have $f(x) = f(y)$ because $x + 1$ must be an odd number and $-y + 3$ an even number. So if $f(x) = f(y)$, then x, y are both odd or both even. If x, y are both even, this means $x + 1 = y + 1$ and hence $x = y$. If they are both odd, this means $-x + 3 = -y + 3$, which means $x = y$. So we see that f is injective.

Is f surjective? Let $z \in \mathbb{Z}$. If z is odd, then $z - 1$ is even and so $f(z-1) = (z-1) + 1 = z$. If z is even, then $3 - z$ is odd and so $f(3-z) = -(3-z) + 3 = z$. So for $z \in \mathbb{Z}$ there is $x \in \mathbb{Z}$ with $f(x) = z$ and hence f is surjective.

Solution to exercise 4.3

Suppose f is surjective and that $hf = gf$. Let $y \in Y$. We show $g(y) = h(y)$. Since y is any element of Y in this argument, this will establish that $g = h$. Because f is surjective, there is some $x \in X$ with $f(x) = y$. Then, because $hf = gf$, we have $h(f(x)) = g(f(x))$, which means that $h(y) = g(y)$. So we've achieved what we needed.

Solution to exercise 4.4

Suppose gf is injective. To show that f is injective we need to show that $f(x) = f(y) \Rightarrow x = y$. Well,

$$f(x) = f(y) \Rightarrow g(f(x)) = g(f(y)) \Rightarrow (gf)(x) = (gf)(y) \Rightarrow x = y,$$

where the last implication is because gf is injective.

Now suppose gf is surjective. So for all $z \in Z$ there is some $x \in X$ with $(gf)(x) = z$. So $g(f(x)) = z$. Denoting $f(x)$ by y, we therefore see that there is $y \in Y$ with $g(y) = z$. Since z was any element of Z, this shows that g is surjective. □

Solution to exercise 4.5

Suppose $|A| = m$ and $|B| = n$. We need to show that $|A \cup B| = m + n$ which means, according to the definition of cardinality, that we need to show there is a bijection from $\mathbb{N}_{m+n}$ to $A \cup B$. Because $|A| = m$, there is a bijection $f : \mathbb{N}_m \to A$ and because $|B| = n$, there is a bijection $g : \mathbb{N}_n \to B$. Let us define $h : \mathbb{N}_{m+n} \to A \cup B$ as follows:

$$\text{for } 1 \le i \le m,\ h(i) = f(i) \quad \text{and for } m + 1 \le i \le m + n,\ h(i) = g(i - m).$$

Then h is injective. We can argue this as follows: if $1 \le i, j \le m$ then

$$h(i) = h(j) \Rightarrow f(i) = f(j) \Rightarrow i = j,$$

because f is injective. If $m + 1 \le i, j \le m + n$ then

$$h(i) = h(j) \Rightarrow g(i - m) = g(j - m) \Rightarrow i - m = j - m \Rightarrow i = j,$$

because g is injective. The only other possibility is that one of i, j is between 1 and m and the other between $m + 1$ and $m + n$. In this case, the image under h of one of i, j belongs to A and the image of the other to B and these cannot be equal because $A \cap B = \emptyset$. So h is indeed an injection. It is also a surjection. For, given $a \in A$, because f is a surjection, there is $1 \le i \le m$ with $f(i) = a$. Then $h(i) = a$ also. If $b \in B$ then there is some $1 \le j \le n$ such that $g(j) = b$. But then, this means that $h(m + j) = g((m + j) - m) = b$, so b is the image under h of some element of $\mathbb{N}_{m+n}$. So h is a bijection from $\mathbb{N}_{m+n}$ to $A \cup B$ and hence $|A \cup B| = m + n$.

Solution to exercise 4.6

Note first that the two sets $(X \setminus Y) \cup (Y \setminus X)$ and $X \cap Y$ are disjoint. Therefore,

$$|X \cup Y| = |(X \setminus Y) \cup (Y \setminus X)| + |X \cap Y|.$$

Now, $(X \setminus Y)$ and $(Y \setminus X)$ are disjoint, so

$$|(X \setminus Y) \cup (Y \setminus X)| = |(X \setminus Y)| + |(Y \setminus X)|$$

and therefore

$$|X \cup Y| = |(X \setminus Y)| + |(Y \setminus X)| + |X \cap Y|.$$

Now, the sets $X \setminus Y$ and $X \cap Y$ are disjoint and their union is X, so

$$|X| = |(X \setminus Y) \cup (X \cap Y)| = |X \setminus Y| + |X \cap Y|.$$

A similar argument shows that

$$|Y| = |(Y \setminus X) \cup (X \cap Y)| = |Y \setminus X| + |X \cap Y|.$$

These mean that

$$|X \setminus Y| = |X| - |X \cap Y| \text{ and } |Y \setminus X| = |Y| - |X \cap Y|.$$

So we have

$$\begin{aligned} |X \cup Y| &= |(X \setminus Y)| + |(Y \setminus X)| + |X \cap Y| \\ &= (|X| - |X \cap Y|) + (|Y| - |X \cap Y|) + |X \cap Y| \\ &= |X| + |Y| - |X \cap Y|. \end{aligned}$$

Solution to exercise 4.7

Let E be the set of even integers, and O the set of odd integers, in the range $\{1, 2, \ldots, 2n+1\}$. Then $|E| = n$ and $|O| = n+1$. If f was such that $f(k)$ was even for all $k \in O$, then $f^* : O \to E$ given by $f^*(x) = f(x)$ would be an injection. But, by the pigeonhole principle, since $|O| > |E|$, such an injection cannot exist. Hence there is some odd k such that $f(k)$ is odd. □

Chapter 5
Equivalence relations and the integers

☞ Biggs, N. L. *Discrete Mathematics.* Chapter 7.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning.* Chapter 22.

5.1 Introduction

In this chapter of the notes we study the important idea of an *equivalence relation*, a concept that is central in abstract mathematics. As an important example, we look at how the integers can be defined from the natural numbers through the use of an equivalence relation. We also study some of the important properties of the integers.

5.2 Equivalence relations

5.2.1 Relations in general

The idea of a *relation* is quite a general one. For example, consider the set of natural numbers $\mathbb{N}$ and let us say that two natural numbers m, n are related, denoted by $m\,R\,n$, if $m+n$ is even. So we have, for instance, $6\,R\,2$ and $7\,R\,5$, but that 6 and 3 are not related. This relation has some special properties. For one thing, since $2n$ is even for all $n \in \mathbb{N}$, $n\,R\,n$ for all $n \in \mathbb{N}$. (We say such a relation is *reflexive*.) Also, if $m\,R\,n$, then $m+n$ is even. But $m+n = n+m$ and hence, also, $n\,R\,m$. (We say such a relation is *symmetric*.) It is because $m\,R\,n \iff n\,R\,m$ that we can simply say that 'm and n are related' rather than 'm is related to n' or 'n is related to m'. The relation R has other important properties that we will come back to later.

Formally, a relation R on a set X is a subset of the Cartesian product $X \times X$ (which, recall, is the set of all ordered pairs of the form (x, y) where $x, y \in X$).

> **Example 5.1** Suppose R is the relation on $\mathbb{R}$ given by $x\,R\,y \iff x > y$. Regarded as a subset of $\mathbb{R} \times \mathbb{R}$, this is the set $\{(x, y) \mid x > y\}$. This relation does not possess the reflexive and symmetric properties we met in the example above. For no $x \in \mathbb{R}$ do we have $x\,R\,x$ because x is not greater than x. Furthermore, if $x\,R\,y$ then $x > y$, and we cannot therefore also have $y\,R\,x$, for that would imply the contradictory statement that $y > x$.

In many cases, we use special symbols for relations. For instance '$=$' is a relation, as is $>$. It is often convenient to use a symbol other than R: for instance, many textbooks use $x \sim y$ rather than $x\,R\,y$ to denote the typical relation.

5.2.2 The special properties of equivalence relations

There are three special properties that a relation might have (two of which we saw in one of the earlier examples):

Definition 5.1 Suppose that R is relation on a set X. Then

- **[The reflexive property]** R is said to be *reflexive* if, for all $x \in X$, $x\,R\,x$.
- **[The symmetric property]** R is said to be *symmetric* if, for all $x, y \in X$, $x\,R\,y$ implies $y\,R\,x$ (equivalently, for all $x, y \in X$, $x\,R\,y \iff y\,R\,x$).
- **[The transitive property]** R is said to be *transitive* if, for all $x, y, z \in X$, whenever $x\,R\,y$ and $y\,R\,z$, we also have $x\,R\,z$; that is, $(x\,R\,y) \wedge (y\,R\,z) \Rightarrow xRz$.

A relation that has all three of these properties is called an *equivalence relation*.

Definition 5.2 A relation is an *equivalence relation* if is reflexive, symmetric and transitive.

Example 5.2 We saw earlier that the relation on $\mathbb{N}$ given by

$$m\,R\,n \iff m + n \text{ is even}$$

is reflexive and symmetric. It is also transitive. To prove that, suppose x, y, z are three natural numbers and that $x\,R\,y$ and $y\,R\,z$. Then $x + y$ is even and $y + z$ is even. To show that $x\,R\,z$ we need to establish that $x + z$ is even. Well,

$$x + z = (x + y) + (y + z) - 2y,$$

and all three terms on the right ($x + y$, $y + z$, and $2y$) are even. Therefore, $x + z$ is even and so $x\,R\,z$.

Example 5.3 Let X be the set of $n \times n$ real matrices. Define a relation $\sim$ on X by:

$$M \sim N \iff \exists r, s \in \mathbb{N} \text{ s.t. } M^r = N^s.$$

Then $\sim$ is an equivalence relation.

Reflexivity and symmetry are easy to see: $M^1 = M^1$ and, if $M^r = N^s$, then $N^s = M^r$. Proving transitivity requires more work. Suppose $M \sim N$ and $N \sim R$. Then there are $r, s, t, u \in \mathbb{N}$ with $M^r = N^s$ and $N^t = R^u$. Then

$$M^{rt} = (M^r)^t = (N^s)^t = (N^t)^s = (R^u)^s = R^{us},$$

so there are integers $w = rt$ and $x = us$ such that $M^w = R^x$ and hence $M \sim R$.

5.3 Equivalence classes

Given an equivalence relation, it is natural to group together objects that are related to each other. The resulting groupings are known as *equivalence classes*. In this section, we formally define equivalence classes and discuss some of their properties.

Definition 5.3 Suppose R is an equivalence relation on a set X and, for $x \in X$, let $[x]$ be the set of all $y \in X$ such that $y\,R\,x$. So,

$$[x] = \{y \in X \mid y\,R\,x\}.$$

Notice that each $[x]$ is a *subset* of X.

Example 5.4 Consider again R on $\mathbb{N}$ given by $m\,R\,n \iff m + n$ is even. Any even number is related to any other even number; and any odd number to any odd number. So there are two equivalence classes:

$$[1] = [3] = [5] = \cdots = \text{set of odd positive integers},$$

$$[2] = [4] = [6] = \cdots = \text{set of even positive integers},$$

Example 5.5 Suppose that $f : X \to Y$ is a surjection. Define relation R on X by $x\,R\,y \iff f(x) = f(y)$. Then R is an equivalence relation and the equivalence classes are given by

$$C_y = \{x \in X : f(x) = y\},$$

for $y \in Y$.

Activity 5.1 Check this!

The equivalence classes have a number of important properties. These are given in the following result.

Theorem 5.1 Suppose R is an equivalence relation on a set X. Then

(i) For $x, y \in X$, $[x] = [y] \iff x\,R\,y$

(ii) For $x, y \in X$, if x and y are not related by R, then $[x] \cap [y] = \emptyset$.

Proof. (i) This is an if and only if statement, so we have two things to prove: namely that $[x] = [y] \Rightarrow x\,R\,y$ and that $x\,R\,y \Rightarrow [x] = [y]$.

Suppose, then, that $[x] = [y]$. The relation R is reflexive, so we have $x\,R\,x$. This means that $x \in [x]$. But if $[x] = [y]$, then we must have $x \in [y]$. But that means (by definition of $[y]$) that $x\,R\,y$.

Conversely, suppose that $x\,R\,y$. We now want to show that $[x] = [y]$. So let $z \in [x]$. (We will show that $z \in [y]$.) Then $z\,R\,x$. But, because $x\,R\,y$ and R is transitive, it follows that $z\,R\,y$ and hence $z \in [y]$. This shows $[x] \subseteq [y]$. We now need to show that $[y] \subseteq [x]$. Suppose $w \in [y]$. Then $w\,R\,y$ and, since $x\,R\,y$, we also have, since R is symmetric, $y\,R\,x$. So $w\,R\,y$ and $y\,R\,x$. By transitivity of R, $w\,R\,x$ and hence $w \in [x]$. This shows that $[y] \subseteq [x]$. Because $[x] \subseteq [y]$ and $[y] \subseteq [x]$, $[x] = [y]$, as required.

(ii) Suppose x and y are not related. We prove by contradiction that $[x] \cap [y] = \emptyset$. So suppose $[x] \cap [y] \neq \emptyset$. Let z be any member of the intersection $[x] \cap [y]$. (The fact that we're assuming the intersection is non-empty means there is such a z.) Then $z \in [x]$, so $z\,R\,x$ and $z \in [y]$, so $z\,R\,y$. Because R is symmetric, $x\,R\,z$. So: $x\,R\,z$ and $z\,R\,y$ and, therefore, by transitivity, $x\,R\,y$. But this contradicts the fact that x, y are not related by R. So $[x] \cap [y] = \emptyset$. □

Theorem 5.1 shows that either two equivalence classes are equal, or they are *disjoint*. Furthermore, because an equivalence relation is reflexive, any $x \in X$ is in some equivalence class (since it certainly belongs to $[x]$ because $x\,R\,x$). So what we see is that the equivalence classes form a *partition* of X: their union is the whole of X, and no two equivalence classes overlap.

> **Example 5.6** Consider again the equivalence relation R on $\mathbb{N}$ given by
>
> $$m\,R\,n \iff m + n \text{ is even.}$$
>
> We have seen that there are precisely two equivalence classes: the set of odd positive integers and the set of even positive integers. Note that, as the theory predicted, these form a partition of all of $\mathbb{N}$ (since every natural number is even or odd, but not both).

5.4 Construction of the integers from the natural numbers

This section might seem at first a little tricky. We know what the integers are, so why do we have to *define* or *construct* them, you might well ask. Well, for one thing the procedure we're about to look at is used very often in mathematics. We will meet it again when we study rational numbers and modular arithmetic. The big idea here is that we can often define an interesting set of objects (and perform operations on them) by considering the set of equivalence classes of some relation. That is, we have a set X and an equivalence relation R on X, and we look at the set C of equivalence classes. In the abstract, this is perhaps confusing, but we will see some examples over the course of the next few chapters. We start with the integers.

(The informal motivation behind what follows is this: imagine you have deposits of a and debts of b. Denote this by (a, b). Then you have a total of $a - b$, which might be negative.)

We can describe, or construct, the integers from the natural numbers, using an equivalence relation. In fact, we consider an equivalence relation on the set $\mathbb{N} \times \mathbb{N}$ of all ordered pairs of natural numbers. Given (a, b) and (c, d) in $X = \mathbb{N} \times \mathbb{N}$, let us say that

$$(a, b)\,R\,(c, d) \iff a + d = b + c.$$

(Informal motivation: we're thinking of $[(a, b)]$ as the (familiar) integer $a - b$. That's why we defined

$$(a, b)\,R\,(c, d) \iff a + d = b + c.$$

This is the 'same' as $a - b = c - d$. **But** we want to make this work, in the set of **natural numbers** and **using addition**, not subtraction, which we haven't defined.)

I've said this is an equivalence relation, but let's check this.

First, R is reflexive because $(a, b)\,R\,(a, b)$ if and only if $a + b = b + a$, which is clearly true.

Next, R is symmetric, for

$$(a, b)\,R\,(c, d) \iff a + d = b + c \iff c + b = d + a \iff (c, d)\,R(a, b).$$

Finally, R is transitive. For suppose that $(a, b)\,R\,(c, d)$ and $(c, d)\,R\,(e, f)$. Then $a + d = b + c$ and $c + f = d + e$. Therefore,

$$(a + d) + (c + f) = (b + c) + (d + e).$$

That is (after cancelling c and d from each side),

$$a + f = b + e,$$

which means $(a, b)\,R\,(e, f)$.

What are the equivalence classes of R? The typical equivalence class $[(a, b)]$ contains all (c, d) for which $a + d = b + c$. For example, $[(2, 1)]$ will be

$$[(2, 1)] = \{(3, 2), (4, 3), (5, 4), \ldots\}.$$

Now, for $n \in \mathbb{N}$, let us denote the equivalence class $[(n + 1, 1)]$ by n and let us denote the equivalence class $[(1, n + 1)]$ by $-n$. Also, we denote by 0 the class $[(1, 1)]$. Then we *define* the integers to be the set

$$\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

We can 'do arithmetic' (addition and multiplication) with the integers defined in this way. First, let's define an addition operation (between equivalence classes) by

$$[(a, b)] + [(c, d)] = [(a + c, b + d)].$$

So, for example, what does this say about the sum of integers $+3$ and -1. Well, $+3 = [(4, 1)]$ and $-1 = [(1, 2)]$ and therefore

$$+3 + -1 = [(4, 1)] + [(1, 2)] = [(5, 3)].$$

Now, $(5, 3)\,R\,(3, 1)$, so $[(5, 3)] = [(3, 1)]$, which is what we called $+2$. No surprise there, then: $3 + (-1) = 2$ in the usual notation for integer arithmetic. (Henceforth, we can simply write $m + (-n)$ as the subtraction $m - n$.)

There is quite a subtle point about this definition of addition of equivalence classes. We know that there are many (infinitely many) pairs (a', b') such that $[(a, b)] = [(a', b')]$. Such an (a', b') is called a *representative* of the equivalence class. (It is always the case, for any equivalence relation, as Theorem 5.1 shows, that $[x]$ and $[x']$ are the same whenever $x'\,R\,x$. So any equivalence class can be represented by potentially many different representatives: any member of the class will do.) So we need to be sure that the definition of addition will give us the same answer if we use different representatives. In other words, suppose that $[(a',b')] = [(a,b)]$ and $[(c',d')] = [(c,d)]$. The definition of addition,

$$[(a,b)] + [(c,d)] = [(a+c, b+d)],$$

will only make sense (or, it will only be 'consistent') if

$$[(a',b')] + [(c',d')] = [(a,b)] + [(c,d)].$$

Thus, we need to prove that if $[(a',b')] = [(a,b)]$ and $[(c',d')] = [(c,d)]$, then

$$[(a+c, b+d)] = [(a'+c', b'+d')].$$

We can see this easily enough in specific cases. For instance, $[(4,1)] = [(6,3)]$ and $[(1,2)] = [(2,3)]$. We have

$$[(4,1)] + [(1,2)] = [(5,3)] \quad \text{and } [(6,3)] + [(2,3)] = [(8,6)].$$

Well, $(5,3)\,R\,(8,6)$, because $5+6 = 3+8$, so $[(5,3)] = [(8,6)]$. So, in this case, and with these choices of representatives for each equivalence class, we end up with the same class when we apply the addition operation.

We can prove, more generally, that the definition works, that it does not depend on the choice of representatives of the classes. Remember, what we need to prove is that if $[(a',b')] = [(a,b)]$ and $[(c',d')] = [(c,d)]$, then

$$[(a+c, b+d)] = [(a'+c', b'+d')].$$

Well, $[(a',b')] = [(a,b)]$ and $[(c',d')] = [(c,d)]$ mean that $(a',b')\,R\,(a,b)$ and $(c',d')\,R\,(c,d)$, so that $a'+b = b'+a$ and $c'+d = d'+c$. We need to show that $[(a+c,b+d)] = [(a'+c',b'+d')]$. This means we need to show that $(a+c,b+d)\,R\,(a'+c',b'+d')$. But

$$\begin{aligned} (a+c,b+d)\,R\,(a'+c',b'+d') &\iff (a+c)+(b'+d') = (b+d)+(a'+c') \\ &\iff (a+b')+(d'+c) = (a'+b)+(c'+d), \end{aligned}$$

which is true, because $a'+b = b'+a$ and $c'+d = d'+c$.

Multiplication, ×, is defined on these equivalence classes by

$$[(a,b)] \times [(c,d)] = [(ac+bd, ad+bc)].$$

5.5 Properties of the integers

We now revert to the usual notation for the integers. In particular, we denote $+n$ by n and consider $\mathbb{N}$ to be a subset of $\mathbb{Z}$, the set of integers. Many of the usual properties of integers follow from the formal definition of integers given in the previous section: see Biggs, Section 7.5, for more details. We give one example:

Example 5.7 With the integers as defined formally in the previous section, we show that for any integer z, $z + 0 = z$. Recall that, for some (a,b), we will have $z = [(a,b)]$ and that 0 is $[(1,1)]$. Now, the definition of addition of integers (that is, of the equivalence classes) means that

$$z + 0 = [(a,b)] + [(1,1)] = [(a+1, b+1)].$$

But $(a+1,b+1)\,R\,(a,b)$ because $(a+1)+b = (b+1)+a$, and hence $[(a+1,b+1)] = [(a,b)] = z$. So $z + 0 = z$.

Activity 5.2 With integers defined in the formal way as these equivalence classes, prove that for any integer z, $z \times 0 = 0$.

5.6 Ordering the integers

For integers $x = [(a,b)]$ and $y = [(c,d)]$, we say $x < y$ if and only if $a + d < b + c$.

We noted in an earlier chapter that any non-empty subset of $\mathbb{N}$ has a least member. But this is not true for subsets of integers.

For a subset S of $\mathbb{Z}$, m is a *lower bound* for S if for all $s \in S$, $m \leq s$; and M is an *upper bound* for S if for all $s \in S$, $s \leq M$. We say that S is *bounded below* if it has a lower bound; and that it is *bounded above* if it has an upper bound. The natural number l is a *least member* of S if $l \in S$ and, for all $s \in S$, $l \leq s$. So a least member will be a lower bound that belongs to S. The natural number g is a *greatest member* of S if $g \in S$ and, for all $s \in S$, $g \geq s$. So a greatest member will be an upper bound that belongs to S.

The following fact is a fundamental property of the integers, known as the *well-ordering principle*. (The well-ordering principle was discussed earlier, when it was presented as an axiom for the natural numbers. This is a generalisation of that principle.)

The Well-ordering Principle: If S is a non-empty set of integers that has a lower bound, then S has a least member.

(The same statement is true with 'lower' replaced by 'upper' and 'least' replaces by 'greatest'.)

