1ACb: Algebra 1

School of Mathematics

UNIVERSITY OF BIRMINGHAM

Simon Goodwin
s.m.goodwin@bham.ac.uk
Office: Watson 107

Spring term 2016

Chapter -1

Module information

Before we get to the interesting mathematics in these notes, I need to cover some of the administrative details. **There's some important stuff here, so you should read it carefully at some point**.

Note that there is a possibility that some of the information given below may change slightly. Any changes will be announced on the 1AC Canvas page.

-1.1 Teaching arrangements

There will be 20 lectures for this course:

- Monday 11:00–12:00 in Physics Poynting Large LT, except in week 6.
- Friday 11:00–12:00 in Vaughan Jeffreys LT, except in week 11.

In addition you have the following opportunities to get help with the course.

- Support classes in weeks 4, 7, 9 and 11 of term.
- The PASS (Peer Assisted Study Scheme), you can find out more about this via the Year 1 Mathematics Canvas page.
- The drop-in sessions with postgraduate teaching assistants in the maths learning centre, you can find out more about this via the Year 1 Mathematics Canvas page.
- through the maths centre drop-in support in the main library, you can find about this from https://intranet.birmingham.ac.uk/as/libraryservices/library/skills/asc/mathematicalsupport.aspx.
- Your weekly tutorial meeting with your personal tutor.
- During my office hours, Monday 12:00–13:30 and Thursday 10:00–11:30.

In the lectures, I'll present all of the material of the course. Sometimes it may be difficult to keep up with the pace of the lectures, so **it's really important that you spend some time looking through the notes afterwards to make sure you understand them**. You will also find it very helpful to read the printed notes before the lecture. I will write on the board in lectures, and strongly recommend that you take notes, as I will present the material slightly differently and give different examples to those in the printed notes, as well as giving more details in some of the proofs.

The formative exercise sheets for 1ACb will be available on Canvas and on the first floor of the Watson building by 12 noon on Wednesday in weeks 3, 6, 8 and 10 of term. You will have one week to complete these exercises and then you will hand in your work to the white pigeonhole on the second floor of the Watson building for your support class. During this week you will be able to ask questions about the exercises in your weekly tutorial meeting, and in your support class.
It is compulsory that you hand in your solutions to the formative exercises on Wednesday at 12 noon in weeks 4, 7, 9 and 11.

On each formative exercise sheet, there will be four exercises for 1ACb and two of these exercises will be marked. There will also be comments on your work giving **feedback**. In addition, you can discuss your marked work in your weekly tutorial to get more **feedback**. A **feedback** sheet, which contains model solutions, comments on common mistakes and advice on how to improve will be posted on the 1AC Canvas page.

Additional exercises will be handed out in week 11 of term and model solutions will be put on the 1AC Canvas page during the Easter break. These exercises will cover the last parts of the course, and it is important that you work through over the Easter break as the topics they cover will be on the exam.

There are many exercises that are included within the printed notes at the end of each chapter. The questions on the formative exercise sheets are taken from those in the notes. You'll benefit from trying the other exercises too. Model solutions to these exercises will be added to the end of these notes at various stages during the term.

In my opinion working through the exercises is the most effective part of the learning process, so it is very important that you make a serious effort at all the exercises. When you work through the exercises you will attain a deeper understanding of the material.

-1.2 Any questions?

If you have any questions about the course, then you are encouraged to ask me after the lecture, email your question to s.m.goodwin@bham.ac.uk, or come to my office Watson 107 during my office hours:

- Monday 12:00–13:30 and Thursday 10:00–11:30.

If your question is urgent, or my office hours are not convenient, then you can try to find me to at another time and see if I am available, or you can email me to make an appointment.

-1.3 Course materials

All course materials will be available on the 1AC Canvas page.

These printed lecture notes will be released during the term, and will be available from the first floor of the Watson building. In preparing these notes, I have used the books listed below and the notes from a previous course by Paul Flavell. There are likely to be some typos or little mistakes, all of which are my fault, so please let me know if you find any. You will be able to make another set of notes during the lectures. As I mentioned

above **you will benefit from reading the printed notes before the lectures, and it is very important that you study your notes after the lectures**.