Furthermore, if S is bounded below, then there is *precisely one* least member. For, if l, l' are least members then $l, l' \in S$ and so (since for all $s \in S$, $l \leq s$ and $l' \leq s$) we have both $l \leq l'$ and $l' \leq l$, so that $l = l'$.

5.7 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- demonstrate that you know what is meant by a relation
- demonstrate that you know what it means to say a relation is reflexive, symmetric or transitive, or that it is an equivalence relation

- verify whether given relations are reflexive, symmetric or transitive
- demonstrate that you know the definition of equivalence classes and that you know some of their basic properties, in particular that they form a partition of the set on which the relation is defined
- determine the equivalence classes that correspond to an equivalence relation
- demonstrate knowledge of the way in which the integers can be formally constructed from the natural numbers through the use of an equivalence relation
- state the Well-Ordering Principle

5.8 Sample exercises

Exercise 5.1

Define a relation R on $\mathbb{Z}$ by: for $x, y \in \mathbb{Z}$, $x \, R \, y \iff x^2 = y^2$. Prove that R is an equivalence relation, and describe the corresponding equivalence classes.

Exercise 5.2

Define the relation R on the set $\mathbb{N}$ by $x \, R \, y$ if and only if there is some $n \in \mathbb{Z}$ such that $x = 2^n y$. Prove that R is an equivalence relation. □

Exercise 5.3

Let X be the set of $n \times n$ real matrices. Define a relation $\sim$ on X by:

$$M \sim N \iff \exists \text{ an invertible } P \in X \text{ s.t. } N = P^{-1}MP.$$

Prove that $\sim$ is an equivalence relation.

Exercise 5.4

Suppose that $f : X \to Y$ is a surjection. Define the relation R on X by $x \, R \, y \iff f(x) = f(y)$. Prove that R is an equivalence relation. What are the equivalence classes? Let C denote the set of equivalence classes $[x]$ for $x \in X$. Prove that if $[x] = [y]$ then $f(x) = f(y)$. This means that we can define a function $g : C \to Y$ by: $g([x]) = f(x)$. Prove that g is a bijection. □

Exercise 5.5

Prove that the set $\{x \in \mathbb{Z} \mid x \text{ is a multiple of } 4\}$ has no lower bound. □

5.9 Comments on selected activities

Learning activity 5.2 Suppose $z = [(a, b)]$. Also, $0 = [(1, 1)]$. The definition of multiplication is that

$$[(a, b)] \times [(c, d)] = [(ac + bd, ad + bc)].$$

So,

$$z \times 0 = [(a, b)] \times [(1, 1)] = [(a + b, a + b)].$$

Now, $(a + b, a + b) \, R \, (1, 1)$ because $a + b + 1 = 1 + a + b$ and therefore $[(a + b, a + b)] = [(1, 1)] = 0$. So we see that $z \times 0 = 0$.

5.10 Solutions to exercises

Solution to exercise 5.1

R is reflexive because for any x, $x^2 = x^2$. R is symmetric because $x^2 = y^2 \iff y^2 = x^2$. To show R is transitive, suppose $x, y, z \in \mathbb{Z}$ and $x \, R \, y$ and $y \, R \, z$. Then $x^2 = y^2$ and $y^2 = z^2$, so $x^2 = z^2$, which means $x \, R \, z$. Thus R is an equivalence relation. Given any $x \in \mathbb{Z}$, the equivalence class $[x]$ consists precisely of those integers y such that $y^2 = x^2$. So $[x] = \{x, -x\}$.

Solution to exercise 5.2

R is reflexive because for any x, $x = 2^0 x$. R is symmetric because if $x \, R \, y$ then $\exists n \in \mathbb{Z}$ with $x = 2^n y$. This means that $y = 2^{-n} x$ and hence, taking $m = -n$, $\exists m \in \mathbb{Z}$ such that $y = 2^m x$. So $y \, R \, x$. To show R is transitive, suppose $x, y, z \in \mathbb{Z}$ and $x \, R \, y$ and $y \, R \, z$. Then there are $m, n \in \mathbb{Z}$ such that $x = 2^n y$ and $y = 2^m z$, so $x = 2^n y = 2^n(2^m z) = 2^{m+n} z$ which, since $m + n \in \mathbb{Z}$, shows that $x \, R \, z$. Thus R is an equivalence relation.

Solution to exercise 5.3

For any M, $M = I^{-1}MI$ where I is the identity matrix, so $M \sim M$. For matrices $M, N \in X$, if $M \sim N$ then there's an invertible P with $N = P^{-1}MP$ and so $M = PNP^{-1}$, which can be written as $M = (P^{-1})^{-1}MP^{-1}$. So there is an invertible matrix Q (equal to P^{-1}) such that $M = Q^{-1}NQ$ and hence $M \sim N$. This shows the relation is symmetric. Suppose $M \sim N$ and $N \sim R$. Then there are invertible matrices P and Q such that $N = P^{-1}MP$ and $R = Q^{-1}NQ$. We therefore have

$$R = Q^{-1}(P^{-1}MP)Q = (Q^{-1}P^{-1})M(PQ) = (PQ)^{-1}M(PQ),$$

so there is an invertible matrix $T = PQ$ so that $R = T^{-1}MT$ and hence $M \sim R$, establishing that $\sim$ is transitive. It follows that $\sim$ is an equivalence relation. (We used here the fact that $(PQ)^{-1} = Q^{-1}P^{-1}$. This follows from the fact that $(Q^{-1}P^{-1})(PQ) = Q^{-1}(P^{-1}P)Q = Q^{-1}IQ = Q^{-1}Q = I$.)

Solution to exercise 5.4

$x \, R \, x$ because $f(x) = f(x)$. If $x \, R \, y$ then $f(x) = f(y)$ so $f(y) = f(x)$ and hence $y \, R \, x$. If $x \, R \, y$ and $y \, R \, z$ then $f(x) = f(y)$ and $f(y) = f(z)$, so $f(x) = f(z)$ and $x \, R \, z$. Hence R is an equivalence relation.

For $x \in X$, $[x]$ is the set of all $y \in X$ with $f(y) = f(x)$, so, since f is a surjection, the equivalence classes are exactly the sets C_z for each $z \in Y$, where $C_z = \{x \in X \mid f(x) = z\}$ is the set of elements of X mapped onto z by f.

The fact that $[x] = [y]$ implies $f(x) = f(y)$ follows directly either from this description of equivalence classes, or from the fact that $[x] = [y]$ implies $x \, R \, y$, which implies $f(y) = f(x)$.

Let g be as defined. It is surjective because for each $z \in Y$, there is some $x \in X$ such that $f(x) = z$ (since f is surjective) and hence $g([x]) = f(x) = z$. Also, g is bijective because $g([x]) = g([y])$ implies $f(x) = f(y)$, which means $x \, R \, y$ and hence that $[x] = [y]$. □

Solution to exercise 5.5

We can prove this by contradiction. Suppose that the set $S = \{x \in \mathbb{Z} \mid x \text{ is a multiple of } 4\}$ has a lower bound, l. Then, for all $x \in S$, $x \geq l$. Now, one of $l - 1, l - 2, l - 3, l - 4$ must be a multiple of 4. So one of these numbers is in S. However, each is less than l, contradicting the fact that l is a lower bound on S.

Chapter 6
Divisibility and prime numbers

☞ Biggs, N. L. *Discrete Mathematics.* Chapter 8.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning.* Chapters 15–17 and Chapter 23 (except section 23.5)

6.1 Introduction

In this chapter we begin to study elements of number theory. We start with a discussion of divisibility and this leads us to discuss common divisors. Prime numbers are the basic building blocks in the theory of numbers: in particular, each number can be written in essentially only one way as a product of primes.

6.2 Divisibility

For integers x, y we say that x is a *multiple* of y or that y *divides* x if, for some $q \in \mathbb{Z}$, $x = yq$. We use the notation $y \,|\, x$ to signify that y divides x.

Note that, for every $x \in \mathbb{Z}$, $x \,|\, 0$. But $0 \,|\, x$ only if $x = 0$.

When y does not divide x, we write $y \nmid x$. In this case, as you will know from elementary arithmetic, dividing x by y will leave a remainder.

6.3 Quotients and remainders

The following theorem is very useful. It formalises the fact that one integer may be divided by another, leaving a remainder.

Theorem 6.1 For any positive integers a and b, there are unique non-negative integers q and r such that

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

This can be proved using some standard properties of integers. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and let

$$Q = \{m \in \mathbb{N}_0 \mid bm \leq a\}.$$

Then Q is non-empty because $0 \in Q$ (since $0b = 0 \leq a$). Also, Q is finite, because if $qb \leq a$, then, given that $b \in \mathbb{N}$, we have $q \leq a$. So Q must have a maximum member. Let's call this q. Let $r = a - bq$. Because $q \in Q$, $bq \leq a$ and hence $r \geq 0$. Now, q is the maximum member of A, so $q + 1 \notin A$, which means that $(q + 1)b > a$, so $qb + b > a$ and hence $r = a - bq < b$. So we have established that $a = bq + r$, and that $0 \leq r < b$. To show that q and r are unique, suppose we have $a = bq + r = bq' + r'$ where $0 \leq r, r' < b$. (We show $q = q'$ and $r = r'$.) Either $q \leq q'$ or $q \geq q'$. Let's suppose that $q \geq q'$ (the argument is similar if $q \leq q'$). Then

$$0 \leq r = a - bq \leq a - bq' = r' < b.$$

So

$$0 \leq (a - bq') - (a - bq) < b,$$

which simplifies to $0 \leq b(q - q') < b$. This implies $0 \leq q - q' < 1$. But $q - q'$ is an integer, and so we must have $q - q' = 0$. So $q = q'$. Then, $r = a - bq = a - bq' = r'$.

The same result holds more generally, without the restriction that $a > 0$, but the proof is simplest when we restrict to the positive case. The general *Division Theorem* is:

Theorem 6.2 (Division Theorem) For any integers a and b with $b > 0$, there are unique integers q and r such that

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

6.4 Representation of integers with respect to a base

Let t be a positive integer. Then any positive integer x can represented uniquely in the form

$$x = x_n t^n + x_{n-1} t^{n-1} + \cdots + r_1 t + r_0$$

for some integer n. The numbers x_i are integers between 0 and $t - 1$ and can be found by repeated division by t: see Biggs Section 8.3. We write

$$x = (x_n x_{n-1} \ldots x_1 x_0)_t \,.$$

Example 6.1 The number 60 (written in base 10) can be written in base 3 as $(2020)_3$ because:

$$60 = 2(3^3) + 0(3^2) + 2(3) + 0(1).$$

The representation can be found by repeated division.

$$\begin{aligned} 60 &= 3 \times 20 + \mathbf{0} \\ 20 &= 3 \times 6 + \mathbf{2} \\ 6 &= 3 \times 2 + \mathbf{0} \\ 2 &= 3 \times 0 + \mathbf{2}. \end{aligned}$$

Example 6.2 What's the representation in base 4 of the number $(201)_{10}$?

$$\begin{aligned} 201 &= 4 \times 50 + \mathbf{1} \\ 50 &= 4 \times 12 + \mathbf{2} \\ 12 &= 4 \times 3 + \mathbf{0} \\ 3 &= 4 \times 0 + \mathbf{3}. \end{aligned}$$

So the answer is $(3021)_4$.

Check: $3(4^3) + 2(4) + 1 = 201$.

6.5 Greatest common divisor

The greatest common divisor of two integers is defined as follows.

Definition 6.1 (Greatest common divisor) Suppose a, b are two integers, at least one of which is not 0. Then the *greatest common divisor* (gcd) of a and b, denoted by $\gcd(a, b)$, is the unique positive integer d with the following properties:

(i) d divides both a and b (that is, it is a *common divisor* of a and b)

(ii) d is greater than every other common divisor of a and b: that is, if $c|a$ and $c|b$ then $c \leq d$.

Implicit here is the fact that the gcd is unique. This easily follows from properties (i) and (ii). For, suppose that d and d' are two positive integers satisfying (i) and (ii). Because d' divides a and b and because d satisfies (ii), we have $d' \leq d$. But also, because d divides a and b and because d' satisfies (ii), we have $d \leq d'$. So we must have $d = d'$.

It's not too hard to see that the gcd exists. For any $n \in \mathbb{Z}$, let $D(n) = \{m \in \mathbb{Z} : m|n\}$. This is the set of positive divisors of n. Since $1|n$, $D(n) \neq \emptyset$. Consider now the set $D(a, b) = D(a) \cap D(b)$, which is the set of positive common divisors of a and b. Then, $D(a, b) \neq \emptyset$ since $1 \in D(a, b)$. Suppose $a \neq 0$. (We know that at least one of a, b is nonzero. If $a \neq 0$, a very similar argument will work using b in place of a.) Then $m \in D(a, b) \Rightarrow m \leq |a|$. So $D(a, b)$ is bounded above and hence has a (unique) maximal element d. That's the gcd.

Note that some textbooks use (a, b) to denote the gcd, rather than $\gcd(a, b)$.

Example 6.3 $\gcd(12, 20) = 4$ because $4|12$ and $4|20$, but there is no common divisor of 12 and 20 that is greater than 4.

Activity 6.1 Convince yourself that $\gcd(35, -77) = 7$.

If two numbers a, b have $\gcd(a, b) = 1$, then we say that a and b are *coprime*. In this case, a and b have no common factors other than 1 and -1. For example, 72 and 77 are coprime.

6.6 The Euclidean algorithm

There is a standard method for computing greatest common divisors, known as the *Euclidean algorithm*. Before presenting this, we state two important properties of greatest common divisors.

- If $b \in \mathbb{N}$ and $a \mid b$, then $\gcd(a, b) = a$.
- For non-zero integers a and b, if $a = bq + r$ where q, r are integers, then $\gcd(a, b) = \gcd(b, r)$.

The first fact is clear: for a is, in this case, a common divisor of a and b. And there's no greater positive divisor of a than a itself.

> **Activity 6.2** Prove this second fact by proving that the set of common divisors of a and b is the same as the set of common divisors of b and r (and hence both sets have the same greatest member.)

6

These observations provide a way to determine gcds. Let's think about a simple example. Suppose we want $\gcd(100, 15)$. (You can see immediately that the answer is 5, but let's try to explain a method that will be useful in rather less easy examples.)

We have $100 = 15 \times 6 + 10$ so, by the second fact above, $\gcd(100, 15) = \gcd(15, 10)$. Next, $15 = 10 \times 1 + 5$, so $\gcd(15, 10) = \gcd(10, 5)$. But $10 = 2 \times 5$, so 5 divides 10 and so, by the first of the two facts above, $\gcd(10, 5) = 5$. It follows, then, that $\gcd(100, 15) = 5$.

The method, known as the *Euclidean Algorithm*, therefore employs successive use of the division theorem. Here's another, more substantial, example.

> **Example 6.4** Let us calculate $\gcd(2247, 581)$. We have
>
> $$\begin{aligned} 2247 &= 581 \times 3 + 504 \\ 581 &= 504 \times 1 + 77 \\ 504 &= 77 \times 6 + 42 \\ 77 &= 42 \times 1 + 35 \\ 42 &= 35 \times 1 + 7 \\ 35 &= 7 \times 5. \end{aligned}$$
>
> It follows that $\gcd(2247, 581) = 7$. The *reason* is that these divisions establish the following equalities:
>
> $$\begin{aligned} \gcd(2247, 581) &= \gcd(581, 504) \\ &= \gcd(504, 77) \\ &= \gcd(77, 42) \\ &= \gcd(42, 35) \\ &= \gcd(35, 7) = 7, \end{aligned}$$
>
> this last equality because $7 \mid 35$.

6.7 Some consequences of the Euclidean algorithm

A useful consequence of the Euclidean algorithm is that we can use it to express d, the gcd of a and b, as an integer linear combination of a and b, by which we mean the following.

Theorem 6.3 Suppose a and b are integers (at least one of which is not 0) and let $d = \gcd(a, b)$. Then there are $m, n \in \mathbb{Z}$ such that $d = am + bn$.

For a formal, general, proof of this, see Biggs Section 8.4 or Eccles Section 17.1. I'll give an example here, which will demonstrate the way in which we can find m, n once we have applied the Euclidean algorithm. Let's work with $a = 2247$ and $b = 581$. We want to find integers m and n such that

$$2247m + 581n = 7.$$

We can use the calculation we had above, by 'working backwards' through the sequence of equations, as follows:

$$\begin{aligned} 7 &= 42 - 35 \\ &= 42 - (77 - 42) = 42 \times 2 - 77 \\ &= (504 - 77 \times 6) \times 2 - 77 = 504 \times 2 - 77 \times 13 \\ &= 504 \times 2 - (581 - 504) \times 13 = 504 \times 15 - 581 \times 13 \\ &= (2247 - 581 \times 3) \times 15 - 581 \times 13 = 2247 \times 15 - 581 \times 58 \\ &= 2247 \times 15 + 581 \times (-58). \end{aligned}$$

6

So we see that $m = 15$ and $n = -58$ will work. Not an answer you could easily have guessed! Notice how, in each line of this calculation, we use one of the lines from the calculation that arises from the Euclidean algorithm (and we also simplify as we go along). You will get used to this with practice. There are many examples you could make up for yourself in order to help you practice.

> **Activity 6.3** In the above procedure to find m and n, figure out exactly which part of the Euclidean algorithm calculation is being used at each stage.

> **Activity 6.4** Choose two particular positive integers a and b. Use the Euclidean algorithm to find $\gcd(a, b)$ and then use your calculation to find integers m and n such that $d = ma + nb$. Do this several times with different choices of numbers until you have mastered it.

The fact that there are $m, n \in \mathbb{Z}$ such that $\gcd(a, b) = am + bn$ (that is, that the gcd of two integers is an integer linear combination of them) is very useful. Here's one nice consequence.

We know that, by definition, if $c \mid a$ and $c \mid b$, then $c \leq d = \gcd(a, b)$. But we can say something stronger, namely that $c \mid d$.

Theorem 6.4 Suppose that $a, b \in \mathbb{Z}$ so that a and b are not both zero, and let $d = \gcd(a, b)$. If $c \mid a$ and $c \mid b$, then $c \mid d$.

Proof. There are integers m, n such that $d = ma + nb$. Suppose that $c \,|\, a$ and $c \,|\, b$. Then $c \,|\, ma$ and $c \,|\, nb$, so $c \,|\, (ma + nb)$. But this says $c \,|\, d$, as required. □

Here's another consequence:

Theorem 6.5 For $a, b \in \mathbb{Z}$, with a, b not both zero, let $d = \gcd(a, b)$. Then, for $c \in \mathbb{N}$, there are integers m and n such that $c = am + bn$ **if and only if** $d \,|\, c$.

Proof. Suppose $c = am + bn$. Now, $d = \gcd(a, b)$ satisfies $d \,|\, a$ and $d \,|\, b$, so, also, $d \,|\, (ma + nb)$ and hence $d \,|\, c$.

Conversely, suppose $d \,|\, c$, so that for some integer k, $c = kd$. Now, there are $m, n \in \mathbb{Z}$ with $d = ma + nb$. Then,

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

where $M, N \in \mathbb{Z}$. This shows that c can be written as an integer linear combination of a and b, as required. □

We also have:

Theorem 6.6 Suppose that $a, b \in \mathbb{N}$ are coprime (meaning $\gcd(a, b) = 1$). If $a \,|\, r$ and $b \,|\, r$, then $ab \,|\, r$.

This is not generally true if the numbers a, b are not coprime. Think of a counterexample!

Proof. Because $\gcd(a, b) = 1$, there are integers m, n such that $1 = ma + nb$. So

$$r = r \times 1 = r(ma + nb) = mra + nrb.$$

Because $a \,|\, r$ and $b \,|\, r$, there are integers k_1, k_2 such that $r = k_1 a$ and $r = k_2 b$. So

$$r = mra + nrb = m(k_2 b)a + n(k_1 a)b = (mk_2 + nk_1)ab,$$

which shows that $ab \,|\, r$. □

6.8 Prime numbers

A *prime number* (or a *prime*) is a natural number $p \geq 2$ with the property that the only divisors of p are 1 and p. In a precise sense, which we'll see shortly, primes are the building blocks of the natural numbers.

One important property of primes is that if a prime divides a product of numbers, then it must divide at least one of the numbers in the product. This isn't true for non-primes: for example, $4 \,|\, 12 = 2 \times 6$, but 4 does not divide either 2 or 6.

Theorem 6.7 Suppose that p is a prime number and that $a, b \in \mathbb{N}$. If $p \,|\, ab$, then $p \,|\, a$ or $p \,|\, b$.

Proof. The proof makes use of the useful fact (seen above) that the gcd of any two numbers can be written as an integer linear combination of the numbers (Theorem 6.3). Suppose, then, that p is prime and that $p \,|\, ab$. If $p \,|\, a$, then the conclusion of the theorem holds, so suppose $p \nmid a$. Then p and a have no common positive divisor other than 1. (The only positive divisors of p are 1 and p because it is prime, and p does not divide a, by assumption.) So $\gcd(p, a) = 1$ and therefore there exist integers m and n with $1 = mp + na$. So, $b = b(mp + na) = (bm)p + n(ab)$. Now, $p \,|\, ab$ and hence $p \,|\, n(ab)$. Clearly, also, $p \,|\, (bm)p$. It follows that $p \,|\, b$, as required. □

This result can easily be extended: if $a_1, a_2, \ldots, a_n \in \mathbb{N}$ and $p \,|\, a_1 a_2 \ldots a_n$, then, for some i between 1 and n, $p \,|\, a_i$.

> **Activity 6.5** Prove this generalisation of Theorem 6.7.

6.9 Prime factorization: the Fundamental Theorem of Arithmetic

6.9.1 The Fundamental Theorem

The *Fundamental Theorem of Arithmetic* is the name given to the following Theorem. (As its name suggests, this is an important theorem!)

Theorem 6.8 (Fundamental Theorem of Arithmetic) Every integer $n \geq 2$ can be expressed as a product of one or more prime numbers. Furthermore, there is essentially only one such way of expressing n: the only way in which two such expressions for n can differ is in the ordering of the prime factors.

The expression of an integer as a product of primes is known as its prime decomposition. For example, the prime decomposition of 504 is

$$504 = 2 \times 2 \times 2 \times 3 \times 3 \times 7 = 2^3.3^2.7.$$

(Note that, in this last expression, the dot, '.' denotes multiplication.)

The proof of the Fundamental Theorem is not very difficult, given the results we already have about prime numbers. Establishing that each positive integer can be written as a product of primes is easy. Showing that such a decomposition is essentially unique (that is, unique up to the ordering of the factors) is a little trickier, but can be established using Theorem 6.7.

6.9.2 Proof of the Fundamental Theorem

There are two things to prove:

- Any $n \geq 2$ can be written as a product of primes: $n = p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r}$, where $p_1 < p_2 < \cdots < p_r$ are primes and $k_1, k_2, \ldots, k_r \in \mathbb{N}$. ('Existence')
- This is essentially unique: if $p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r} = q_1^{l_1} q_2^{l_2} \ldots q_s^{l_s}$, are two equal such expressions, then $r = s$, $p_i = q_i$ and $k_i = l_i$ for all i. ('Uniqueness').

*

Fundamental Theorem: 'Existence'

We use (strong) induction. For $n \geq 2$, let $P(n)$ be the statement:

n can be written as a product of primes

Base case: $n = 2$ is a product of a single prime, so $P(2)$ is true.

Assume, inductively, that $k \in \mathbb{N}$ and that $P(s)$ is true for all $s \leq k$. (We're using strong induction.) Consider $k + 1$. This could be a prime number, in which case we're done and $P(k + 1)$ is true. Otherwise $k + 1 = ab$ where $1 < a, b < k + 1$. But then $P(a)$ and $P(b)$ are true (by assumption) so each of a, b is a product of primes. So, therefore, is $k + 1$.

*

Fundamental Theorem: 'Uniqueness'

The idea here is simple enough, but it is notationally difficult to write a detailed proof. I'll try to give you the basic idea.

Suppose $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} = q_1^{l_1} q_2^{l_2} \dots q_s^{l_s}$. Cancel as much as possible all (powers of) common primes. We want to show both resulting sides are 1 (so that the expressions were the same.)

Suppose not. Then the resulting LHS (left-hand side) is a product of some powers of some p_i and the RHS is a product of some powers of some q_i, and no p_i equals any q_j (because we've cancelled). Take any p appearing in the LHS. Then $p \,|\, \text{RHS}$ so (by an earlier result, 'If $p \,|\, a_1 a_2 \dots a_n$, then, for some i, $p \,|\, a_i$') p divides one of the q_j. This isn't possible since $p_i \neq q_j$.

We can use the Fundamental Theorem of Arithmetic to prove that there are infinitely many primes. You can find an outline of this proof in Chapter 2.

> **Activity 6.6** Look again at the proof, in Chapter 2, that there are infinitely many primes, and understand where the Fundamental Theorem of Arithmetic is used in the proof.

6.10 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- state clearly what it means to say that one number divides another
- state the Division Theorem and, given two numbers, find the remainder and quotient when one is divided by the other
- understand what is meant by the representation of an integer with respect to a particular basis and be able to work with this definition
- state the definition of the greatest common divisor of two numbers

- use the Euclidean algorithm to find the gcd of two numbers
- demonstrate that you know that the gcd of two numbers can always be expressed as an integer linear combination of the two numbers; and be able to express the gcd in this way for any two numbers.
- use the Euclidean algorithm to express the gcd of two numbers as an integer linear combination of them
- state what is meant by a prime number
- state the Fundamental Theorem of Arithmetic.

6.11 Sample exercises

Exercise 6.1
Find $d = \gcd(2406, 654)$. Express d in the form $d = 2406m + 654n$ for integers m, n. □

Exercise 6.2
Suppose that $a, b \in \mathbb{N}$, both non-zero, and let $d = \gcd(a, b)$. We know that, by definition, if $c \,|\, a$ and $c \,|\, b$, then $c \leq d$. Prove, in fact, that $c \,|\, d$. □

Exercise 6.3
Suppose $a, b \in \mathbb{N}$ and that $d = \gcd(a, b)$. Prove that, for $c \in \mathbb{N}$, there are integers m and n such that $c = am + bn$ *if and only if* $d \,|\, c$.

Exercise 6.4
Suppose $a, b \in \mathbb{N}$. Prove that if there are integers m and n such that $am + bn = 1$ then a and b are coprime. □

Exercise 6.5
Prove that for all $n \in \mathbb{N}$, the numbers $9n + 8$ and $6n + 5$ are coprime.

Exercise 6.6
Suppose that $a, b \in \mathbb{N}$ and that $\gcd(a, b) = 1$. Suppose that $a \,|\, r$ and $b \,|\, r$. Prove that $ab \,|\, r$.

Exercise 6.7
The Fibonacci numbers $f_1, f_2, f_3, \dots$ are defined as follows: $f_1 = f_2 = 1$ and, for $n \geq 3$, $f_n = f_{n-1} + f_{n-2}$. Prove that for all $n \in \mathbb{N}$, $\gcd(f_n, f_{n+1}) = 1$. □

Exercise 6.8

Suppose that $p_1, p_2, \dots, p_k$ are primes and that $a, b \in \mathbb{N}$ are given by

$$a = p_1^{l_1} p_2^{l_2} \dots p_k^{l_k}, \quad b = p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}.$$

Prove that

$$\gcd(a, b) = p_1^{r_1} p_2^{r_2} \dots p_k^{r_k},$$

where, for $i = 1$ to k, r_i is the smaller of the two numbers l_i and m_i. □

Exercise 6.9

Suppose $a, b \in \mathbb{N}$ satisfy $\gcd(a, b) = 1$ and, for some $k \in \mathbb{N}$, $ab = k^2$. Prove that for some integers m, n, $a = m^2$ and $b = n^2$.

6.12 Comments on selected activities

6

Learning activity 6.1 Certainly, 7 divides both 35 and -77, but there is no larger common divisor. (For, the only larger divisor of 35 is 35 itself, and this does not divide -77.)

Learning activity 6.2 We prove that $D(a, b) = D(b, r)$. The result on gcds will follow since $\gcd(x, y)$ is the maximal element of $D(x, y)$.

Suppose $m \in D(a, b)$. Then $m \,|\, a$ and $m \,|\, b$. It follows that $m \,|\, (a - bq)$; that is, $m \,|\, r$. So $m \,|\, b$ and $m \,|\, r$ and hence $m \in D(b, r)$. Therefore $D(a, b) \subseteq D(b, r)$.

Suppose $m \in D(b, r)$. Then $m \,|\, b$ and $m \,|\, r$. It follows that $m \,|\, (bq + r)$; that is, $m \,|\, a$. So $m \,|\, b$ and $m \,|\, a$ and hence $m \in D(a, b)$. Therefore $D(b, r) \subseteq D(a, b)$.

Learning activity 6.5 To prove this, we can (unsurprisingly) use induction on n. Let $P(n)$ be the statement:

If $a_1, a_2, \dots, a_n \in \mathbb{N}$ and $p \,|\, a_1 a_2 \dots a_n$, then, for some i between 1 and n, $p \,|\, a_i$.

Then $P(1)$ is clearly true (and $P(2)$ is the theorem just proved). Suppose $P(n)$ is true and let's show $P(n + 1)$ follows. So, suppose $a_1, a_2, \dots, a_{n+1} \in \mathbb{N}$ and $p \,|\, a_1 a_2 \dots a_n a_{n+1}$. Well, since $p \,|\, A a_{n+1}$, where $A = a_1 a_2 \dots a_n$, we can apply the $n = 2$ case to see that $p \,|\, A$ or $p \,|\, a_{n+1}$. But, by $P(n)$, if $p \,|\, A = a_1 a_2 \dots a_n$ then $p \,|\, a_i$ for some i between 1 and n. So we're done: p divides at least one of the a_i for i between 1 and $n + 1$.

Learning activity 6.6 Here's the proof, with explicit reference to the Fundamental Theorem.

Suppose there were *not* infinitely many primes, so there's a largest prime, M, say. Let

$$X = (2 \times 3 \times 5 \times 7 \times 11 \times \cdots \times M) + 1.$$

Since $X > M$, X is not a prime. By the Fundamental Theorem of Arithmetic, X has a prime divisor p which satisfies $1 < p < X$. This p must be one of the numbers $2, 3, 5, \dots, M$ (since these are the only primes). However, X is **not** divisible by any of these numbers. So we have a contradiction. We conclude there are infinitely many primes.

6.13 Solutions to exercises

Solution to exercise 6.1

See Biggs, Section 8.4. The gcd is 6 and we have $6 = 28 \times 2406 + (-103) \times 654$.

Solution to exercise 6.2

We know that there are integers m, n such that $d = ma + nb$. Suppose that $c \,|\, a$ and $c \,|\, b$. Then $c \,|\, ma$ and $c \,|\, nb$, so $c \,|\, (ma + nb)$. But this says $c \,|\, d$, as required. □

Solution to exercise 6.3

Suppose first that $c = am + bn$ for some some integers m and n. Now, $d = \gcd(a, b)$ satisfies $d \,|\, a$ and $d \,|\, b$, so we also have $d \,|\, (ma + nb)$ and hence $d \,|\, c$. Conversely, suppose $d \,|\, c$, so that for some integer k, $c = kd$. Now, the gcd $d = \gcd(a, b)$ can be written as an integer linear combination of a and b, so there are $m, n \in \mathbb{Z}$ with $d = ma + nb$. Then,

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

where $M, N \in \mathbb{Z}$. This shows that c can be written as an integer linear combination of a and b, as required. □

6

Solution to exercise 6.4

This follows from the previous exercise, but we can prove it directly. Suppose that $d \in \mathbb{N}$, that $d \,|\, a$ and $d \,|\, b$. Then $d \,|\, (am + bn)$, which means $d \,|\, 1$. Therefore, we must have $d = 1$. That is, the only positive common divisor of a and b is 1 and hence $\gcd(a, b) = 1$ and the numbers are coprime.

Solution to exercise 6.5

We have $2(9n + 8) - 3(6n + 5) = 1$. So, if $d = \gcd(9n + 8, 6n + 5)$, then $d \,|\, (9n + 8)$ and $d \,|\, (6n + 5)$, so $d \,|\, 2(9n + 8) - 3(6n + 5)$. But this says $d \,|\, 1$ and hence $d = 1$. □

Solution to exercise 6.6

Before we begin, let's just note that this property does not hold if a and b are not coprime. For example, $6 \,|\, 12$ and $4 \,|\, 12$ but $24 \nmid 12$. Suppose then that $a \,|\, r$ and $b \,|\, r$. The fact that $\gcd(a, b) = 1$ means that there are integers m, n such that $1 = ma + nb$. So

$$r = r \times 1 = r(ma + nb) = mra + nrb.$$

Now, because $a \,|\, r$ and $b \,|\, r$ there are integers k_1, k_2 such that $r = k_1 a$ and $r = k_2 b$. So

$$r = mra + nrb = m(k_2 b)a + n(k_1 a)b = (mk_2 + nk_1)ab,$$

which shows that r is an integer multiple of ab and hence $ab \,|\, r$.

Solution to exercise 6.7

We prove this by induction on n. Let $P(n)$ be the statement that $\gcd(f_n, f_{n+1}) = 1$. When $n = 1$, this is true, because $\gcd(f_1, f_2) = \gcd(1, 1) = 1$. It is true also when $n = 2$ because $f_3 = 2$ and hence $\gcd(f_2, f_3) = \gcd(1, 2) = 1$. Suppose, inductively, that $k \geq 2$ and $\gcd(f_k, f_{k+1}) = 1$. We want to show that $\gcd(f_{k+1}, f_{k+2}) = 1$. Now, $f_{k+2} = f_{k+1} + f_k$. Therefore, if $d \,|\, f_{k+1}$ and $d \,|\, f_{k+2}$ then $d \,|\, f_{k+1}$ and $d \,|\, (f_{k+2} - f_{k+1}) = f_k$. So any common divisor of f_{k+1} and f_{k+2} is also a common divisor or f_k and f_{k+1}. Also, if $d \,|\, f_k$ and $d \,|\, f_{k+1}$

then we also have $d\,|\,f_{k+1}$ and $d\,|\,(f_k + f_{k+1}) = f_{k+2}$, so any common divisor of f_k and f_{k+1} is also a common divisor or f_{k+1} and f_{k+2}. This all shows that the common divisors of the pair $\{f_{k+1}, f_{k+2}\}$ are precisely the same as the common divisors of the pair $\{f_{k+1}, f_k\}$. Therefore, the greatest common divisors of each pair are equal. That is,

$$\gcd(f_{k+1}, f_{k+2}) = \gcd(f_{k+1}, f_k) = \gcd(f_k, f_{k+1}) = 1,$$

where we have used the inductive hypothesis for the last equality.

You can also establish this result by thinking about the way in which the Euclidean algorithm would work in finding the gcd of f_k and f_{k+1}. □

Solution to exercise 6.8

Let $d = p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k}$. Then because r_i is the smaller of l_i and m_i, we have $r_i \leq l_i$ and $r_i \leq m_i$. So $p^{r_i}\,|\,p^{l_i}$ and $p^{r_i}\,|\,p^{m_i}$, for each i. Therefore $d\,|\,a$ and $d\,|\,b$. Explicitly, for example,

$$\begin{aligned} a &= p_1^{l_1} p_2^{l_2} \ldots p_k^{l_k} \\ &= p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k} \times p_1^{l_1 - r_1} p_2^{l_2 - r_2} \ldots p_k^{l_k - r_k} \\ &= d(p_1^{l_1 - r_1} p_2^{l_2 - r_2} \ldots p_k^{l_k - r_k}), \end{aligned}$$

and because $l_k - r_k$ is a non-negative integer for each i, the number in parentheses is an integer. This shows $d\,|\,a$. The fact that $d\,|\,b$ can be similarly shown. So d is a common divisor of a and b.

Suppose D is any common divisor of a and b. Then, by the Fundamental Theorem of Arithmetic, D can be written as a product of primes. Let p be any one of these. Then $p\,|\,a$ and $p\,|\,b$. Now, we know that if $p\,|\,a_1 a_2 \ldots a_n$ then $p\,|\,a_i$ for some i. (This follows from the results of Section 6.8.) The only primes appearing in the decomposition of a and b are $p_1, p_2, \ldots, p_k$, so we can deduce that for some i, $p\,|\,p_i$ which means $p = p_i$ (given that p and p_i are primes). So the prime decomposition of D is of the form

$$D = p_1^{s_1} p_2^{s_2} \ldots p_k^{s_k}$$

for some non-negative integers s_i. Now, suppose that, for some i, $s_i > l_i$. Then, for some integers M and N,

$$D = p_i^{s_i} M, \quad a = p_i^{l_i} N,$$

where, because they involve only products of the other primes, neither N or M is divisible by p_i. Now, the fact that $D\,|\,a$ means that there's an integer L with $a = LD$, so

$$p_i^{l_i} N = L p_i^{s_i} M$$

and hence

$$N = L p_i^{s_i - l_i} M.$$

But this shows, since $s_i - l_i \geq 1$ (because $s_i, l_i \in \mathbb{Z}$ and $s_I > l_i$), that $p_i\,|\,N$, contradicting the observation that $p_i \nmid N$. So we must have $s_i \leq l_i$ for all i. A similar argument shows that $s_i \leq m_i$ for all i. So $s_i \leq r_i = \min(l_i, m_i)$ and hence $D \leq d$. The result follows. □

Solution to exercise 6.9

We use the Fundamental Theorem of Arithmetic. Let k have prime decomposition $k = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$. Then

$$ab = k^2 = p_1^{2\alpha_1} p_2^{2\alpha_2} \ldots p_r^{2\alpha_r}.$$

It follows that a and b must have prime decompositions involving only the primes $p_1, p_2, \ldots, p_r$, and that each of a, b takes the form $p_1^{s_1} p_2^{s_2} \ldots p_r^{s_r}$ where s_i is a non-negative integer. But we cannot have, for any i, $p_i\,|\,a$ and also $p_i\,|\,b$, for this would mean that $p_i > 1$ is a common divisor of a, b, contradicting $\gcd(a, b) = 1$. So, for each i, $p_i^{2\alpha_i}$ divides precisely one of a and b and p_i does not divide the other of the two numbers. In other words, each of a, b takes the form $p_1^{2\beta_1} p_2^{2\beta_2} \ldots p_r^{2\beta_r}$ where $\beta_i = 0$ or $\beta_i = \alpha_i$. This can be written as $(p_1^{\beta_1} p_2^{\beta_2} \ldots p_r^{\beta_r})^2$, and hence there are integers m, n such that $a = m^2$ and $b = n^2$.

Chapter 7
Congruence and modular arithmetic

☞ Biggs, N. L. *Discrete Mathematics.* Chapter 13, Sections 13.1–13.3.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning.* Chapters 19–21.

7.1 Introduction

In this chapter, we study *congruence* and we describe *modular arithmetic.* This builds on the ideas and results on divisibility and equivalence relations that we met in earlier chapters.

You go to sleep at 10 o'clock and you sleep for 8 hours. At what time do you wake. Well, this is simple: you wake at 6 o'clock. What you're doing in this calculation is you're doing what's called *arithmetic modulo* 12. The answer is not $10 + 8 = 18$, because the clock re-starts once the hour of 12 is reached. This is a fairly simple idea, when expressed in these terms, and it's the key concept behind *modular arithmetic*, a topic later in this chapter. We now take a more abstract approach.

7.2 Congruence modulo m

7.2.1 The congruence relation

Suppose that m is a fixed natural number, and let's define a relation R on the integers by $a\,R\,b$ if and only if $b - a$ is a multiple of m. That is, $a\,R\,b \iff m\,|\,(b-a)$. Then R is an equivalence relation.

Activity 7.1 Prove that R is an equivalence relation on $\mathbb{Z}$.

This relation is so important that it has a special notation. If $a\,R\,b$, we say that a and b are *congruent modulo m* and we write $a \equiv b \pmod{m}$.

If a and b are not congruent modulo m, then we write $a \not\equiv b \pmod{m}$.

The division theorem tells us that for any integers a and for any $m \in \mathbb{N}$, there are unique integers q and r such that

$$a = qm + r \quad \text{and} \quad 0 \leq r < m.$$

What this implies for congruence is that, for any a, there is precisely one integer r in the range $0, 1, \ldots, m-1$ such that $a \equiv r (\text{mod } m)$.

Congruence relations can be manipulated in many ways like equations, as the following Theorem shows.

Theorem 7.1 Suppose that $m \in \mathbb{N}$ and that $a, b, c, d \in \mathbb{Z}$ with $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$. Then

(i) $a + c \equiv b + d \pmod{m}$

(ii) $a - c \equiv b - d \pmod{m}$

(iii) $ac \equiv bd \pmod{m}$

(iv) $\forall k \in \mathbb{Z},\ ka \equiv kb \pmod{m}$

(v) $\forall n \in \mathbb{N},\ a^n \equiv b^n \pmod{m}$

Proof. I leave (i) and (ii) for you to prove. Here's how to prove (iii): because $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, we have $m \,|\, (b - a)$ and $m \,|\, (d - c)$. So, for some integers k, l, $b - a = km$ and $d - c = lm$. That is, $b = a + km$ and $d = c + lm$. So

$$bd = (a + km)(c + lm) = ac + (kmc + alm + klm^2) = ac + m(kc + al + klm).$$

Now, $kc + al + klm \in \mathbb{Z}$, so $bd - ac = (kc + al + klm)m$ is a multiple of m; that is, $m \,|\, (bd - ac)$ and $ac \equiv bd \pmod{m}$. Part (iv) follows from (iii) by noting that $k \equiv k \pmod{m}$, and part (v) follows by repeated application of (iii) (or, by (iii) and induction.). □

Activity 7.2 Prove parts (i) and (ii) of Theorem 7.1.

Theorem 7.1 is useful, and it enables us to solve a number of problems. Here are two examples.

Example 7.1 Suppose that the natural number x has digits $x_n x_{n-1} \dots x_0$ (when written, normally, in 'base 10'). So, for example, if $x = 1246$ then $x_0 = 6, x_1 = 4, x_2 = 2, x_3 = 1$. Then 9 divides x if and only if $x_0 + x_1 + \cdots + x_n$ is a multiple of 9. (So this provides a quick and easy way to check divisibility by 9. For example, 127224 is divisible by 9 because $1 + 2 + 7 + 2 + 2 + 4 = 18$ is.)

How do we prove that this test works? We can use congruence modulo 9. Note that

$$x = x_0 + (10)x_1 + (10)^2 x_2 + \cdots + (10)^n x_n.$$

Now, $10 \equiv 1 \pmod{9}$, so, for each $k \in \mathbb{N}$, $10^n \equiv 1 \pmod{9}$. Hence

$$x = x_0 + (10)x_1 + (10)^2 x_2 + \cdots + (10)^n x_n \equiv x_0 + x_1 + \cdots + x_n \pmod{9}.$$

A number is divisible by 9 if and only if it is congruent to 0 modulo 9, so

$$\begin{aligned} 9 \,|\, x &\iff x \equiv 0 \pmod{9} \\ &\iff x_0 + x_1 + \cdots + x_n \equiv 0 \pmod{9} \\ &\iff 9 \,|\, (x_0 + x_1 + \cdots + x_n). \end{aligned}$$

This is precisely what the test says.

Example 7.2 We can use congruence to show that there are no integers x and y satisfying the equation $7x^2 - 15y^2 = 1$.

We prove this by contradiction. So suppose such x and y did exist. Then, because $15y^2$ is a multiple of 5, we'd have $7x^2 \equiv 1 \pmod{5}$. Now, x is congruent to one of the numbers $0, 1, 2, 3, 4$ modulo 5. That is, we have:

$$x \equiv 0 \text{ or } x \equiv 1 \text{ or } x \equiv 2 \text{ or } x \equiv 3 \text{ or } x \equiv 4 \pmod{5}.$$

So,

$$x^2 \equiv 0 \text{ or } x^2 \equiv 1 \text{ or } x^2 \equiv 4 \text{ or } x^2 \equiv 9 \text{ or } x^2 \equiv 16 \pmod{5}.$$

But $9 \equiv 4 \pmod{5}$ and $16 \equiv 1 \pmod{5}$, so, in every case, either $x^2 \equiv 0$ or $x^2 \equiv 1$ or $x^2 \equiv 4 \pmod{5}$. It follows, then, that in all cases, we have

$$7x^2 \equiv 0 \text{ or } 7x^2 \equiv 7 \equiv 2 \text{ or } 7x^2 \equiv 28 \equiv 3 \pmod{5}.$$

So there does not exist an integer x with $7x^2 \equiv 1 \pmod{5}$, and hence there are no integer solutions to the original equation.