The formative exercise sheets for 1ACb will be available on Canvas and on the first floor of the Watson building by 12 noon on Wednesday in weeks 3, 6, 8 and 10 of term. Feedback sheets for the formative exercises will be available on the 1AC Canvas page. The additional exercises will be available in week 11.

-1.4 Books

Most of the material from the course is covered in

- Martin Liebeck, *A concise introduction to pure mathematics*, Fourth Edition, Chapman and Hall.

This is the recommended book for the course.

There are a lot of other good books that cover material in this course, many of which are in the library; for example,

- Carol Whitehead, *Guide*^{2} *abstract algebra*, Second Edition, Palgrave MacMillan.

A book that may be useful for learning about the structures of proofs, and contains some of the material in this course is

- Daniel J. Velleman, *How to prove it: a structured approach*, Second Edition, Cambridge University Press.

I have used the book

- Niels Lauritzen, *Concrete abstract algebra*, Third Edition, Chapman and Hall.

when preparing the course. However, this book is a bit more advanced and only mainly recommended for reading beyond the scope of the course, though it would be helpful for some parts of the course.

Another book that covers a lot of the material in the course (and much more) is

- Peter J. Cameron, *Introduction to Algebra*, Second Edition, Oxford University Press.

It is quite brief in places, so may be a bit too advanced for this course. This is also the recommended book for Algebra 2, which you can take next year.

-1.5 Assessment

1AC (1ACa and 1ACb combined) is worth 20 credits. The assessment is divided up as follows:

- 80% from one 3-hour examination in the summer;
- 20% from class tests during term.

You have taken two class tests in 1ACa last term, which contributed a total of 10%. There will be two class tests that you take this term:

- a written class test on Tuesday 13:00–14:00 in week 6 of term in Haworth 101;
- an online class test in your computer lab session in week 11 of term.

Each of these tests will contribute 5%, giving a total of 10%.

Feedback for the class tests will be provided on the 1AC Canvas page

-1.6 Syllabus

The syllabus below contains a list of topics that are covered in the course.

- **Proofs and prime numbers:** Some interesting examples of proofs and counterexamples about prime numbers.
- **The integers:** factors and prime numbers; the division theorem; highest common factors; the Euclidean algorithm; primes and products; the fundamental theorem of arithmetic.
- **Modular arithmetic:** congruence modulo n; arithmetic with congruences; congruence equations; the Chinese remainder theorem; the ring of integers modulo n; Fermat's little theorem; RSA cryptography.
- **Permutations:** two-row and cycle notation; composition and inversion; permutations can be written as a products of transpositions; even and odd permutations.
- **Groups:** permutation groups; axioms for a group; examples of groups; subgroups; cosets; Lagranges theorem.

-1.7 Learning outcomes

By the end of this course you should be able to:

- understand elementary number theory and make calculations in examples;
- understand congruences of integers and make calculations in examples;
- calculate with both the two-row and cycle notation for permutations;
- understand basic group theory, and make calculations in examples;
- write proofs or provide counterexamples for statements concerning the material in the course; and
- apply the material in the course to solve weakly-posed problems and prove related statements.

The learning outcomes above are statements of what "a learner is expected to know, understand and be able to demonstrate after completion of a process of learning". These are important as they describe what you are expected to be able to do in order to demonstrate that you have understood the course. The assessment of the course is based on these outcomes, so **if you can do the things on this list, then you should do well in the assessment.** Also at the end of each chapter, I have provided a list of more specific learning aims in a summary.

A bit of a warning

A few changes to the lecture notes have been made recently, so there are a very likely to be some typos and some bits of the notes that are not set out as well as they could be. The version on canvas will be kept updated with any changes and corrections. Please let me know if you spot any typos, or anything that you think may be a mistake. It may be possible to provide an updated version of these notes at the end of term.

Chapter 0

Notation

We recap a little bit of notation about sets, which has been covered in previous courses. **This stuff isn't very interesting so you should skip it at first and just look back to it if you need to.**

Definition 0.1. A *set* is a collection of objects.
We write:

- $a \in A$ to mean a is an element of A;
- $a \notin A$ to mean a is not an element of A; and
- $A \subseteq B$ to mean that A is a subset of B; this means that all elements of A are elements of B.

Now we define some sets of numbers that you are familiar with.