7.2.2 Congruence classes

What are the equivalence classes of the congruence relation, modulo a particular positive integer m? These are often called the *congruence classes modulo m*. Let's denote these by $[x]_m$. For a particular $x \in \mathbb{Z}$, $[x]_m$ will be all the integers y such that $y \equiv x \pmod{m}$. We know that each x is congruent to precisely one of the integers in the range $0, 1, 2, \dots, m - 1$, and we know (from the general theory of equivalence relations) that if $x \equiv y \pmod{m}$ then $[x]_m = [y]_m$. So it follows that for each $x \in \mathbb{Z}$, we'll have

$$[x]_m = [0]_m, \text{ or } [x]_m = [1]_m, \dots \text{ or } [x]_m = [m-1]_m.$$

So there are precisely m equivalence classes,

$$[0]_m, [1]_m, \dots, [m-1]_m.$$

The theory of equivalence relations tells us that these form a partition of $\mathbb{Z}$: they are disjoint and every integer belongs to one of them. But what are they? Well, $[0]_m$ is the set of all x such that $x \equiv 0 \pmod{m}$, which means $m \,|\, x$. So $[0]_m$ is the set of all integers divisible by m. Generally, for $0 \le r \le m - 1$, $[r]_m$ will be the set of x such that $x \equiv r \pmod{m}$. It follows that $[r]_m$ is the set of all integers x which have remainder r on division by m.

Example 7.3 Suppose $m = 4$. Then the congruence classes are:

$$[0]_4 = \{\dots, -8, -4, 0, 4, 8, \dots\},$$

$$[1]_4 = \{\dots, -7, -3, 1, 5, 9, \dots\},$$

$$[2]_4 = \{\dots, -6, -2, 2, 6, 10, \dots\},$$

$$[3]_4 = \{\dots, -5, -1, 3, 7, 11, \dots\}.$$

7.3 $\mathbb{Z}_m$ and its arithmetic

When, in an earlier chapter, we looked at how the integers may be *constructed* from the natural numbers through using an equivalence relation, we also saw that we could 'do arithmetic' with the equivalence classes. We can also do this here, and the resulting addition and multiplication operations are known as *modular arithmetic*.

First, let's introduce a new piece of notation. For $m \in \mathbb{N}$, $\mathbb{Z}_m$ is called the set of *integers modulo* m, and is the set of equivalence classes

$$[0]_m, [1]_m, \ldots, [m-1]_m.$$

So $\mathbb{Z}_m$ has m members. We can define operations $\oplus$ and $\otimes$ on $\mathbb{Z}_m$ as follows:

$$[x]_m \oplus [y]_m = [x+y]_m, \quad [x]_m \otimes [y]_m = [xy]_m.$$

For example, when $m = 4$, $[2]_4 \oplus [3]_4 = [5]_4 = [1]_4$ and $[2]_4 \otimes [3]_4 = [6]_4 = [2]_4$.

In practice, if we do not use the $\oplus$ and $\otimes$ symbols and we simply write x instead of $[x]_m$, using values of x between 0 and $m-1$. We would say that we are 'in $\mathbb{Z}_m$' if that's not clear from the context. So the above two calculations may be written:

$$\text{in } \mathbb{Z}_4, \quad 2+3 = 1, \quad \text{and } 2 \times 3 = 2.$$

Note that we use only the symbols $0, 1, \ldots, m-1$, so we do *not* write $2 + 3 = 5$. Instead we replace 5 by the number that it is congruent to modulo 4 and which lies between 0 and 3.

The equations we've just written are entirely equivalent to the statements

$$2 + 3 \equiv 1 \pmod 4, \quad 2 \times 3 \equiv 2 \pmod 4.$$

The addition and multiplication operations we've defined on $\mathbb{Z}_m$ obey a number of rules that are familiar from normal addition and multiplication of integers.

Theorem 7.2 Let $m \in \mathbb{N}$. In $\mathbb{Z}_m$, for all a, b, c,

(i) $a + b = b + a$

(ii) $a \times b = b \times a$

(iii) $(a + b) + c = a + (b + c)$

(iv) $(a \times b) \times c = a \times (b \times c)$

(v) $a + 0 = a$

(vi) $a \times 1 = a$

(vii) $a \times (b + c) = (a \times b) + (a \times c)$

(viii) for each $a \in \mathbb{Z}_m$ there is a unique element $-a \in \mathbb{Z}_m$ such that $a + (-a) = 0$.

Let's think a little about rule (viii) in this Theorem. Suppose we're in $\mathbb{Z}_4$ and that $a = 3$. What is $-a$? Well, what we want is an element of $\mathbb{Z}_m$ which when added to 3 gives 0 when we are doing arithmetic in $\mathbb{Z}_4$. Now, $3 + 1 = 0$ in $\mathbb{Z}_4$ because $3 + 1$ is congruent to 0 modulo 4, so $-3 = 1$. (Alternatively, we can note that $-3 = -1(4) + 1$, so -3 has remainder 1 on division by 4, so $-3 \equiv 1 \pmod 4$.)

Activity 7.3 In $\mathbb{Z}_9$, what is -4?

It is important to realise that arithmetic in $\mathbb{Z}_m$ does not obey *all* the nice properties that normal arithmetic of integers obeys. In particular, we cannot generally *cancel*. For example, in $\mathbb{Z}_4$,

$$2 \times 3 = 2 = 2 \times 1,$$

but we cannot 'cancel the 2' (that is, divide both sides by 2) to deduce that $3 = 1$, because $3 \neq 1$ in $\mathbb{Z}_4$. (The reason we cannot 'cancel' the 2 is that 2 has no inverse in $\mathbb{Z}_4$. Existence of inverses is the topic of the next section.)

7.4 Invertible elements in $\mathbb{Z}_m$

A member x of $\mathbb{Z}_m$ is *invertible* if there is some $y \in \mathbb{Z}_m$ such that (in $\mathbb{Z}_m$) $xy = yx = 1$. If such a y exists, it is called the *inverse* of x and is denoted by x^{-1}.

Example 7.4 In $\mathbb{Z}_{10}$, 3 has inverse 7 because, in $\mathbb{Z}_{10}$, $3 \times 7 = 1$ (because $3 \times 7 = 21 \equiv 1 \pmod{10}$).

Example 7.5 In $\mathbb{Z}_{10}$, 5 has no inverse. There is no x such that $5x = 1$. For, modulo 10, for any $x \in \mathbb{Z}$, $5x \equiv 0$ or 5. (This is just the familiar fact that any multiple of 5 has last digit 0 or 5.)

If $x \in \mathbb{Z}_m$ is invertible, then it is possible to *cancel* x from both sides of an equation in $\mathbb{Z}_m$. That is, we have

$$xa = xb \Rightarrow a = b \quad (\text{in } \mathbb{Z}_m).$$

This is not, as we have seen, generally true, but it works when x has an inverse because in this case

$$xa = xb \Rightarrow x^{-1}(xa) = x^{-1}(xb) \Rightarrow (x^{-1}x)a = (x^{-1}x)b \Rightarrow 1a = 1b \Rightarrow a = b.$$

Which $x \in \mathbb{Z}_m$ are invertible? The answer is given by the following theorem.

Theorem 7.3 Suppose $m \in \mathbb{N}$. Then an element x of $\mathbb{Z}_m$ is invertible if and only if x and m are coprime (that is, $\gcd(x, m) = 1$).

Proof. Suppose x is invertible, so that there is y with $xy = 1$ in $\mathbb{Z}_m$. This means that $xy \equiv 1 \pmod m$, so, for some $k \in \mathbb{Z}$, $xy = 1 + km$. Let $d = \gcd(x, m)$. Then $d|x$ and $d|m$, so $d|(xy - km)$. That is, $d|1$, from which it follows that $d = 1$ and x and m are coprime.

Conversely, suppose $\gcd(x, m) = 1$. Then (by the fact that the gcd can be written as an integer linear combination), there are integers y, z such that $1 = yx + zm$. But this means $yx \equiv 1 \pmod m$, so, in $\mathbb{Z}_m$, $yx = 1$ and x has inverse y. □

As a result of this theorem, we see that if p is a prime, then every non-zero element x of $\mathbb{Z}_p$ is invertible. This is because $\gcd(x, p) = 1$.

7.5 Solving equations in $\mathbb{Z}_m$

7.5.1 Single linear equations

Suppose we want to solve, in $\mathbb{Z}_m$, the equation $ax = b$. That is, we want to find x between 0 and $m - 1$ such that $ax \equiv b \pmod{m}$. This may have no solutions. Indeed, suppose we take $b = 1$. Then the equation we're confronted with is $ax = 1$, which has a solution if and only if a is invertible (by definition of inverse). So if a has no inverse in $\mathbb{Z}_m$, then such a linear equation will not always have a solution. If, however, a is invertible, then we can see that the equation $ax = b$ in $\mathbb{Z}_m$ has solution $x = a^{-1}b$, because $a(a^{-1}b) = (aa^{-1})b = 1b = b$.

How do you find a solution? Trial and error is not efficient if the numbers involved are large. But we can use the Euclidean algorithm. For suppose we want to solve $ax = b$ in $\mathbb{Z}_m$, where $\gcd(a, m) = 1$. Then, by using the Euclidean algorithm, we have seen how we can find integers k, l such that $1 = ak + ml$. So, $b = abk + mlb$, from which it follows that, modulo m, $a(bk) \equiv 1$. So it looks like bk will be a solution. But remember that we're looking for a solution in $\mathbb{Z}_m$, so we will want to find $x \in \mathbb{Z}_m$ such that $x \equiv bk \pmod{m}$. Here's an example.

Example 7.6 Suppose we want to solve the equation $83x = 2$ in $\mathbb{Z}_{321}$. We can check that 83 and 321 are coprime by the Euclidean algorithm, as follows: We have

$$\begin{aligned} 321 &= 83 \times 3 + 72 \\ 83 &= 72 \times 1 + 11 \\ 72 &= 11 \times 6 + 6 \\ 11 &= 6 \times 1 + 5 \\ 6 &= 5 \times 1 + 1 \\ 5 &= 1 \times 5. \end{aligned}$$

It follows that $\gcd(321, 83) = 1$. Now, working backwards, we can express 1 as an integer linear combination of 83 and 321:

$$\begin{aligned} 1 &= 6 - 5 \\ &= 6 - (11 - 6) = 6 \times 2 - 11 \\ &= (72 - 11 \times 6) \times 2 - 11 = 72 \times 2 - 11 \times 13 \\ &= 72 \times 2 - (83 - 72) \times 13 = 72 \times 15 - 83 \times 13 \\ &= (321 - 83) \times 3) \times 15 - 83 \times 13 = 321 \times 15 - 83 \times 58. \end{aligned}$$

This tells us that

$$83 \times (-58) \equiv 1 \pmod{321}.$$

So,

$$83 \times (-116) \equiv 2 \pmod{321}.$$

Now, we want to find x in the range 0 to 321 such that $-116 \equiv x \pmod{321}$. The answer is $x = 205$. So, finally, then, we see that, in $\mathbb{Z}_{321}$, the equation $83x = 2$ has solution $x = 205$.

Activity 7.4 This calculation also reveals that 83 is invertible in $\mathbb{Z}_{321}$. Why? And what is the inverse of 83 in $\mathbb{Z}_{321}$?

More generally, we can ask: when does $ax = b$ have a solution in $\mathbb{Z}_m$?

The answer is: when $d = \gcd(a, m)$ divides b.

So a special case is when $d = 1$.

That is, we have the following theorem.

Theorem 7.4 In $\mathbb{Z}_m$, $ax = b$ has a solution if and only if $d \mid b$, where $d = \gcd(a, m)$.

Proof. First part, $\Longrightarrow$: Suppose $ax_0 = b$ in $\mathbb{Z}_m$. Then $ax_0 - b = km$ for some $k \in \mathbb{Z}$. So, $b = ax_0 - km$. Since $d = \gcd(a, m)$, $d \mid a$ and $d \mid m$, so $d \mid (ax_0 - km)$. That is, $d \mid b$.

Second part, $\Longleftarrow$: Suppose $d \mid b$, so $b = db_1$ for some $b_1 \in \mathbb{Z}$. There are x_1, y_1 such that $d = x_1a + y_1m$. Then, $b = db_1 = (x_1b_1)a + (y_1b_1)m$.
So, in $\mathbb{Z}_m$, $a(x_1b_1) = b$. That is, x_1b_1 (reduced modulo m) is a solution. □

This theorem suggests a general method for solving $ax = b$ in $\mathbb{Z}_m$:

- Find $d = \gcd(a, m)$.
- If $d \not| b$, there's no solution.
- If $d \mid b$, write $b = db_1$. Use Euclidean algorithm to find $x, y \in \mathbb{Z}$ such that $d = xa + ym$. Then a solution is xb_1, reduced modulo m.

7.5.2 Systems of linear equations

We can also consider simultaneous linear equations in $\mathbb{Z}_m$. It should be realised that there might be no solutions, or more than one solution.

Example 7.7 Let's solve the following two equations simultaneously in $\mathbb{Z}_6$:

$$2x + 3y = 1, \quad 4x + 3y = 5.$$

Subtracting the first equation from the second gives $2x = 4$. You might be tempted to cancel the 2 and deduce that x must be 2. But wait! You can't cancel unless 2 and 6 are coprime, and they are not (since their gcd is 2). Instead, you can check for each of the elements of $\mathbb{Z}_6$ whether $2x = 4$. Of course, $x = 2$ is a solution, but so also is $x = 5$ because, in $\mathbb{Z}_6$, $2(5) = 10 = 4$. You can also check by calculating the other values of $2x$ that $x = 2, 5$ are the only solutions. Now, from the first equation, $3y = 1 - 2x$. When $x = 2$ or 5, $2x = 4$, so this is $3y = 1 - 4 = -3 = 3$. So we now have $3y = 3$. Again, we cannot cancel the 3. Instead we check, for each $y \in \mathbb{Z}_6$, whether it is a solution, and we find that $1, 3$ and 5 are all solutions. What this argument shows is that the *possible* solutions are

$$(x, y) = (2, 1), (2, 3), (2, 5), (5, 1), (5, 3), (5, 5).$$

In fact, it can easily be checked (by substituting these pairs of values into the original equations) that these are indeed solutions. So this system has 6 different solutions.

Activity 7.5 Check, by substituting into the original equations, that each of these six possible solution pairs (x, y) is indeed a solution.

Example 7.8 Consider the following system of simultaneous equations in $\mathbb{Z}_7$:

$$3x + y = 1, \quad 5x + 4y = 1.$$

If we multiply the first equation by 4, we obtain $12x + 4y = 4$, which is the same (in $\mathbb{Z}_7$) as $5x + 4y = 4$. But the second equation says $5x + 4y = 1$. Since 1 and 2 are not equal in $\mathbb{Z}_7$, these equations are inconsistent, so there are no solutions to this system.

7.6 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- state the definition of the equivalence relation congruence modulo m
- prove that congruence modulo m is an equivalence relation
- demonstrate an understanding of the links between congruence modulo m and remainder on division by m
- state and prove standard properties of congruence
- apply congruence to show, for example, that equations have no solutions
- demonstrate an understanding of what the congruence classes are
- demonstrate an understanding of what $\mathbb{Z}_m$ means and how its addition and multiplication are defined
- find the negatives of elements of $\mathbb{Z}_m$
- state the definition of an invertible element of $\mathbb{Z}_m$
- demonstrate that you know an element x in $\mathbb{Z}_m$ is invertible if and only if x and m are coprime
- find the inverse of an invertible element
- solve linear equations and systems of linear equations in $\mathbb{Z}_m$

7.7 Sample exercises

Exercise 7.1
Show that $n \equiv 7 \pmod{12} \Rightarrow n \equiv 3 \pmod{4}$. Is the converse true? □

Exercise 7.2
Show that for all $n \in \mathbb{Z}$, $n^2 \equiv 0$ or $1 \pmod{3}$. Hence show that if 3 divides $x^2 + y^2$ then $3 \,|\, x$ and $3 \,|\, y$. Use this to prove that there are no integers x, y, z such that $x^2 + y^2 = 3z^2$, other than $x = y = z = 0$.

Exercise 7.3
Show that, for all $n \in \mathbb{N}$, $3^{3n+1} \equiv 3 \times 5^n \pmod{11}$ and that $2^{4n+3} \equiv 8 \times 5^n \pmod{11}$. Hence show that for all $n \in \mathbb{N}$, $11 \,|\, (3^{3n+1} + 2^{4n+3})$. □

Exercise 7.4
By working modulo 7, prove that $2^{n+2} + 3^{2n+1}$ is divisible by 7. (This result was proved in a different way, using induction, in the exercises at the end of Chapter 3)

Exercise 7.5
Prove that 290 is an invertible element of $\mathbb{Z}_{357}$ and find its inverse. □

Exercise 7.6
Solve the equation $10x = 3$ in $\mathbb{Z}_{37}$. □

7.8 Comments on selected activities

Learning activity 7.1 We have $m \,|\, 0 = a - a$, so the relation is reflexive. Symmetry follows from $m \,|\, (b - a) \iff m \,|\, (a - b)$. Suppose that $a \,R\, b$ and $b \,R\, c$. Then $m \,|\, (b - a)$ and $m \,|\, (c - b)$, so $m \,|\, ((b - a) + (c - b)) = (c - a)$ and hence $a \,R\, c$. Thus, R is transitive.

Learning activity 7.2 We have $m \,|\, (b - a)$ and $m \,|\, (d - c)$. So $m \,|\, ((b - a) + (d - c))$, which is the same as $m \,|\, ((b + d) - (a + c))$, so $a + c \equiv b + d \pmod{m}$. We also have $m \,|\, ((b - a) - (d - c))$, which is the same as $m \,|\, ((b - d) - (a - c))$ so $a - c \equiv b - d \pmod{m}$.

Learning activity 7.3 Because $4 + 5 = 9 \equiv 0 \pmod{9}$, we have $-4 = 5$ in $\mathbb{Z}_9$.

Learning activity 7.4 The calculation shows that $83 \times (-58) \equiv 1 \pmod{321}$. We also know that $-58 \equiv 263 \pmod{321}$ because $58 + 263 = 312 \equiv 0 \pmod{321}$. So we'll have $83 \times 263 \equiv 1 \pmod{321}$. This shows that 83 is invertible in $\mathbb{Z}_{321}$ and that its inverse is 263.

7.9 Solutions to exercises

Solution to exercise 7.1

If $n \equiv 7 \pmod{12}$ then, for some integer k, $m = 7 + 12k$ and so $m = 3 + 4 + 12k = 3 + 4(1 + 3k)$. This means that $n \equiv 3 \pmod 4$, because $1 + 3k$ is an integer. The converse is false, because, for example, $3 \equiv 3 \pmod 4$, but $3 \not\equiv 7 \pmod{12}$

Solution to exercise 7.2

We have, modulo 3, $n \equiv 0$ or $n \equiv 1$ or $n \equiv 2$. So, respectively, $n^2 \equiv 0^2 = 0$ or $n^2 \equiv 1^2 = 1$ or $n^2 \equiv 2^2 = 4 \equiv 1$. So in all cases n^2 is congruent to 0 or 1. Suppose $3 \mid (x^2 + y^2)$. Then, modulo 3, $x^2 + y^2 \equiv 0$. But each of x^2 and y^2 is congruent to 0 or 1. If either or both are congruent to 1, then we'd have $x^2 + y^2 \equiv 1$ or $x^2 + y^2 \equiv 2$. So we can see that we must have $x^2 \equiv 0$ and $y^2 \equiv 0$. This means $x \equiv 0$ and $y \equiv 0$, which is the same as $3 \mid x$ and $3 \mid y$.

Now suppose that, for integers x, y, z, $x^2 + y^2 = 3z^2$, where not all of x, y, z are zero. If d is a common factor of x, y and z, then we can write $x = dx_1$, $y = dy_1$ and $z = dz_1$, where $x_1, y_1, z_1 \in \mathbb{Z}$. We can then see that $d^2x_1^2 + d_2y_1^2 = 3d^2z_1^2$, so that $x_1^2 + y_1^2 = 3z_1^2$. What this shows is that if there are any integer solutions, then there is one in which x, y, z have no common divisors (for any common divisors can be cancelled). So assume we're dealing with such a solution. Now, $x^2 + y^2 = 3z^2$ implies $3 \mid (x^2 + y^2)$ (noting that neither side of the equation is 0 because not all of x, y, z are). What we've shown earlier in this exercise establishes that $3 \mid x$ and $3 \mid y$. So $x = 3x_1$ and $y = 3y_1$ for some $x_1, y_1 \in \mathbb{Z}$. Then the equation $x^2 + y^2 = 3z^2$ becomes $9x_1^2 + 9y_1^2 = 3z^2$ and so $z^2 = 3x_1^2 + 3y_1^2$. This implies $3 \mid z^2$. But this means $3 \mid z$. (You can see this either by the Fundamental Theorem of Arithmetic, or by the fact that if $z \not\equiv 0$ modulo 3 then $z^2 \not\equiv 0$, as we see from the calculations at the start of this solution.) So what we see, then, is that x, y, z are all divisible by 3. But we assumed that they constituted a solution with no common factor and we've reached a contradiction. So there are no solutions other than $x = y = z = 0$.

Solution to exercise 7.3

Modulo 11, we have $3^3 = 27 \equiv 5$ and so $3^{3n} \equiv (3^3)^n \equiv 5^n$ and hence $3^{3n+1} = 3(3^n) \equiv 3 \times 5^n$. Also, $2^4 = 16 \equiv 5$ and so $2^{4n+3} = 8 \times (2^4)^n \equiv 8 \times 5^n$. It follows that

$$3^{3n+1} + 2^{4n+3} \equiv 3 \times 5^n + 8 \times 5^n = 11(5^n) \equiv 0 \pmod{11},$$

which means that $11 \mid (3^{3n+1} + 2^{4n+3})$. □

Solution to exercise 7.4

Modulo 7, $3^2 = 9 \equiv 2$, so $3^{2n+1} = 3(3^{2n}) \equiv 3(2^n)$ and $2^{n+2} = 4(2^n)$, so

$$2^{n+2} + 3^{2n+1} \equiv 4(2^n) + 3(2^n) = 7(2^n) \equiv 0,$$

and hence $7 \mid (2^{n+2} + 3^{2n+1})$. □

Solution to exercise 7.5

By the Euclidean algorithm, we have

$$\begin{aligned} 357 &= 290 + 67 \\ 290 &= 4 \times 67 + 22 \\ 67 &= 3 \times 22 + 1, \end{aligned}$$

so 290 and 357 are coprime, from which it follows that 290 is invertible. Now, from the calculations just given,

$$\begin{aligned} 1 &= 67 - 3 \times 22 \\ &= 67 - 3(290 - 4 \times 67) = 13 \times 67 - 3 \times 290 \\ &= 13(357 - 290) - 3 \times 290 = 13 \times 357 - 16 \times 290. \end{aligned}$$

The fact that $13 \times 357 - 16 \times 290 = 1$ means that, modulo 357, $-16 \times 290 \equiv 1$. So, in $\mathbb{Z}_{357}$, $(290)^{-1} = -16 = 341$.

Solution to exercise 7.6

Because 37 is prime, we certainly know that 10 and 37 are coprime, so the equation has a solution. The quickest way to find it is simply to note that the equation is equivalent to the congruence, modulo 37, that $10x \equiv 40$, and the 10 can then be cancelled because 10 and 37 are coprime. But suppose we didn't spot that. The Euclidean algorithm tells us that:

$$\begin{aligned} 37 &= 3 \times 10 + 7 \\ 10 &= 1 \times 7 + 3 \\ 7 &= 2 \times 3 + 1 \\ 3 &= 3 \times 1, \end{aligned}$$

and so

$$\begin{aligned} 1 &= 7 - 2 \times 3 \\ &= 7 - 2(10 - 7) = 3 \times 7 - 2 \times 10 \\ &= 3(37 - 3 \times 10) - 2 \times 10 \\ &= 3 \times 37 - 11 \times 10. \end{aligned}$$

So we see that $-11 \times 10 \equiv 1 \pmod{37}$ and hence $-33 \times 10 \equiv 3 \pmod{37}$. Now, $-33 \equiv 4 \pmod{37}$, so the solution is $x = 4$. This is easily checked: $10(4) = 40 = 3$ in $\mathbb{Z}_{37}$.□

Chapter 8
Rational, real and complex numbers

☞ Biggs, N. L. *Discrete Mathematics*. Chapter 9.

☞ Eccles, P.J. *An Introduction to Mathematical Reasoning*. Chapters 13 and 14.

The treatment in Biggs is probably better for the purposes of this course.

Neither of these books covers complex numbers. You do not have to know very much about complex numbers for this course, but because this topic is not in these books, I have included quite a bit of material on complex numbers in this chapter.

You can find useful reading on complex numbers in a number of books, including the following (which you might already have, given that it is the MA100 text).

☞ Anthony, M. and M. Harvey. *Linear Algebra: Concepts and Methods*. Cambridge University Press 2012. Chapter 13.

8.1 Introduction

In this chapter, we explore rational numbers, real numbers and complex numbers. In this course, we started with natural numbers and then we showed how to construct the set of all integers from these. This construction used an equivalence relation, together with a suitable way of adding and multiplying the equivalence classes. In a similar way, the rational numbers can be constructed from the integers by means of an equivalence relation. In this course, we do not take a very formal approach to the definition or construction of the real numbers (which can, in fact, be quite complicated). But we study properties of real numbers, and in particular we shall be interested in whether real numbers are rational or not. We also consider the 'cardinality' of infinite sets.

8.2 Rational numbers

8.2.1 An important equivalence relation

Rational numbers are, essentially, fractions. You'll certainly be aware that there are many ways of representing a given rational number. For instance, $\frac{2}{5}$ represents the same number as $\frac{4}{10}$. We can capture these sorts of equivalences more formally by using an equivalence relation on pairs of integers (m, n), where $n \neq 0$. So let $X = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ be the set of all pairs (m, n) where $m, n \in \mathbb{Z}$ and $n \neq 0$, and define a relation R on X by:

$$(m, n)\, R\, (m', n') \iff mn' = m'n.$$

(Yes, it looks like what we're really saying here is $m/n = m'/n'$, but we want to work in the world of the integers for now, so we don't want to do division.) Then we can *define* the set of rational numbers $\mathbb{Q}$ to be the set of equivalence classes of the relation R.

Let's just pause for a moment to prove that R is indeed an equivalence relation.

R **is Reflexive:** $(m,n)R(m,n)$ because $mn = nm$.

R **is Symmetric:** $(m,n)R(p,q) \Rightarrow mq = np \Rightarrow pn = qm \Rightarrow (p,q)R(m,n)$.

R **is Transitive:** Suppose $(m,n)R(p,q)$ and $(p,q)R(s,t)$. Then $mq = np$ and $pt = qs$. So, $(mq)(pt) = (np)(qs)$ and, after cancelling qp, this gives $mt = ns$, so $(m,n)R(s,t)$. But, wait a minute: can we cancel pq? Sure, if it's nonzero. If it *is* zero then that means $p = 0$ (since we know that $q \neq 0$). But then $mq = 0$, so $m = 0$; and $qs = 0$, so $s = 0$. So, in this case $mt = ns = 0$.

8.2.2 Rational numbers as equivalence classes

We represent the equivalence class $[(m,n)]$ by $\dfrac{m}{n}$. For example, we then have the (familiar) fact that $\dfrac{2}{5} = \dfrac{4}{10}$ which follows from the fact that $[(2,5)] = [(4,10)]$, something that is true because $(2,5)\,R\,(4,10)$ $(2 \times 10 = 4 \times 5)$. What we've done here is construct the set of rational numbers without reference to division. In an abstract approach, this is the logically sound thing to do. Once we have constructed the rational numbers, we can then make sense of the division of integers: the division of m by n is the rational number m/n.

What, for instance, is the equivalence class $[(1,2)]$? Well, $(m,n)R(1,2)$ means $m \times 2 = n \times 1$, or $n = 2m$. So it consists of

$$(1,2), (-1,-2), (2,4), (-2,-4), (3,6), (-3,-6)\ldots.$$

Denoting the equivalence class $[(m,n)]$ by $\dfrac{m}{n}$, we therefore have

$$\frac{1}{2} = \{(1,2), (-1,-2), (2,4), (-2,-4), (3,6), (-3,-6), \ldots\}.$$

Recall that if $x' \in [x]$ then $[x'] = [x]$. So we can say

$$\frac{1}{2} = \frac{-1}{-2} = \frac{2}{4} = \frac{-2}{-4} = \frac{3}{6} = \frac{-3}{-6} = \cdots.$$

We can think of the integers as particular rational numbers by identifying the integer n with the rational number $\dfrac{n}{1}$ (that is, with the equivalence class $[(n,1)]$). So $\mathbb{Z} \subseteq \mathbb{Q}$.

8.2.3 Doing arithmetic

How do we 'do arithmetic' with rational numbers. Well, you've been doing this for years, but how would we define addition and multiplication of rational numbers in an abstract setting? Just as we defined operations on equivalence classes in earlier chapters (in the construction of $\mathbb{Z}$ from $\mathbb{N}$ and in the construction of $\mathbb{Z}_m$), we can define addition and multiplication as an operation on the equivalence classes of R. Here's how: let $\oplus$ and $\otimes$ be defined on the set of rational numbers as follows:

$$\frac{a}{b} \oplus \frac{c}{d} = \frac{ad+bc}{bd}, \quad \frac{a}{b} \otimes \frac{c}{d} = \frac{ac}{bd}.$$

In practice, we just use normal addition and multiplication symbols (and we often omit the multiplication symbol), so we have

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \quad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}.$$

Well, no surprises there, but remember that what we are doing here is *defining* additional and multiplication of rational numbers (and remember also that these rational numbers are, formally, equivalence classes). Now, if you think hard about it, one issue that is raised is whether these definitions depend on the choice of representatives from each equivalence class. They should not, but we ought to check that. What I mean is that we really should have

$$\frac{2}{5} + \frac{2}{6} = \frac{4}{10} + \frac{1}{3},$$

for example, because

$$\frac{2}{5} = \frac{4}{10} \text{ and } \frac{2}{6} = \frac{1}{3}.$$

Well, let's see. Consider the addition definition. Suppose that

$$\frac{a}{b} = \frac{a'}{b'} \text{ and } \frac{c}{d} = \frac{c'}{d'}.$$

What we need to check is that

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

Now, the fact that $\dfrac{a}{b} = \dfrac{a'}{b'}$ means precisely that $[(a,b)] = [(a',b')]$, which means that $ab' = a'b$. Similarly, we have $cd' = c'd$. Now,

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \quad \text{and} \quad \frac{a'}{b'} + \frac{c'}{d'} = \frac{a'd'+b'c'}{b'd'}$$

and we need to prove that

$$\frac{ad+bc}{bd} = \frac{a'd'+b'c'}{b'd'}.$$

This means we need to prove that

$$(ad+bc, bd)\,R\,(a'd'+b'c', b'd').$$

Now,

$$\begin{aligned}(ad+bc, bd)\,R\,(a'd'+b'c', b'd') &\iff (ad+bc)b'd' = (a'd'+b'c')bd \\ &\iff adb'd' + bcb'd' = a'd'bd + b'c'bd \\ &\iff (ab')dd' + (cd')bb' = (a'b)dd' + (c'd)bb'.\end{aligned}$$

Now, the first terms on each side are equal to each other because $ab' = a'b$ and the second terms are equal to each other because $cd' = c'd$, so we do indeed have 'consistency' (that is, the definition of addition is independent of the choice of representatives chosen for the equivalence classes).

Activity 8.1 Show that the definition of multiplication of rational numbers is 'consistent': that is, that it does not depend on the choice of representatives chosen for the equivalence classes. Explicitly, show that if

$$\frac{a}{b} = \frac{a'}{b'} \text{ and } \frac{c}{d} = \frac{c'}{d'},$$

then

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

By the way, the rational numbers are described as such because they are (or, more formally, can be represented by) *ratios* of integers.

8.3 Rational numbers and real numbers

8.3.1 Real numbers: a 'sketchy' introduction

For our purposes, we will assume that the set of real numbers $\mathbb{R}$ is given. That is, we shall not construct the real numbers formally. (This can be done, but is outside the scope of this course.) We will regard the rational numbers (or, as we'll often call them, the rationals) $\mathbb{Q}$ as a subset of $\mathbb{R}$. There are various ways of thinking about real numbers. One way is to think about a real number as a point on the infinite real line. Another is to think of real numbers as the 'limits' of rational numbers. The formal idea of limit will be encountered in a later chapter, but we can encapsulate most of what we need by thinking about the decimal representation of real numbers.

First, let's note that if $a_i \in \mathbb{N} \cup \{0\}$ and $a_i \leq 9$ for $1 \leq i \leq n$, then the (finite) decimal expansion

$$a_0.a_1a_2 \dots a_n$$

represents the rational number

$$a_0 + \frac{a_1}{10} + \frac{a_2}{(10)^2} + \cdots + \frac{a_n}{(10)^n}.$$

For example, what we mean by 1.2546 is the number

$$1 + \frac{2}{10} + \frac{5}{100} + \frac{4}{1000} + \frac{6}{10000}.$$

Every positive real number can be represented by an infinite decimal expansion

$$a_0.a_1a_2a_3 \dots a_i \dots,$$

where $a_i \in \mathbb{N} \cup \{0\}$ and $a_i \leq 9$ for $i \geq 1$. We allow for a_i to be 0, so, in particular, it is possible that $a_i = 0$ for all $i \geq N$ where N is some fixed number: such an expansion is known as a *terminating* expansion. Given such an infinite decimal expansion, we say that it represents a real number a if, for all $n \in \mathbb{N} \cup \{0\}$,

$$a_0.a_1a_2 \dots a_n \leq a \leq a_0.a_1a_2 \dots a_n + 1/(10)^n.$$

This formalism allows us to see that the infinite decimal expansion $0.99999\dots$, all of whose digits after the decimal point are 9, is in fact the same as the number $1.0000000\dots$.

For example, two infinite decimal expansions are

$$3.1415926535\dots$$

and

$$0.1833333333333\dots.$$

(You'll probably recognise the first as being the number π.) Suppose, in this second decimal expansion, that every digit is 3 after the first three (that is, $a_i = 3$ for $i \geq 3$). Then we write this as $0.18\overline{3}$ (or, in some texts, $0.18\dot{3}$). We can extend this notation to cases in which there is a repeating pattern of digits. For example, suppose we have

$$0.1123123123123\dots,$$

where the '123' repeats infinitely. Then we denote this by $0.1\overline{123}$.

8.3.2 Rationality and repeating patterns

You probably have heard stories of strange, obsessive mathematicians working out the expansion of π to millions and millions of decimal places. (This has been the subject of a novel, a play, a film, and a song!) This is relevant because the digits of π have no repeating pattern, which you might think quite remarkable. In fact, it turns out that a real number will have an infinitely repeating pattern in its decimal expansion (which includes the case in which the pattern is 0, so that it includes terminating expansions) *if and only if* the number is rational.

Let's look at part of this statement: if a number is rational, then its decimal expansion will have a repeating pattern (which might be 0). Why is that? Well, let's look at an example.

Example 8.1 We find the decimal expansion of 4/7 by 'long-division'.

$$\begin{array}{rl}
 & 0.5714285\cdots \\
7 & \overline{)4.0000000} \\
 & \underline{3.5} \\
 & .50 \\
 & \underline{.49} \\
 & 10 \\
 & \underline{7} \\
 & 30 \\
 & \underline{28} \\
 & 20 \\
 & \underline{14} \\
 & 60 \\
 & \underline{56} \\
 & 40 \\
 & \underline{35} \\
 & 50
\end{array}$$

So,

$$4/7 = 0.\overline{571428}.$$

Notice: we must have the same remainder re-appear at some point, and then the calculation repeats. Here's the calculation again, with the repeating remainder highlighted in bold.

$$\begin{array}{r|l} & 0.5714285\cdots \\ \hline 7 & 4.0000000 \\ & \underline{3.5} \\ & \mathbf{.50} \\ & \underline{.49} \\ & 10 \\ & \underline{7} \\ & 30 \\ & \underline{28} \\ & 20 \\ & \underline{14} \\ & 60 \\ & \underline{56} \\ & 40 \\ & \underline{35} \\ & \mathbf{50} \end{array}$$

Next, we think about the second part of the statement: that if the decimal expansion repeats, then the number is rational.

Clearly, if the decimal expansion is terminating, then the number is rational. But what about the infinite, repeating, case? We've given two examples above. let's consider these in more detail.

Example 8.2 Consider $a = 0.18\overline{3}$. Let $x = 0.00\overline{3}$. Then $10x = 0.0\overline{3}$ and so $10x - x = 0.0\overline{3} - 0.00\overline{3} = 0.03$. So, $9x = 0.03$ and hence $x = (3/100)/9 = 1/300$, so

$$0.18\overline{3} = 0.18 + 0.00\overline{3} = \frac{18}{100} + \frac{1}{300} = \frac{55}{300} = \frac{11}{60},$$

and this is the rational representation of a.

Example 8.3 Consider the number $0.1\overline{123}$. If $x = 0.0\overline{123}$, then $1000x = 12.3\overline{123}$ and $1000x - x = 12.3$. So $999x = 12.3$ and hence $x = 123/9990$. So,

$$0.1\overline{123} = \frac{1}{10} + x = \frac{1}{10} + \frac{123}{9990} = \frac{1122}{9990}.$$

In general, if the repeating block is of length k, then an argument just like the previous two, in which we multiply by 10^k, will enable us to express the number as a rational number.

8.3.3 Irrational numbers

A real number is *irrational* if it is not a rational number. (So, given what we said above, an irrational number has no infinitely repeating pattern in its decimal expansion.) What's clear from above is that any real number can be approximated well by rational numbers: for the rational number $a_0.a_1a_2\ldots a_n$ is within $1/(10)^n$ of the real number with infinite decimal expansion $a_0.a_1a_2\ldots$.

We can, in some cases, prove that particular numbers are irrational. Here is a classic result of this type.

Theorem 8.1 The real number $\sqrt{2}$ is irrational. That is, there are no positive integers m, n with $\left(\frac{m}{n}\right)^2 = 2$.

Proof. Suppose there were such m, n. If m, n are divisible by some $d > 1$, we may divide both m and n to obtain m', n' such that the rational number m'/n' equals m/n. So we may assume that m, n have no common divisors other than 1; that is, $\gcd(m, n) = 1$. Now, the equation $(m/n)^2 = 2$ means $m^2 = 2n^2$. So we see that m^2 is even. We know (from Chapter 2) that this means m must be even. (For, the square of an odd integer is easily shown to be odd.) So there is some m_1 such that $m = 2m_1$. Then, $m^2 = 2n^2$ becomes $4m_1^2 = 2n^2$, and so $n^2 = 2m_1^2$. Well, this means n^2 is even and hence n must be even. So m, n are both divisible by 2. But we assumed that $\gcd(m, n) = 1$, and this is contradicted by the fact that m and n have 2 as a common divisor. So our assumption that $(m/n)^2 = 2$ must have been wrong and we can deduce no such integers m and n exist. □

Isn't this theorem a thing of beauty?

Activity 8.2 Make sure you understand that this is a proof by contradiction, and that you understand what the contradiction is.

Many other important numbers in mathematics turn out to be irrational. I've already mentioned π, and there is also e (the base of the natural logarithm).

8.3.4 'Density' of the rational numbers

As we've seen, some important numbers in mathematics are not rational. An intuitive question that arises is 'how many real numbers are rational' and this is a difficult question to answer. There are infinitely many real numbers and infinitely many rationals, and infinitely may real numbers are not rational. More on this in the next section!

For the moment, let's make one important observation: not only are there infinitely many rational numbers, but there are no 'gaps' in the rational numbers. If you accept the view of real numbers as (possibly) infinite decimal expansions, then this is quite clear: you can get a very good approximation to any real number by terminating its decimal expansion after a large number of digits. (And we know that a terminating decimal expansion is a rational number.) The following theorem makes sense of the statement that there are no 'rational-free' zones in the real numbers. Precisely, between any two rational numbers, no matter how close together they are, there is always another rational number.

Theorem 8.2 Suppose $q, q' \in \mathbb{Q}$ with $q < q'$. Then there is $r \in \mathbb{Q}$ with $q < r < q'$.

Proof. Consider $r = (1/2)(q + q')$. Details are left to you! □

Activity 8.3 Complete this proof.

8.4 Countability of rationals and uncountability of real numbers

We know that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$, and that each inclusion is strict (there are integers that are not natural numbers, rational numbers that are not integers, and real numbers that are not rational). All of these sets are infinite. But there is a sense in which the sets $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$ have the same 'size' and $\mathbb{R}$ is 'larger'. Clearly we have to define what this means in precise terms, because right now all we know is that there are more real numbers than rationals, for instance, but there are infinitely many of each type of number.

The following definition helps us.

Definition 8.1 (Countable sets) A set is *countable* if there is a bijection between the set and $\mathbb{N}$.

For instance, $\mathbb{Z}$ is countable: we can define $f : \mathbb{N} \to \mathbb{Z}$ by

$$f(1) = 0,\ f(2) = 1,\ f(3) = -1,\ f(4) = 2,\ f(5) = -2,\ f(6) = 3,\ f(7) = -3, \ldots .$$

(In general, $f(n) = (-1)^n \lfloor n/2 \rfloor$, where $\lfloor n/2 \rfloor$ means the largest integer that is no more than $n/2$.) It is straightforward to show that f is a bijection. Hence $\mathbb{Z}$ is countable. So, in this sense, the sets $\mathbb{Z}$ and $\mathbb{N}$ have the same 'cardinality' (even though $\mathbb{Z}$ is strictly larger than $\mathbb{N}$). Working with infinite sets is not the same as working with finite sets: two finite sets, one of which was a strict subset of the other, could not have the same cardinality!

What does 'countable' mean? The formal definition is given above. But one way of thinking about it is that if S is countable, then the members of S can be *listed*:

$$s_1, s_2, s_3, \ldots, .$$

For, suppose S is countable. Then there is a bijecion $f : \mathbb{N} \to S$. Let $s_i = f(i)$ for $i \in \mathbb{N}$. Then, because f is a bijection, the list $s_1, s_2, s_3, \ldots$ will include every element of S, each precisely once.

What is more surprising is that $\mathbb{Q}$ is also countable.

8.4.1 Countability of the rationals

Theorem 8.3 The set $\mathbb{Q}$ of rational numbers is countable.

How can we prove $\mathbb{Q}$ is countable?

By constructing a bijection $\mathbb{N} \to \mathbb{Q}$. We won't do so by means of a formula, but instead by thinking of a way in which all the rational numbers could be listed.

Arrange all the ordered pairs of natural numbers as follows:

$$\begin{array}{ccccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) & \cdots \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

Traverse this array as indicated in the following diagrams, where traversed elements are emboldened and underlined as they are traversed.

$$\begin{array}{ccccccc}
\underline{(\mathbf{1},\mathbf{1})} & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) & \cdots \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{(\mathbf{1},\mathbf{1})} & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{(\mathbf{2},\mathbf{1})} & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{(\mathbf{1},\mathbf{1})} & \underline{(\mathbf{1},\mathbf{2})} & (1,3) & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{(\mathbf{2},\mathbf{1})} & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
\underline{\mathbf{(3,1)}} & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
\underline{\mathbf{(3,1)}} & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
\underline{\mathbf{(4,1)}} & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & (2,3) & (2,4) & (2,5) & (2,6) & \cdots \\
\underline{\mathbf{(3,1)}} & \underline{\mathbf{(3,2)}} & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
\underline{\mathbf{(4,1)}} & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & (1,4) & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & \underline{\mathbf{(2,3)}} & (2,4) & (2,5) & (2,6) & \cdots \\
\underline{\mathbf{(3,1)}} & \underline{\mathbf{(3,2)}} & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
\underline{\mathbf{(4,1)}} & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

$$\begin{array}{ccccccc}
\underline{\mathbf{(1,1)}} & \underline{\mathbf{(1,2)}} & \underline{\mathbf{(1,3)}} & \underline{\mathbf{(1,4)}} & (1,5) & (1,6) & \cdots \\
\underline{\mathbf{(2,1)}} & \underline{\mathbf{(2,2)}} & \underline{\mathbf{(2,3)}} & (2,4) & (2,5) & (2,6) & \cdots \\
\underline{\mathbf{(3,1)}} & \underline{\mathbf{(3,2)}} & (3,3) & (3,4) & (3,5) & (3,6) & \cdots \\
\underline{\mathbf{(4,1)}} & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) & \cdots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) & \cdots \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

We get a listing of all the ordered pairs of natural numbers:

$$(1,1),(2,1),(1,2),(1,3),(2,2),(3,1),(4,1),(3,2),(2,3),\ldots$$

From this we can get a listing of all positive rational numbers m/n: simply write down the corresponding rationals and ignore any (m,n) such that the rational m/n has already earlier appeared in the list:

$$(1,1),(2,1),(1,2),(1,3),(2,2),(3,1),(4,1),(3,2),(2,3),\ldots$$

gives

$$\frac{1}{1},\frac{2}{1},\frac{1}{2},\frac{1}{3},\frac{2}{2},\frac{3}{1},\frac{4}{1},\frac{3}{2},\frac{2}{3},\frac{1}{4},\frac{1}{5},\frac{2}{4},\frac{3}{3},\frac{4}{2},\frac{5}{1},\frac{6}{1},\ldots$$

which becomes

$$\frac{1}{1},\frac{2}{1},\frac{1}{2},\frac{1}{3},\qquad\frac{3}{1},\frac{4}{1},\frac{3}{2},\frac{2}{3},\frac{1}{4},\frac{1}{5},\qquad\qquad\frac{5}{1},\frac{6}{1},\ldots$$

We can then include the negative rational numbers and 0 by starting with 0 and replacing m/n by m/n and $-m/n$:

$$0,\frac{1}{1},-\frac{1}{1},\frac{2}{1},-\frac{2}{1},\frac{1}{2},-\frac{1}{2},\frac{1}{3},-\frac{1}{3},\frac{3}{1},-\frac{3}{1},\frac{4}{1},-\frac{4}{1},\frac{3}{2},-\frac{3}{2},\frac{2}{3},-\frac{2}{3},$$
$$\frac{1}{4},-\frac{1}{4},\frac{1}{5},-\frac{1}{5},\frac{5}{1},-\frac{5}{1},\frac{6}{1},-\frac{6}{1},\ldots$$

It's clear that this listing describes a bijection from $\mathbb{N}$ to $\mathbb{Q}$. So this proves $\mathbb{Q}$ is countable.

So, informally speaking, $\mathbb{N},\mathbb{Z}$ and $\mathbb{Q}$ all have the same 'size'. What about $\mathbb{R}$? Well, here it gets very interesting: the set of real numbers is *not* countable. (It is said to be *uncountable*.)

8.4.2 Uncountability of the real numbers

Theorem 8.4 The set $\mathbb{R}$ is not countable. (That is, $\mathbb{R}$ is uncountable.)

This is probably not too surprising: a 'randomly' written decimal expansion will not terminate or have a repeating pattern, so, intuitively, 'most' real numbers are not rational.

The proof uses the famous 'Cantor diagonal' argument.

Suppose that $f : \mathbb{N} \to \mathbb{R}$ and that

$$f(n) = x_{n0}.x_{n1}x_{n2}x_{n3}\ldots.$$

We show there's a number in $\mathbb{R}$ which isn't the image under f of any element of $\mathbb{N}$ (and hence f is not a surjection). Consider

$$y = 0.y_1y_2y_3\ldots$$

where

$$y_i = \begin{cases} 1 & \text{if } x_{ii} \neq 1 \\ 2 & \text{if } x_{ii} = 1. \end{cases}$$

For all $n \in \mathbb{N}$, $y \neq f(n)$ since y_n is different from x_{nn}.

Since f was arbitrary, this shows that there can be no function $f : \mathbb{N} \to \mathbb{R}$ that is surjective. Hence $\mathbb{R}$ is not countable.

8.5 Complex numbers

8.5.1 Introduction

Consider the two quadratic polynomials,

$$p(x) = x^2 - 3x + 2 \qquad \text{and} \qquad q(x) = x^2 + x + 1$$

If you sketch the graph of $p(x)$ you will find that the graph intersects the x-axis at the two real solutions (or roots) of the equation $p(x) = 0$, and that the polynomial factors into the two linear factors,

$$p(x) = x^2 - 3x + 2 = (x-1)(x-2)$$

Sketching the graph of $q(x)$, you will find that it does not intersect the x-axis. The equation $q(x) = 0$ has no solution in the real numbers, and it cannot be factorised (or factored) over the reals. Such a polynomial is said to be *irreducible*. In order to solve this equation, we need to define the complex numbers.

8.5.2 Complex numbers: a formal approach

In a formal approach to complex numbers, we define the set $\mathbb{C}$ of complex numbers to be the set of all ordered pairs (x, y) of real numbers, with addition and multiplication operations defined as follows:

$$(a,b) + (c,d) = (a+c, b+d), \quad (a,b) \times (c,d) = (ac - bd, ad + bc).$$

By identifying the complex number $(x, 0)$ with the real number x, we regard $\mathbb{R}$ as a subset of $\mathbb{C}$. We give the special symbol i to the complex number $(0, 1)$. Note that we can then re-write the complex number (x, y) as $(x, y) = x + yi$.

The complex number i satisfies

$$i^2 = i \times i = (0,1) \times (0,1) = (0 \times 0 - 1 \times 1, 0 \times 1 + 1 \times 0) = (-1, 0),$$

so i^2 is the real number -1.

8.5.3 Complex numbers: a more usual approach

Rather than the ordered pairs approach outlined above, it is more common to define the complex numbers as follows. We begin by defining the *imaginary* number i which has the property that $i^2 = -1$. The term 'imaginary' is historical, and not an indication that this is a figment of someone's imagination. With this, we can then say what we mean by the complex numbers.

Definition 8.2 A complex number is a number of the form $z = a + ib$, where a and b are real numbers, and $i^2 = -1$. The set of all such numbers is

$$\mathbb{C} = \{a + ib \ : \ a, b \in \mathbb{R}\}\ .$$

If $z = a + ib$ is a complex number, then the real number a is known as the real part of z, denoted $\mathsf{Re}(z)$, and the real number b is the imaginary part of z, denoted $\mathsf{Im}(z)$. Note that $\mathsf{Im}(z)$ is a *real* number.

If $b = 0$, then z is a real number, so $\mathbb{R} \subseteq \mathbb{C}$. If $a = 0$, then z is said to be *purely imaginary.*

The quadratic polynomial $q(z) = x^2 + x + 1$ can be factorised over the complex numbers, because the equation $q(z) = 0$ has two complex solutions. Solving in the usual way, we have

$$x = \frac{-1 \pm \sqrt{-3}}{2}.$$

We write, $\sqrt{-3} = \sqrt{(-1)3} = \sqrt{-1}\ \sqrt{3} = i\sqrt{3}$, so that the solutions are

$$w = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \qquad \text{and} \qquad \overline{w} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Notice the form of these two solutions. They are what is called a *conjugate pair*. We have the following definition.

Definition 8.3 If $z = a + ib$ is a complex number, then the *complex conjugate* of z is the complex number $\overline{z} = a - ib$.

We can see by the application of the quadratic formula, that the roots of an irreducible quadratic polynomial with real coefficients will always be a conjugate pair of complex numbers.

Addition, multiplication, division

Addition and *multiplication* of complex numbers are defined as for polynomials in i using $i^2 = -1$.

> **Example 8.4** If $z = (1+i)$ and $w = (4-2i)$ then
>
> $$z + w = (1+i) + (4-2i) = (1+4) + i(1-2) = 5 - i$$
>
> and
>
> $$zw = (1+i)(4-2i) = 4 + 4i - 2i - 2i^2 = 6 + 2i$$

If $z \in \mathbb{C}$, then $z\overline{z}$ is a real number:

$$z\overline{z} = (a+ib)(a-ib) = a^2 + b^2.$$

> **Activity 8.4** Carry out the multiplication to verify this: let $z = a + ib$ and calculate $z\overline{z}$.

Division of complex numbers is then defined by $\dfrac{z}{w} = \dfrac{z\overline{w}}{w\overline{w}}$ since $w\overline{w}$ is real.

> **Example 8.5**
>
> $$\frac{1+i}{4-2i} = \frac{(1+i)(4+2i)}{(4-2i)(4+2i)} = \frac{2+6i}{16+4} = \frac{1}{10} + \frac{3}{10}i$$

Properties of the complex conjugate

A complex number is real if and only if $z = \overline{z}$. Indeed, if $z = a + ib$, then $z = \overline{z}$ if and only if $b = 0$.

The complex conjugate of a complex number satisfies the following properties:

- $z + \overline{z} = 2\,\mathsf{Re}(z)$ is real
- $z - \overline{z} = 2i\,\mathsf{Im}(z)$ is purely imaginary
- $\overline{\overline{z}} = z$
- $\overline{z + w} = \overline{z} + \overline{w}$
- $\overline{zw} = \overline{z}\,\overline{w}$
- $\overline{\left(\dfrac{z}{w}\right)} = \dfrac{\overline{z}}{\overline{w}}$

> **Activity 8.5** Let $z = a + ib,\ w = c + id$ and verify all of the above properties.

8.5.4 Roots of polynomials

The *Fundamental Theorem of Algebra* asserts that a polynomial of degree n with complex coefficients has n complex roots (not necessarily distinct), and can therefore be factorised into n linear factors. Explicitly, any equation

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$$

where $a_i \in \mathbb{C}$ has n solutions $z \in \mathbb{C}$. Contrast this with the difficulty of solving polynomial equations in $\mathbb{R}$. So, the introduction of i enables us to solve **all** polynomial equations: there's no need to introduce anything else. A fancy way of saying this is: ‘The field of complex numbers is algebraically closed.’

If the coefficients of the polynomial are restricted to real numbers, the polynomial can be factorised into a product of linear and irreducible quadratic factors over $\mathbb{R}$ and into a product of *linear* factors over $\mathbb{C}$. The proof of the *Fundamental Theorem of Algebra* is beyond the scope of this course. However, we note the following useful result.

Theorem 8.5 Complex roots of polynomials with real coefficients appear in conjugate pairs.

Proof. Let $P(x) = a_0 + a_1 x + \cdots + a_n x^n,\ a_i \in \mathbb{R}$, be a polynomial of degree n. We shall show that if z is a root of $P(x)$, then so is $\overline{z}$.

Let z be a complex number such that $P(z) = 0$, then

$$a_0 + a_1 z + +a_2 z^2 \cdots + a_n z^n = 0$$

Conjugating both sides of this equation,

$$\overline{a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n} = \overline{0} = 0$$

Since 0 is a real number, it is equal to its complex conjugate. We now use the properties of the complex conjugate: that the complex conjugate of the sum is the sum of the conjugates, and the same is true for the product of complex numbers. We have

$$\overline{a_0} + \overline{a_1 z} + \overline{a_2 z^2} + \cdots + \overline{a_n z^n} = 0,$$

and

$$\overline{a}_0 + \overline{a}_1\overline{z} + \overline{a}_2\overline{z}^2 + \cdots + \overline{a}_n\overline{z}^n = 0.$$

Since the coefficients a_i are real numbers, this becomes

$$a_0 + a_1\overline{z} + a_2\overline{z}^2 + \cdots + a_n\overline{z}^n = 0.$$

That is, $P(\overline{z}) = 0$, so the number $\overline{z}$ is also a root of $P(x)$. □

> **Example 8.6** Let us consider the polynomial
>
> $$x^3 - 2x^2 - 2x - 3 = (x-3)(x^2 + x + 1).$$
>
> If $w = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$, then
>
> $$x^3 - 2x^2 - 2x - 3 = (x-3)(x-w)(x-\overline{w})$$

> **Activity 8.6** Multiply out the last two factors above to check that their product is the irreducible quadratic $x^2 + x + 1$.

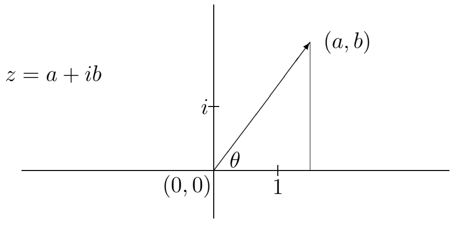
8.5.5 The complex plane

The following theorem shows that a complex number is uniquely determined by its real and imaginary parts.

Theorem 8.6 Two complex numbers are equal if and only if their real and imaginary parts are equal.

Proof. Two complex numbers with the same real parts and the same imaginary parts are clearly the same complex number, so we only need to prove this statement in one direction. Let $z = a + ib$ and $w = c + id$. If $z = w$, we will show that their real and imaginary parts are equal. We have $a + ib = c + id$, therefore $a - c = i(d - b)$. Squaring both sides, we obtain $(a - c)^2 = i^2(d - b)^2 = -(d - b)^2$. But $a - c$ and $(d - b)$ are real numbers, so their squares are non-negative. The only way this equality can hold is for $a - c = d - b = 0$. That is, $a = c$ and $b = d$. □

As a result of this theorem, we can think of the complex numbers geometrically, as points in a plane. For, we can associate the vector $(a, b)^T$ uniquely to each complex number $z = a + ib$, and all the properties of a two-dimensional real vector space apply. A complex number $z = a + ib$ is represented as a point (a, b) in the complex plane; we draw two axes, a horizontal axis to represent the real parts of complex numbers, and a vertical axis to represent the imaginary parts of complex numbers. Points on the horizontal axis represent real numbers, and points on the vertical axis represent purely imaginary numbers.



Complex plane or Argand diagram

> **Activity 8.7** Plot $z = 2 + 2i$ and $w = 1 - i\sqrt{3}$ in the complex plane.

8.5.6 Polar form of z

If the complex number $z = a + ib$ is plotted as a point (a, b) in the complex plane, then we can determine the polar coordinates of this point. We have

$$a = r\cos\theta, \quad b = r\sin\theta$$

where $r = \sqrt{a^2 + b^2}$ is the length of the line joining the origin to the point (a, b) and θ is the angle measured anticlockwise from the real (horizontal) axis to the line joining the origin to the point (a, b). Then we can write $z = a + ib = r\cos\theta + i\,r\sin\theta$.

Definition 8.4 The *polar form* of the complex number z is

$$z = r(\cos\theta + i\sin\theta).$$

The length $r = \sqrt{a^2 + b^2}$ is called the *modulus* of z, denoted $|z|$, and the angle θ is called the *argument* of z.

Note the following properties:

- z and $\overline{z}$ are reflections in the real axis. If θ is the argument of z, then $-\theta$ is the argument of $\overline{z}$.
- $|z|^2 = z\overline{z}$.
- θ and $\theta + 2n\pi$ give the same complex number.

We define the *principal argument* of z to be the argument in the range, $-\pi < \theta \leq \pi$.

> **Activity 8.8** Express $z = 2 + 2i$, $w = 1 - i\sqrt{3}$ in polar form.
>
> Describe the following sets of z: (a) $|z| = 3$, (b) argument of z is $\frac{\pi}{4}$.

Multiplication and division using polar coordinates gives

$$\begin{aligned} zw &= r(\cos\theta + i\sin\theta)\cdot\rho(\cos\phi + i\sin\phi) \\ &= r\rho(\cos(\theta + \phi) + i\sin(\theta + \phi)) \end{aligned}$$

$$\frac{z}{w} = \frac{r}{\rho}\big(\cos(\theta - \phi) + i\sin(\theta - \phi)\big)$$

> **Activity 8.9** Show these by performing the multiplication and the division as defined earlier, and by using the facts that $\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$ and $\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$.

DeMoivre’s Theorem

We can consider explictly a special case of the multiplication result above, in which $w = z$. If we apply the multiplication to $z^2 = zz$, we have

$$\begin{aligned} z^2 &= zz \\ &= (r(\cos\theta + i\sin\theta))(r(\cos\theta + i\sin\theta)) \\ &= r^2(\cos^2\theta + i^2\sin^2\theta + 2i\sin\theta\cos\theta) \\ &= r^2(\cos^2\theta - \sin^2\theta + 2i\sin\theta\cos\theta) \\ &= r^2(\cos 2\theta + i\sin 2\theta). \end{aligned}$$

Here we have used the double angle formulae for $\cos 2\theta$ and $\sin 2\theta$.

Applying the product rule n times, where n is a positive integer, we obtain *DeMoivre's Formula*

Theorem 8.7

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

Proof.

$$\begin{aligned} z^n &= \underbrace{z\cdots z}_{n \text{ times}} = (r(\cos\theta + i\sin\theta))^n \\ &= r^n\Big(\cos(\underbrace{\theta + \cdots + \theta}_{n \text{ times}}) + i\sin(\underbrace{\theta + \cdots + \theta}_{n \text{ times}})\Big) \end{aligned}$$

□

8.5.7 Exponential form of z

Functions of complex numbers can be defined by the power series (Taylor expansions) of the functions:

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \qquad z \in \mathbb{C}$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots \qquad \cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots$$

If we use the expansion for e^z to expand $e^{i\theta}$, and then factor out the real and imaginary parts, we find:

$$\begin{aligned} e^{i\theta} &= 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3} + \frac{\theta^5}{5!} - \cdots\right) \end{aligned}$$

From which we conclude:

Euler's Formula: $\boxed{e^{i\theta} = \cos\theta + i\sin\theta}$

Definition 8.5 The *exponential form* of a complex number $z = a + ib$ is

$$z = re^{i\theta}$$

where $r = |z|$ is the modulus of z and θ is the argument of z.

In particular, the following equality is of note because it combines the numbers e, π and i in a single expression: $e^{i\pi} = -1$.

If $z = re^{i\theta}$, then its complex conjugate is given by $\overline{z} = re^{-i\theta}$. This is because, if $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$, then

$$\overline{z} = r(\cos\theta - i\sin\theta) = r(\cos(-\theta) + i\sin(-\theta)) = re^{-i\theta}.$$

We can use either the exponential form, $z = re^{i\theta}$, or the standard form, $z = a + ib$, according to the application or computation we are doing. For example, addition is simplest in the form $z = a + ib$, but multiplication and division are simpler in exponential form. To change a complex number between $re^{i\theta}$ and $a + ib$, use Euler's formula and the complex plane (polar form).

Example 8.7

$$e^{i\frac{2\pi}{3}} = \cos\tfrac{2\pi}{3} + i\sin\tfrac{2\pi}{3} = -\tfrac{1}{2} + i\tfrac{\sqrt{3}}{2}.$$

$$e^{2+i\sqrt{3}} = e^2e^{i\sqrt{3}} = e^2\cos\sqrt{3} + ie^2\sin\sqrt{3}.$$

Activity 8.10 Write each of the following complex numbers in the form $a + ib$:

$$e^{i\frac{\pi}{2}} \qquad e^{i\frac{3\pi}{2}} \qquad e^{i\frac{3\pi}{4}} \qquad e^{i\frac{11\pi}{3}} \qquad e^{1+i} \qquad e^{-1}$$

Example 8.8 Let $z = 2 + 2i = 2\sqrt{2}\,e^{i\frac{\pi}{4}}$ and $w = 1 - i\sqrt{3} = 2e^{-i\frac{\pi}{3}}$, then

$$w^6 = (1 - i\sqrt{3})^6 = (2e^{-i\frac{\pi}{3}})^6 = 2^6e^{-i2\pi} = 64$$

$$zw = (2\sqrt{2}e^{i\frac{\pi}{4}})(2e^{-i\frac{\pi}{3}}) = 4\sqrt{2}e^{-i\frac{\pi}{12}}$$

and

$$\frac{z}{w} = \sqrt{2}e^{i\frac{7\pi}{12}}.$$

Notice that in the above example we are using certain properties of the complex exponential function, that if $z, w \in \mathbb{C}$,

$$e^{z+w} = e^ze^w \qquad \text{and} \qquad (e^z)^n = e^{nz} \quad \text{for } n \in \mathbb{Z}.$$

This last property is easily generalised to include the negative integers.

Example 8.9 Solve the equation $z^6 = -1$ to find the 6th roots of -1.

Write $z^6 = (re^{i\theta})^6 = r^6e^{i6\theta}$, $\qquad -1 = e^{i\pi} = e^{i(\pi+2n\pi)}$

Equating these two expressions, and using the fact that r is a real positive number, we have

$$r = 1 \qquad 6\theta = \pi + 2n\pi, \quad \theta = \frac{\pi}{6} + \frac{2n\pi}{6}$$

This will give the six complex roots by taking $n = 0, 1, 2, 3, 4, 5$.

> **Activity 8.11** Show this. Write down the six roots of -1 and show that any one raised to the power 6 is equal to -1. Show that $n = 6$ gives the same root as $n = 0$.
>
> Use this to factor the polynomial $x^6 + 1$ into linear factors over the complex numbers and into irreducible quadratics over the real numbers.

8.6 Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- demonstrate that you understand how the rational numbers can be formally constructed by means of an equivalence relation and that addition and multiplication of rational numbers can be defined as operations on the equivalence classes
- indicate that you know that a real number is rational if and only if it has an infinitely repeating pattern in its decimal expansion
- find the decimal expansion of a rational number
- determine, in the form m/n, a rational number from its decimal expansion
- prove that certain numbers are rational or irrational
- demonstrate that you understand that there are rational numbers arbitrarily close to any real number
- state what it means to say that a set is countable or uncountable
- demonstrate that you know that the rationals are countable and the reals uncountable
- show that you know what is meant by complex numbers, and demonstrate that you can add, subtract, multiply and divide complex numbers
- state the definition of the complex conjugate of a complex number
- show that you know that every polynomial of degree n has n complex roots and that these occur in conjugate pairs
- indicate complex numbers on the complex plane
- determine the polar and exponential form of complex numbers
- state and use DeMoivre's theorem and Euler's formula

8.7 Sample exercises

Exercise 8.1
Prove that $\sqrt{5}$ is irrational.

Exercise 8.2
Express the complex number $\dfrac{1+2i}{4-5i}$ in the form $a + bi$.

Exercise 8.3
Solve the equation $x^2 - 2ix + 3 = 0$.

Exercise 8.4
Write each of the following complex numbers in the form $a + ib$:

$$e^{i\frac{\pi}{2}} \qquad e^{i\frac{3\pi}{2}} \qquad e^{i\frac{3\pi}{4}} \qquad e^{i\frac{11\pi}{3}} \qquad e^{1+i} \qquad e^{-1}.$$

Exercise 8.5
Express $1 + \sqrt{3}i$ in exponential form. Hence find $(1 + \sqrt{3}i)^{30}$.

8.8 Comments on selected activities

Learning activity 8.1 Suppose that

$$\frac{a}{b} = \frac{a'}{b'} \text{ and } \frac{c}{d} = \frac{c'}{d'}.$$

What we need to check is that

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

Now, the fact that $\frac{a}{b} = \frac{a'}{b'}$ means precisely that $[(a,b)] = [(a',b')]$, which means that $ab' = a'b$. Similarly, we have $cd' = c'd$. Now,

$$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}, \quad \text{and} \quad \frac{a'}{b'} \times \frac{c'}{d'} = \frac{a'c'}{b'd'}$$

and so we need to prove that

$$\frac{ac}{bd} = \frac{a'c'}{b'd'}.$$

This means we need to prove that

$$[(ac, bd)]\, R\, [(a'c', b'd')].$$

Now,

$$\begin{aligned}
[(ac, bd)]\, R\, [(a'c', b'd')] &\iff acb'd' = a'c'bd \\
&\iff (ab')(cd') = (a'b)(c'd),
\end{aligned}$$

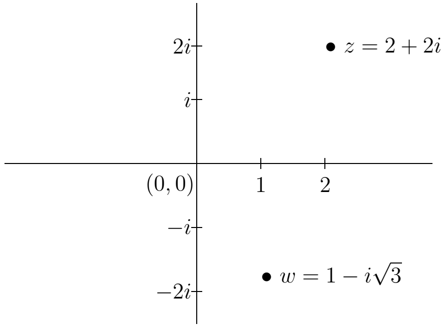
which is true because $ab' = a'b$ and $cd' = c'd$.

Learning activity 8.6 We have

$$(x-w)(x-\overline{w}) = x^2 - (w+\overline{w})x + w\overline{w}.$$

Now, $w+\overline{w} = 2\,\mathsf{Re}(w) = 2(-\frac{1}{2})$ and $w\overline{w} = \frac{1}{4}+\frac{3}{4}$ so the product of the last two factors is x^2+x+1.

Learning activity 8.7



Learning activity 8.8 Draw the line from the origin to the point z in the diagram above. Do the same for w. For z, $|z| = 2\sqrt{2}$ and $\theta = \frac{\pi}{4}$, so $z = 2\sqrt{2}(\cos(\frac{\pi}{4}) + i\sin(\frac{\pi}{4}))$. The modulus of w is $|w| = 2$ and the argument is $-\frac{\pi}{3}$, so that

$$w = 2(\cos(-\frac{\pi}{3}) + i\sin(-\frac{\pi}{3})) = 2(\cos(\frac{\pi}{3}) - i\sin(\frac{\pi}{3})).$$

The set (a) $|z| = 3$, is the circle of radius 3 centered at the origin. The set (b), argument of z is $\frac{\pi}{4}$, is the half line from the origin through the point (1,1).

Learning activity 8.10 The roots are:

$$z_1 = 1\cdot e^{i\frac{\pi}{6}}, \qquad z_2 = 1\cdot e^{i\frac{3\pi}{6}}, \qquad z_3 = 1\cdot e^{i\frac{5\pi}{6}},$$
$$z_4 = 1\cdot e^{i\frac{7\pi}{6}}, \qquad z_5 = 1\cdot e^{i\frac{9\pi}{6}}, \qquad z_6 = 1\cdot e^{i\frac{11\pi}{6}}.$$

These roots are in conjugate pairs, and $e^{i\frac{13\pi}{6}} = e^{i\frac{\pi}{6}}$:

$$z_4 = \overline{z}_3 = e^{-i\frac{5\pi}{6}}, \qquad z_5 = \overline{z}_2 = e^{-i\frac{\pi}{2}}, \qquad z_6 = \overline{z}_1 = e^{-i\frac{\pi}{6}}.$$

The polynomial factors as

$$x^6+1 = (x-z_1)(x-\overline{z}_1)(x-z_2)(x-\overline{z}_2)(x-z_3)(x-\overline{z}_3),$$

Using the $a+ib$ form of each complex number, for example, $z_1 = \frac{\sqrt{3}}{2} + i\frac{1}{2}$, you can carry out the multiplication of the linear terms pairwise (conjugate pairs) to obtain x^6+1 as a product of irreducible quadratics with real coefficients,

$$x^6+1 = (x^2 - \sqrt{3}\,x + 1)(x^2 + \sqrt{3}\,x + 1)(x^2+1)$$

8.9 Solutions to exercises

Solution to exercise 8.1

Suppose we have $\sqrt{5} = m/n$ where $m, n \in \mathbb{Z}$. Since $\sqrt{5} > 0$, we may assume that $m, n > 0$. (Otherwise, both are negative, and we can multiply each by -1.) We can also suppose that m, n have greatest common divisor 1. (For, we can cancel any common factors.) Then $(m/n)^2 = 5$ means that $m^2 = 5n^2$. So $5\,|\,m^2$. Now m can, by the Fundamental Theorem of Arithmetic, be written as a product of primes $m = p_1p_2\ldots p_k$. Then $m^2 = p_1^2p_2^2\ldots p_k^2$. If no p_i is 5, then 5 does not appear as a factor in m^2 and so 5 does not divide m^2. So some p_i is equal to 5. So $5\,|\,m$. Now, this means that $m = 5r$ for some $r \in \mathbb{N}$ and hence $m^2 = (5r)^2 = 25r^2$ and so $25r^2 = 5n^2$. Then, $n^2 = 5r^2$, so $5\,|\,n^2$. Arguing as before, $5\,|\,n$. So 5 is a common factor if m and n, which contradicts $\gcd(m,n) = 1$. Hence $\sqrt{5}$ is not rational.

Solution to exercise 8.2

We have

$$\begin{aligned}\frac{1+2i}{4-5i} &= \frac{1+2i}{4-5i}\frac{4+5i}{4+5i}\\ &= \frac{(1+2i)(4+5i)}{(4-5i)(4+5i)}\\ &= \frac{4+8i+5i+10i^2}{16-25i^2}\\ &= \frac{-6+13i}{41}\\ &= -\frac{6}{41}+\frac{13}{41}i.\end{aligned}$$

You can *check* that this is the correct answer by calculating the product

$$\left(-\frac{6}{41}+\frac{13}{41}i\right)(4-5i)$$

and observing that the answer is $1+2i$. □

Solution to exercise 8.3

To solve the equation $x^2 - 2ix + 3 = 0$, we could use the formula for the solutions of a quadratic equation. Or we could note that the equation is equivalent to $(x-i)^2 = -4$, so the solutions are given by $x - i = 2i$ and $x - i = -2i$, so they are $x = 3i$ and $x = -i$.

Solution to exercise 8.4

We have

$$e^{i\pi/2} = i, \qquad e^{i3\pi/2} = -i, \qquad e^{i3\pi/4} = -\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}},$$

$$e^{i(11\pi/3)} = e^{-i(\pi/3)} = \frac{1}{2} - i\frac{\sqrt{3}}{2}, \qquad e^{1+i} = e^1e^i = e\cos(1) + i\,e\sin(1),$$

$$e^{-1} = e^{-1} + 0i \text{ is real, so already in the form } a+ib.$$

Solution to exercise 8.5

To express $z = 1 + \sqrt{3}i$ in exponential form, we first note that

$$1 + \sqrt{3}i = 2\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)$$

and this is $r(\cos\theta + i\sin\theta)$ when $r = 2, \theta = \pi/3$. So $z = 2e^{\pi i/3}$. Then,

$$(1+\sqrt{3}i)^{30} = z^{30} = \left(2e^{\pi i/3}\right)^{30} = 2^{30}e^{30\pi i/3} = 2^{30}e^{10\pi i} = 2^{30}.$$

Introduction to Abstract Mathematics
MA 103

Exercises 1

- Relevant parts of the **Lecture notes**: Chapter 1 and Sections 2.1 – 2.8, 2.12 and 2.13.
- Relevant parts of the text books: **Biggs**: Chapter 1 and Sections 3.1 – 3.5;
  Eccles: Chapters 1 – 4.

- Try to do as many questions as you can, and hand in whatever you have.
- Hand in your homework by putting it in the **homework box of your class teacher** on the Ground Floor of Columbia House.
- Always write your name, the course number (MA 103), your class group number and your class teacher's name on your homework.
- Use standard A4 paper and leave room in the margins for remarks, corrections, etc. from your class teacher. Use paperclips or staples to keep the pages of your homework together.

Here are some general remarks about writing mathematics:

- **Always** justify your answers, whatever the wording of the question.
- Most answers to exercises in this course will require a careful explanation of why something is true or false. You should give your argument in enough detail for somebody else to follow what you write.
- Write your answers in **English**. That is, don't just use symbols, but use **words** to explain how you get from each line to the next.
- Avoid using the symbols "$\therefore$" or "$\because$" (if you don't know what they mean, that's just fine), or arrows like "$\longrightarrow$" or "$\Longrightarrow$" in mathematical arguments. Use words!

In the questions below, we assume that we are always dealing with natural numbers. So if a statement is made about certain numbers n, say, then the fuller version of this statement would start: *for all natural numbers n, ...*

1 Consider the following (false) statement:

If n is a multiple of 16, then n is not a multiple of 6.

What properties must the number n have, for n to be a *counterexample* to this statement?

Find a counterexample to the given statement.

2 Show that the following statement is true, by giving a proof:

If n is a multiple of 4, then $9n + 30$ is a multiple of 6.

3 (a) Construct the truth table for the statement $p \Rightarrow (q \wedge r)$.

(b) Decide whether the statement in (a) is logically equivalent to $(p \Rightarrow q) \wedge (p \Rightarrow r)$ or not.

4 (a) Construct the truth table for the statement $(p \Rightarrow q) \Rightarrow r$.

(b) Construct the truth table for the statement $p \Rightarrow (q \Rightarrow r)$.

(c) Construct the truth table for the statement $(p \Rightarrow q) \wedge (q \Rightarrow r)$.

(d) In view of your answers, why do you think we should never write $p \Rightarrow q \Rightarrow r$?

5 (a) Write down the **converse** of the statements in Questions 1 and 2.

(b) For each of the converses, determine whether the converse statement is true or false. Justify your answer by giving a proof or a counterexample.

6 (a) Write down the **contrapositive** of the statements in Questions 1 and 2.

(b) For each of the contrapositives, determine whether the statement is true or false.

7 For which natural numbers n is $3^n - 1$ a prime number? Justify your answer.

8 (a) Explain what is wrong with the following proof of the statement:

If n and m are natural numbers, then $4 = 3$.

Proof: Denote the sum $n + m$ by t. Then we certainly have $n + m = t$.
This last statement can be rewritten as $(4n - 3n) + (4m - 3m) = 4t - 3t$.
Rearrangement leads to $4n + 4m - 4t = 3n + 3m - 3t$.
Taking out common factors on each side gives $4(n + m - t) = 3(n + m - t)$.
Removing the common term now leads to $4 = 3$. □

(b) Explain what is wrong with the following proof of the statement:

If n and m are natural numbers, then $n = m$.

Proof: Denote the sum $n + m$ by t. Then we certainly have $n + m = t$.
Multiplying both sides by $n - m$ yields $(n + m)(n - m) = t(n - m)$.
Multiplying out gives $n^2 - m^2 = tn - tm$.
This can be rewritten as $n^2 - tn = m^2 - tm$.
Adding $\frac{1}{4}t^2$ to both sides gives $n^2 - tn + \frac{1}{4}t^2 = m^2 - tm + \frac{1}{4}, t^2$.
This is the same as $(n - \frac{1}{2}t)^2 = (m - \frac{1}{2}t)^2$.
So we get $n - \frac{1}{2}t = m - \frac{1}{2}t$.
Adding $\frac{1}{2}t$ to both sides gives $n = m$. □

Introduction to Abstract Mathematics
MA 103

Exercises 2

- Relevant parts of the **Lecture notes**: Sections 2.9 – 2.17.
- Relevant parts of the text books: **Biggs**: Chapter 2 and Sections 3.6 – 3.7;
  Eccles: Sections 6.1 – 6.2, 7.1 – 7.4 and 7.6.
- Always justify your answers.
- In the questions below you must take care to note whether the question deals with 'natural numbers' or with 'integers'.

1 (a) Write down the mathematical notation for the following sets:

(i) the set A of all natural numbers that are multiples of five;

(ii) the set B of all integers whose fourth power is less than 1000.

(b) List the elements of the set B in (a)(ii).

2 For each of the following pairs of sets A, B, determine whether the statement "$A \subseteq B$" is true or false. Make sure you explain why!

(a) $A = \{-1, -3, -9\}, \quad B = \{\, m \mid m \text{ is an odd integer} \,\}$;

(b) $A = \{0, 1, 3, 7, 1023\}, \quad B = \{\, m \mid m = 2^n - 1 \text{ for some natural number } n \,\}$;

(c) $A = \varnothing, \quad B = \{0\}$;

(d) $A = \{n \in \mathbb{N} \mid n \text{ is even}\}, \quad B = \{n \in \mathbb{N} \mid n^2 \text{ is even}\}$;

(e) $A = \{1\}, \quad B = \{\, \mathbb{N} \,\}$.

3 Consider the set $X = \{\, A \mid A \subseteq \{\, 0, 1 \,\} \,\}$. In words: X is the set consisting of those A which are subsets of $\{0, 1\}$.

(a) Exactly what does it mean to say that A is an *element* of X?

Exactly what does it mean to say that B is a *subset* of X?

(*If you are not sure, go back and look at the definitions of these terms.*)

(b) Is $\{0\}$ an *element* of X?

(c) Is $\{0\}$ a *subset* of X?

(d) Is there a set which is both an element of X and a subset of X?

Justify all your answers carefully.

4 (a) Write out the following statements using mathematical notation for the sets and quantifiers. (Ordinary phrases like 'such that' are acceptable.)

For every integer x, there is an integer z such that $x \leq z \leq 2x$.

There is a natural number n such that, for every natural number m, we have $n \leq m$.

(b) Prove that the first statement in (a) is false by giving a counterexample.

(c) Is the second statement in (a) true or false? (Justify your answer in detail.)

(d) Formulate the negation for each of the two statements, and express these as simply as you can.

Write these negations out in mathematical notation as well.

5 Decide whether the following statement is True or False, and justify your assertion by means of a proof or a counterexample:

For all sets A, B and C, we have $A \cup (B \cap C) = (A \cup B) \cap C$.

6 (a) Give a proof that, for all sets A, B, C, we have

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C).$$

(b) Give a proof that, for all sets A, B, C, we have

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

7 *This is an extra question, and it is not directly related to the material discussed this week. Try to see how far you can get; it's a good way to test your mathematical skills.*

The numbers 1 to 25 are arranged in a square array of five rows and five columns in an arbitrary way. The greatest number in each row is determined, and then the least number of these five is taken; call that number s. Next, the least number in each column is determined, and then the greatest number of these five is taken; this number is called t.

(a) Arrange the numbers 1 to 25 in a square array by writing $1, 2, 3, 4, 5$ in the first row, then $6, 7, 8, 9, 10$ in the second row, etc. Determine s and t for this arrangement. (You should have $s = t$.)

(b) Find a way to arrange the numbers 1 to 25 in a square array which leads to values s, t with $s \neq t$.

(c) Write out carefully a proof of the following statement:

For every possible arrangement in a square array of the numbers 1 to 25,
if we obtain s and t as described above, then we have $s \geq t$.

Introduction to Abstract Mathematics
MA 103

Exercises 3

- Relevant parts of the **Lecture notes**: Sections 3.1 – 3.8 and 3.10 – 3.12.
- Relevant parts of the text books: **Biggs**: Chapter 5;
  Eccles: Chapter 5.
- Always justify your answers.

1 (a) Use the principle of induction to prove that, for all $n \in \mathbb{N}$, $n^3 + 5n$ is a multiple of 3.

(b) Use the principle of induction to prove that, for all $n \in \mathbb{N}$, $n^3 + 5n$ is a multiple of 6.

(*You may use the result that, for every $k \in \mathbb{N}$, $k^2 + k$ is an even number.*)

2 Prove that $\sum_{k=1}^{n} (4k-3) = n(2n-1)$, for all $n \geq 1$.

3 (a) Prove that $2m \leq 2^m$, for all natural numbers m.

Now, for m a natural number, let $P(m)$ be the statement: $m(m-1) + 10 \leq 2^m$.

(b) Use (a) to show that, for any natural number m, if $P(m)$ is True, then $P(m+1)$ is True.

(c) For which natural numbers m is $P(m)$ True? Justify your answer carefully.

4 The natural numbers u_n are defined recursively as follows.

$$u_0 = 2; \quad u_n = 2u_{n-1} + 1 \quad (n \geq 1).$$

(a) Write down the first few values of u_n.

(b) On the basis of these values, guess a formula for u_n of the form $u_n = a2^n + b$, where you need to guess the value of the integers a and b.

(c) Use induction to show that your formula is correct.

5 Write down explicit formulae for the expressions u_n, v_n and w_n defined as follows:

$$\begin{aligned} u_1 &= 1; \quad u_n = u_{n-1} + 3 \quad (n \geq 2); \\ v_1 &= 1; \quad v_n = n^2 v_{n-1} \quad (n \geq 2); \\ w_0 &= 0; \quad w_n = w_{n-1} + 3n \quad (n \geq 1). \end{aligned}$$

(*Start by working out the first few values in each case – without doing the arithmetic, so you might write, for instance, $u_2 = 1 + 3$, $u_3 = (1+3) + 3, \ldots$ – and try to spot the pattern emerging. It is a good idea to check that the formula you give is correct for the first one or two values of n.*)

6 (a) For each of the sets X, Y, Z below, find the least element in that set.

$$\begin{aligned} X &= \{\, x \in \mathbb{N} \mid x^3 \leq 500 \,\}; \\ Y &= \{\, x \in \mathbb{N} \mid x = k^2 - 100 \text{ for some } k \in \mathbb{N} \,\}; \\ Z &= \{\, x \in \mathbb{N} \mid x^2 + 100 < 25x \,\}. \end{aligned}$$

(b) And, if it exists, find the greatest element in X, Y and Z. (Make sure you justify your answers.)

Note: make sure you read the definitions of the sets carefully. What exactly does it mean, for instance, to say that 11 is an element of Y?

7 *This is an extra question, and it is not directly related to the material discussed this week. Try to see how far you can get; it's a good way to test your mathematical skills.*

Eddie and Ayesha, a married couple, go to a party in which there are four other couples. Some people from the group know each other, others have never met.

People who meet for the first time are introduced and shake hands. No one shakes hands with her/his own partner, and no one shakes hands more than once with the same person.

At the end of the party, Eddie asks the nine others how many people they had shaken hands with. To his surprise, all answers are different.

(a) How many people shook hands with Eddie's wife, Ayesha?

(b) If you find part (a) hard, you can try the following easier variants first:

(i) What would the answer be if there was only **one** other couple?
(Use the same assumptions on whom can shake hands with whom. But now assume at the end Eddie asks the three others how many people they shook hands with, and he gets three different answers. And the question is to find out how many people shook hands with Ayesha.)

(ii) And what is the answer if there are **two** other couples?
(Again, use the same assumptions: Eddie asks five people and gets five different answers.)

(iii) Finally, try to answer the original question.

Introduction to Abstract Mathematics
MA 103

Exercises 4

- The material for this week is not included in the Lecture notes. You will be give an additional set of notes.
- Relevant parts of the text books: **Eccles**: Section 5.4.
- Always justify your answers.

1 (a) Consider female cats that reproduce by the following rule: A cat born at time n has two daughters both born at time $n+1$, and otherwise no further female offspring.

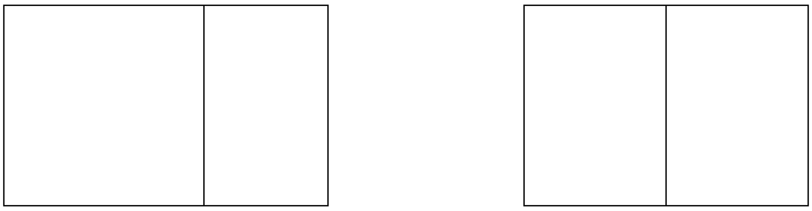
Starting with a single female cat born at time 1, what is the number of female cats born at time n, for any natural number n? Justify your answer.

(b) Suppose rabbit pairs reproduce by the following rule: A rabbit pair born at time n gives birth to one further pair in every month starting from month $n+1$. Rabbits never die.

Starting with a single rabbit pair at time 1, what is the total number of rabbit pairs at time n, for any natural number n? Justify your answer.

2 If a rectangle has sides of lengths A and B where $A \geq B$, then we call A/B the *proportion* of the rectangle (so a square has proportion 1, and any non-square rectangle has proportion larger than 1).

The Golden Ratio ϕ is defined as the proportion of a rectangle with the property that, after removing a square, the remaining smaller rectangle has the same proportion ϕ, as shown in the picture on the left:



Consider now the different proportion α of a rectangle that has the property that when dividing the longer side in half, the resulting two rectangles each have that same proportion α, as shown in the above picture on the right.

(a) Determine α.

The standard paper formats A0, A1, A2, A3, A4 etc. are defined by the rule that they have this proportion α and each is the previous format cut in half. So A4 is obtained by cutting A3 in half, for example. In addition, A0 is exactly one square metre in area.

(b) Determine the sizes (rounded to a millimetre) of A0, A1, A2, A3, A4 (the familiar A4 format allows you to check your result for plausibility).

(c) What is the weight of an A4 sheet of paper that weighs 80 grams per square metre?

3 Analogous to the method for finding Binet's explicit formula for the Fibonacci numbers, show how to derive explicit formulas for the following recursively defined sequences (which, apart from (c), have been given, and proved to hold by induction, in an earlier lecture):

(a) $a_1 = 9,\ a_2 = 13,$ and $a_n = 3a_{n-1} - 2a_{n-2}$ for $n \geq 3$;

(b) $b_1 = 7,\ b_2 = 23,$ and $b_n = 5b_{n-1} - 6b_{n-2}$ for $n \geq 3$;

(c) $c_1 = 1,\ c_2 = 2,$ and $c_n = 2c_{n-1} + 3c_{n-2}$ for $n \geq 3$.

4 Recall that a *unit fraction* is of the form $1/a$ for a natural number a. An *Egyptian fraction* is a sum of distinct unit fractions.

(a) Show, using Egyptian fractions, how to divide 3 loaves of bread among 11 people so that everyone gets the same kinds of pieces.

The "greedy method" to present a fraction p/q, where $1 \leq p < q$, as an Egyptian fraction finds the smallest integer a so that $p/q \geq 1/a$, which is $a = \lceil q/p \rceil$, and applies this recursively to the difference $p/q - 1/a$ until that difference is zero.

(b) Represent $7/15$ as an Egyptian fraction using the greedy method.

(c) Represent $3/7$ as an Egyptian fraction using the greedy method.

How does this help you to distribute 3 loaves of bread among 7 people? If this does not work, how would you modify the method? (Think of how to divide the bread.)

Introduction to Abstract Mathematics
MA 103

Exercises 5

- Relevant parts of the **Lecture notes**: Sections 4.1 – 4.5 and 4.8.
- Relevant parts of the text books: **Biggs**: Chapter 5 and Sections 6.1 – 6.3;
  Eccles: Sections 8.1 – 8.2, 8.4 – 8.5, 9.1 – 9.2 and 10.1.

- Always justify your answers.
- Before you start these exercises, make sure you have in front of you the *precise* definitions you need. In particular, you shoud know exactly what it means for a set to have m elements.

1 The functions α and β from $\mathbb{N}$ to $\mathbb{N}$ are defined by

$$\alpha(n) = 2^n, \qquad \beta(n) = n^2 \qquad (n \in \mathbb{N}).$$

(a) Find formulae for the compositions $\alpha\beta$ and $\beta\alpha$.

(b) Determine whether $\alpha\beta$ and $\beta\alpha$ are different functions or not.

2 Which of the following functions from $\mathbb{Z}$ to $\mathbb{Z}$ are injections, which are surjections, and which are bijections?

(a) $f(x) = x^3 \ (x \in \mathbb{Z})$;

(b) $g(x) = 3 - x \ (x \in \mathbb{Z})$;

(c) $h(x) = 2x + 3 \ (x \in \mathbb{Z})$;

(d) $j(x) = \begin{cases} x+1, & \text{if } x \text{ is even,} \\ x+3, & \text{if } x \text{ is odd,} \end{cases} \quad (x \in \mathbb{Z})$.

For each of the functions that are bijections, give the inverse function.

3 The function z from $\mathbb{N}$ to $\mathbb{N}$ is defined recursively by the rules

$$z(1) = 1 \qquad \text{and} \qquad z(n+1) = \begin{cases} \frac{1}{2}(z(n)+3), & \text{if } z(n) \text{ is odd,} \\ z(n)+5 & \text{if } z(n) \text{ is even,} \end{cases} \qquad \text{for all } n \geq 1.$$

Show that z is neither an injection nor a surjection.

4 Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c, d\}$.

(a) How many functions are there from X to Y?

(b) How many bijections are there from X to Y?

(c) How many injections are there from X to Y?

Explain your answers.

5 In each of the following cases, find a value for m such that a bijection $f : \mathbb{N}_m \longrightarrow X$ exists, and give a formula for such a bijection:

(a) $X = \{ 12, 15, 18, 21, 24, 27, 30 \}$;

(b) $X = \{ x \in \mathbb{Z} \mid -2 \leq x \leq 4 \}$;

(c) $X = \{ x \in \mathbb{Z} \mid x^2 \leq 16 \}$.

6 Let A be a finite set with m elements, for some $m \in \mathbb{N}$. And suppose x is an object that is not a member of A.

Prove, using the definitions, that $A \cup \{x\}$ has $m + 1$ elements.

(All you are told about A is that it has m elements. You need to show that there is a bijection from $\mathbb{N}_{m+1}$ to $A \cup \{x\}$.)

7 Explain what is wrong with the following proof of the statement:

If at least one person in the world has blue eyes, then everybody has blue eyes.

Proof: Instead of the statement above, we prove the following:

Theorem: *For all $n \in \mathbb{N}$ we have: if in a group of n persons there is at least one person with blue eyes, then everybody in that group has blue eyes.*

Once the theorem is proved, the original statement follows if we take for n the number of people in the world.

Proof of Theorem: We prove the theorem using induction.

First, the result is true for $n = 1$. Since if we have a group of one person, and that person has blue eyes, then everybody in the group has blue eyes.

Now suppose the result is true for $n = k$, hence we suppose that: *if in a group of k persons at least one person has blue eyes, then everybody in that group has blue eyes.*
We are going to prove that this gives the statement for $n = k + 1$: *if in a group of $k + 1$ persons there is at least one person with blue eyes, then everybody in that group has blue eyes.*
So take a group of $k + 1$ persons in which at least one person has blue eyes. Let's use the name A for one of the persons with blue eyes. Since $k + 1 \geq 2$, there is at least one person in the group who is not A. Call that person B. Now remove B from the group. What is left is a group of k people in which at least one has blue eyes (since A is in that group). Hence everybody in that group has blue eyes (induction hypothesis).
Now put B back in, but remove another person. Again we have a group of k people with at least one person with blue eyes. (In fact, all people except B are known to have blue eyes.) So also for this group we must conclude that everybody has blue eyes. In particular B must have blue eyes as well.
So we see that everybody in the original group of $k + 1$ persons has blue eyes.

By the principle of induction it follows that the result is true for all $n \in \mathbb{N}$. □

Introduction to Abstract Mathematics
MA 103

Exercises 6

- Relevant parts of the **Lecture notes**: Sections 4.6 – 4.12. and 5.1 – 5.4.
- Relevant parts of the text books: **Biggs**: Section 6.3 – 6.4, 7.2 – 7.3 and 7.6;
  Eccles: Section 11.1 and Chapter 22.

  (The treatment of equivalence relations in **Eccles** is somewhat different of the way we do it, so read Chapter 22 with care.)
- Always justify your answers.

  (Hint: in the first three exercises, try using the Pigeonhole Principle.)

1 Vladimir has 5 pairs of blue socks, 10 pairs of red socks, and 3 pairs of grey socks, all loose in a drawer. He picks a number of socks in the dark (i.e. he can't see the colours of the socks).

What is the minimum number of socks he has to pick to be guaranteed to have two of the same colour?

2 Let T be a set of 11 different natural numbers. Show that there are two elements $t_1, t_2 \in T$, $t_1 \neq t_2$, such that $t_2 - t_1$ is divisible by 10.

3 In Question 7 of Exercises 3 we met Eddie and Ayesha. A couple of weeks after that question they attend another party. Again, some pairs of people shake hands, and some pairs don't. The total number of people at the party is n, for some $n \geq 2$.

Explain why, no matter what handshakes take place, there are certain to be two people at the party who shake the same number of hands.

4 For each of the following relations on the set $\mathbb{N}$, decide whether the relation is reflexive, whether it is symmetric, whether it is transitive, and whether it is an equivalence relation. Justify your answers, giving short proofs or explicit counterexamples as appropriate.

(a) the relation xRy given by "$x \geq y + 2$";

(b) the relation xSy given by "$x \times y$ is even";

(c) the relation xTy given by "x is a multiple of y";

(d) the relation xUy given by "$x = 1$ and $y = 1$".

5 Show that the relation V on $\mathbb{Z}$ defined by setting xVy if $|x^2 - y^2| \leq 2$ is an equivalence relation.

Describe the equivalence classes of this relation.

6 Explain what is wrong with the following proof of the statement:

> *If S is a relation on a set X that is both symmetric and transitive,*
> *then S is an equivalence relation.*

Proof: Suppose S is a symmetric and transitive relation on a set X, and let a be any element of X. Now, if aSb, then bSa (since S is symmetric), and so aSa (since S is transitive). Therefore S is reflexive, as well as being symmetric and transitive. So S is an equivalence relation. □

Hint: consider the relation U in Question 4(d).
Metahint: why might we suggest you do that?

7 For a subset S of $\mathbb{Z}$, we say that m is a *lower bound of* S if for all $s \in S$ we have $m \leq S$. And M is an *upper bound of* S if for all $s \in S$ we have $M \geq S$.

For each of the following (false) statements, give a counterexample (i.e. a set S of integers with the first property, but not the second).

(a) If S is a set of integers with an upper bound, then S has a lower bound.

(b) If S is a set of integers with an upper bound, then $\overline{S} = \{x \in \mathbb{Z} \mid x \notin S\}$ has a lower bound.

(c) If S is a set of integers with no upper bound, then $\overline{S}$ has an upper bound.

Introduction to Abstract Mathematics
MA 103

Exercises 7

- Relevant parts of the **Lecture notes**: Chapter 6.
- Relevant parts of the text books: **Biggs**: Chapter 8;
  Eccles: Chapters 15 – 17 and Section 23.1.

- Always justify your answers.

1 (a) Write down the definition of the statement "$a|b$" (a divides b) for two integers a, b.

Use this definition to answer the following questions.

(b) For which integers z is it true that $0|z$?

(c) For which integers z is it true that $z|0$?

(d) For which integers z is it true that $1|z$?

(e) For which integers z is it true that $z|1$?

2 Which of the following statements, for integers a, b, c, d, p, q, are true in general? Justify your answers, referring again to the definition of divisibility.

(a) If $d|a$ and $d|b$, then $d|(pa + qb)$.

(b) If $a|c$ and $b|c$, then $(ab)|c$.

(c) If $a|c$ and $b|d$, then $(a + b)|(c + d)$.

(d) If $a|c$ and $b|d$, then $(ab)|(cd)$.

3 (a) Use Euclid's algorithm to find the greatest common divisor of 2009 and 820.

(b) Find integers x, y so that $\gcd(2009, 820) = 2009x + 820y$.

4 (a) Show that $\gcd(96, 42) = 6$, and find integers x and y satisfying $96x + 42y = 6$. Hence find integers p and q satisfying $96p + 42q = 60$.

(b) Let a and b be positive integers, and let $d = \gcd(a, b)$. Prove that, for $c \in \mathbb{Z}$, we have $d|c$ if and only if there are integers x, y so that $c = xa + yb$.

5 Let a, b be integers, $b > 0$. In the lecture notes, it is proved that

Theorem: *if $a > 0$, then there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$ with $0 \leq r < b$.*

Show how you can use this statement to deduce that

if $a < 0$, then there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$ with $0 \leq r < b$.

6 Let $n \geq 2$ be a number that is not a prime.

(a) Show that there is a divisor q of n, with $q \geq 2$, so that $q^2 \leq n$.

(b) Show that there is a divisor p of n so that $p^2 \leq n$, $p \geq 2$ and p is prime.

7 Let $F_0, F_1, F_2, F_3, \ldots$ be the *Fibonacci numbers*, defined by

$$F_0 = 1; \qquad F_1 = 1; \qquad F_n = F_{n-1} + F_{n-2}, \quad \text{for } n \geq 2.$$

(a) Prove that for all $n \geq 0$ we have $\gcd(F_{n+1}, F_n) = 1$.

(b) Prove that for all $n \geq 0$ we also have $\gcd(F_{n+2}, F_n) = 1$.

(Hint: check what Euclid's algorithm would do if you started to compute $\gcd(F_{n+1}, F_n)$ or $\gcd(F_{n+2}, F_n)$.)

(c) Is it true that, for all $n \geq 0$, we have $\gcd(F_{n+3}, F_n) = 1$?

(d) Show that, for $n \geq 3$, $F_{n+3} = 4F_n + F_{n-3}$. What does this tell us about $\gcd(F_{n+3}, F_n)$?

Maple questions

Since all students following this course are also supposed to follow (or have followed) MA100 *Mathematical Methods*, everybody should have some experience with using the computer package "Maple". This program is available on all PCs in the school.

Future exercises sheets will contain some questions in which you are asked to use Maple to explore aspects of the theory. In your answers, you should say what commands you used and what the output is. You may attach a printout of your work as well if you wish.

8 For all $k \in \mathbb{N}$, let p_k be the k-th prime, so $p_1 = 2$, $p_2 = 3$, etc.

(a) The maple command to find the k-th prime is "`ithprime(k)`".

Use this command to find p_{15}, p_{150} and $p_{1,500}$.

(b) The maple command to check if a number n is prime is "`isprime(n)`".

Use this command to find all prime numbers between 1,000,000 and 1,000,010.

We prove in lectures that the set of primes is infinite, using a proof by contradiction. This proof looks at the numbers

$$N_m = p_1 \cdot p_2 \cdot \cdots \cdot p_m + 1, \qquad \text{for } m \geq 1.$$

So $N_1 = p_1 + 1 = 2 + 1 = 3$, $N_2 = p_1 p_2 + 1 = 2 \cdot 3 + 1 = 7$, etc.

We prove that N_m is not divisible by any of $p_1, p_2, \ldots, p_m$, so that N_m is either a prime or it is divisible by a prime larger than p_m.

(c) Use Maple to find out which of these numbers N_m, for $m = 1, 2, \ldots, 15$, is actually prime.

The Maple command to find the prime factors of n is "`numtheory[factorset](n)`". (It is easier to first load the "number theory" package by typing "`with(numtheory)`", and then use the command "`factorset(n)`".)

(d) Use Maple to compare p_m with the smallest prime number that divides N_m, for $m = 1, 2, \ldots, 15$.

Introduction to Abstract Mathematics
MA 103

Exercises 8

- Relevant parts of the **Lecture notes**: Chapter 7.
- Relevant parts of the text books: **Biggs**: Sections 13.1 – 13.3;
  Eccles: Chapters 19 – 21.
- Always justify your answers.

1 (a) Find the representations of $(2008)_{10}$ in base 2, base 3 and base 11.

(b) What is the normal decimal notation for $(100011)_2$ and $(100011)_3$?

(c) For any integer $m \geq 2$, what is 1 in base m? And what is m in base m?

2 Let m be an integer, $m \geq 2$.

(a) Prove that, for all elements $r, s, t \in \mathbb{Z}_m$, we have $(r \oplus s) \oplus t = r \oplus (s \oplus t)$.
(Remember that an element of $\mathbb{Z}_m$ is an equivalence class $[x]_m$ for some $x \in \mathbb{Z}$.)

(b) Prove that for all elements $r, s \in \mathbb{Z}_m$ we have $r \otimes s = s \otimes r$.

(c) Show that in $\mathbb{Z}_m$ we have $(m-1)^{-1} = m - 1$.

3 Solve the following systems of equations; i.e., in each case, find **all** pairs (x, y) satisfying both equations. Make sure to justify each step carefully.

(a) $\left\{ \begin{array}{r} 2x + y = 1, \\ x + 2y = 3, \end{array} \right.$ in $\mathbb{Z}_7$;

(b) $\left\{ \begin{array}{r} 2x + y = 1, \\ x + 2y = 3, \end{array} \right.$ in $\mathbb{Z}_6$;

(c) $\left\{ \begin{array}{r} 2x + y = 2, \\ x + 2y = 4, \end{array} \right.$ in $\mathbb{Z}_6$.

4 Find the inverses (if they exist) of

(a) 3 in $\mathbb{Z}_{11}$; (b) 11 in $\mathbb{Z}_{16}$; (c) 14 in $\mathbb{Z}_{16}$. (d) 100 in $\mathbb{Z}_{2007}$.

5 For part (a) and (b) of this question, we work in $\mathbb{Z}_p$, where p is a prime. (*So what result from lectures is likely to be useful?*)

(a) Suppose $a, b \in \mathbb{Z}_p$ have the property that $ab = 0$ in $\mathbb{Z}_p$. Show that $a = 0$ or $b = 0$ in $\mathbb{Z}_p$.

(b) Using the result in (a), show that the equation $x^2 = 1$ has only the solutions $x = 1$ and $x = -1$ in $\mathbb{Z}_p$ (which is the same as $x = 1$ and $x = p - 1$ in $\mathbb{Z}_p$).

(c) How many solutions are there for the equation $x^2 = 1$ in $\mathbb{Z}_2$?

(d) Find an integer $m \geq 2$ so that the equation $x^2 = 1$ in $\mathbb{Z}_m$ has more than two solutions.

6 We want to solve the quadratic equation $y^2 + y - 11 = 0$ in $\mathbb{Z}_{41}$.

(a) Solve the equation $2a = 1$ in $\mathbb{Z}_{41}$.

(b) Find elements $a, b \in \mathbb{Z}_{41}$ such that $y^2 + y - 11 = (y + a)^2 - b$ in $\mathbb{Z}_{41}$.

(c) Use (b) and Question 5(b) to write down the solution(s) of $y^2 + y - 11 = 0$ in $\mathbb{Z}_{41}$.

Maple questions

7 (a) Find all primes p with $41 \leq p \leq 107$. (Remember that the Maple command to find the i-th prime number is "`ithprime(i)`".)

The Maple command to solve equations in modular arithmetic $\mathbb{Z}_m$ is "`msolve`". Use the Maple help (or type "`?msolve`") to find out how to use this command for different values of m.

(b) Solve the equation $y^2 + y - 11 = 0$ in $\mathbb{Z}_p$ for all primes p with $41 \leq p \leq 107$.

You will notice that there are no solutions (or at least, Maple doesn't give solutions) for certain values of p.

(c) Use your results (and if necessary check for more prime values) to make a conjecture describing the primes p for which there are solutions to $y^2 + y - 11 = 0$ in $\mathbb{Z}_p$.

Introduction to Abstract Mathematics
MA 103

Exercises 9

- Relevant parts of the **Lecture notes**: Section 8.1 – 8.4.
- Relevant parts of the text books: **Biggs**: Sections 6.6, 9.1 – 9.5 and 9.7;
  Eccles: Chapters 13 – 14.

  (The treatment of this material in **Eccles** is somewhat different of the way we do it, so read those chapters with care.)
- Always justify your answers.

1 In the lectures, we gave a construction for the rational numbers. This started by looking at the set S of all pairs of the form (a, b), with $a, b \in \mathbb{Z}$ and $b \neq 0$, and then considering the relation R on R defined by:

$$(a,b)R(c,d) \qquad \text{if and only if} \qquad ad = bc.$$

(a) Explain why things would go badly wrong if we allow S to include pairs (a, b) with $b = 0$.

(The answer is **not** that if $b = 0$ the division a/b doesn't make sense. We define "division" only **after** we've established certain properties of the relation R, and defined $\mathbb{Q}$ and its arithmetic operations.)

We define the set $\mathbb{Q}$ to be the set of equivalence classes of the relation R, and defined an "addition" operation on $\mathbb{Q}$ by setting $[a, b] + [e, f] = [af + be, bf]$, for all $(a, b), (e, f) \in S$. (We're using the shorthand $[a, b]$ here for $[(a, b)]$.)

(b) Suppose that (a, b), (c, d) and (e, f) are in S, and that $(a, b)R(c, d)$. Show that this means that $(af + be, bf)R(cf + de, df)$.

Use this to show that if $(r, s)R(t, u)$ and $(v, w)R(x, y)$, then

$$[r,s] + [v,w] \;=\; [t,u] + [x,y].$$

(This means that the addition operation defined on $\mathbb{Q}$ is *well-defined*.)

2 (a) Write down 5 rational numbers lying strictly between 4/3 and 41/30.

(b) Explain how you would extend the method you used in (a) to find 5,000 rational numbers lying strictly between 4/3 and 41/30.

3 (a) Express the following rational numbers in decimal representation:

(i) $\dfrac{14}{13}$; (ii) $\dfrac{7}{-25}$.

(b) Express the recurring decimal $-0.8\overline{765}$ as a rational number p/q, where p and q are integers.

4 In the lectures and text books (Example 1 of Section 9.4 in Biggs; Theorem 13.3.2 in Eccles), it is proved that $\sqrt{2}$ is an irrational number. We also observe that $1 < \sqrt{2} < 2$.

(a) Prove that for any rationals x, y, with $y \neq 0$, also $x + y\sqrt{2}$ is an irrational number.

(b) Explain how you can use these facts about $\sqrt{2}$ to find an irrational number lying strictly between any two rational numbers a/b and c/d, where $a/b < c/d$.

5 If z is a real number, then z^n is defined recursively for non-negative integers n: $z^0 = 1$, and, for $n \geq 0$, $z^{n+1} = z \times z^n$.

If z is a positive real number, and b a natural number, then we define the positive b^{th} root $y = z^{1/b}$ to be the positive real number y such that $y^b = z$. You may assume that a positive real number always has a positive b^{th} root.

(a) Given a positive real number z, is it possible for there to be two different positive b^{th} roots of z? Justify your answer briefly.

(b) Suppose that a, b, c, d are natural numbers with $ad = bc$, and x is a positive real number. Show, from the definitions above, that $(x^a)^{1/b} = (x^c)^{1/d}$.

Explain why this result allows you to define x^q when x is a positive real number and q is a positive rational number.

(c) Show that $2^{3/4}$ is irrational.

Maple questions

6 The numbers x_n for $n \in \mathbb{N}$ are defined by

$$\begin{aligned} x_1 &= 1; \\ x_n &= \frac{1}{2}\left(x_{n-1} + \frac{3}{x_{n-1}}\right), \qquad \text{for } n \geq 2. \end{aligned}$$

(a) Prove that x_n is a rational number for all $n \geq 1$.

(b) Use Maple to calculate $x_1, x_2, \ldots, x_8$ as rational numbers.

The Maple command to find approximate decimal representations of numbers is "`evalf(.)`". By default it gives 10 digits of the results. Check the Maple help of "`evalf`" to find out how to obtain results with a different (prescribed) number of digits.

(c) Approximate the decimal representation of $x_1, x_2, \ldots, x_8$ with 20 digits precision. Calculate $\sqrt{3}$ with 20 digits precision as well.

(d) For what x_n does it appear that x_n gives an approximation of $\sqrt{3}$ with 20 or more digits precision?

(e) By further increasing the number of digits, try to find out in how many digits $\sqrt{3}$ and x_8 coincide.

Introduction to Abstract Mathematics
MA 103

Exercises 10

- Relevant parts of the **Lecture notes**: Section 8.5.
- The two text books **Biggs** and **Eccles** don't do complex numbers.

  You can find some further reading on complex numbers in one of the MA100 text books: M. Anthony and M. Harvey, *Linear Algebra: Concepts and Methods*, Chapter 13.
- Always justify your answers.

1 Write the following sum, product and quotient of complex numbers as complex numbers:

(a) $(2+3i)+(-5+4i)$; (b) $(2+3i)(-5+4i)$; (c) $\dfrac{2+3i}{-5+4i}$.

2 Find the square roots of the following numbers. (That is, for each of the numbers, find the complex numbers z such that z^2 equals the number.)

(a) -9; (b) i; (c) $4-3i$.

3 Solve the quadratic equation

$$x^2 - 2ix - 5 + 3i \;=\; 0.$$

4 DeMoivre's formula tells us that

$$(\cos\theta + i\sin\theta)^3 \;=\; \cos(3\theta) + i\sin(3\theta).$$

By expanding the left hand-side of this equation, and by using $\sin^2\theta + \cos^2\theta = 1$, prove the following trigonometrical identities:

$$\cos(3\theta) \;=\; 4\cos^3\theta - 3\cos\theta \qquad \text{and} \qquad \sin(3\theta) \;=\; 3\sin\theta - 4\sin^3\theta.$$

5 Write the following complex numbers in the standard form $x + yi$:

(a) $\left(\cos\left(\dfrac{3\pi}{14}\right) + i\sin\left(\dfrac{3\pi}{14}\right)\right)^7$; (b) $(\sqrt{3}+i)^{57}$.

6 Find the modulus and principal argument of the following complex numbers:

(a) $1+i$, (b) $-1-\sqrt{3}i$.

7 (a) Consider the equation $z^5 = a$, where a is a positive real number.

What can you say about the modulus of any solution z? What can you say about the principal argument?

Draw all the five solutions of $z^5 - 32 = 0$ on the Argand diagram.

(b) Write down a polynomial $P(z)$ such that

$$(z-2)P(z) \;=\; z^5 - 32.$$

What can you say about the solutions of $P(z) = 0$?

8 Consider the polynomial $P(z) = z^4 + z^2 - 2z + 6$.

(a) Show that $z^* = 1 + i$ is a root of $P(z)$, i.e. a solution of $P(z) = 0$.

(b) Hence find all the roots of $P(z)$.