Definition 0.2. (a) We write $\mathbb{N}$ for the set of *natural numbers*:

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

Note that 0 is not a natural number, though sometimes we want to consider the set consisting of 0 and the natural numbers, and we use the notation

$$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}.$$

(b) We write $\mathbb{Z}$ for the set of *integers*:

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

(c) We write $\mathbb{R}$ for the set of *real numbers*. These are numbers that can be written using a decimal expansion.

(d) We write $\mathbb{Q}$ for the set of *rational numbers*. These are the real numbers that can be written as a fraction, so

$$\mathbb{Q} = \left\{ \frac{a}{b} \in \mathbb{R} \ : \ a \in \mathbb{Z} \text{ and } b \in \mathbb{N} \right\}.$$

(e) We write $\mathbb{C}$ for the set of *complex numbers*. These are expressions of the form $a+bi$, where i is a square root of -1, so

$$\mathbb{C} = \{a + bi \ : \ a, b \in \mathbb{R}\}.$$

We demonstrate some notation for sets by example. Sometimes we give a set by listing its elements. For example

$$\{1, 3, 6\}$$

is a set with three elements. We used similar notation for $\mathbb{N}$ and $\mathbb{Z}$ above. Another example, the set of positive real numbers, is neatly described by the notation

$$\{x \in \mathbb{R} \ : \ x > 0\}.$$

The colon : can be read as "such that", so the set above is the set of $x \in \mathbb{R}$ such that $x > 0$. We used similar notation to describe $\mathbb{Q}$. Sometimes a vertical line | is written rather than a colon :. Another example is

$$\{x \in \mathbb{Z} \ : \ x^2 < 5\} = \{-2, -1, 0, 1, 2\}.$$

You shouldn't write a colon instead of "such that" outside of sets though.

One piece of notation that we'll use and may be a bit different from what you've done before is that we often use a dot to denote multiplication of integers. For example we write $3 \cdot 5$ to denote 3 multiplied by 5, which is of course equal 15, so we have $3 \cdot 5 = 15$.

Chapter 1

Introduction to proofs and prime numbers

The first part of this course covers some elementary number theory and includes some really nice theorems and applications; later we will cover some other interesting areas of algebra. As with some of the others previous courses that you have taken at university, the material in this course is likely to be different from the mathematics that you have seen before university. Therefore, it may seem difficult to begin with, but with some perseverance you will be able to grasp the topic and I hope you will enjoy it!

An emphasis in this course is put on setting out pure mathematics well and writing proofs, and this is the most important skill for you to improve during this course. In this first chapter we recap some ideas about definitions, proofs and counterexamples, by giving some examples of interesting proofs about prime numbers.

1.1 What is a prime number?

You all should know what a prime number is. As we are going to be proving theorems about prime numbers, we need to make sure that a prime number means the same thing to all of us. For example, we need to decide whether 1 is a prime number. Therefore, we need a *definition* of a prime number.

Definition 1.1. A natural number $n \in \mathbb{N}$ is a *prime number* if $n \neq 1$ and the only positive factors of n are 1 and n.

We immediately see that there is a problem with this definition, as we don't yet know what we mean by factors. So we better define this now.

Definition 1.2. Let $a, b \in \mathbb{Z}$. We say that a *is a factor of* b if there exists $z \in \mathbb{Z}$ such that $b = az$. We write $a \mid b$ to mean that a is a factor of b, and $a \nmid b$ to mean that a is not factor of b.

Sometimes we say a *divides* b or b *is divisible by* a to mean the same thing as a is a factor of b.

For $a \neq 0$, we remark that saying that a is a factor of b is equivalent to $\frac{b}{a} \in \mathbb{Z}$; you may find it helpful to think of it in this way when we are learning more about factors later.

Examples 1.3.

(a) $7 \mid 21$, because $21 = 7 \cdot 3$. (The dot here is a shorthand for multiplication.)
(b) $4 \nmid 19$, because if $19 = 4z$, then $z = \frac{19}{4}$, which is not an integer.
(c) Let $n \in \mathbb{N}$. Then $n \mid n$, because $n = n \cdot 1$.

Definition 1.1 may seem very formal, but it is important that we have an agreement of exactly what it means for a natural number to be prime. From now we do not argue about what a prime number is – the above definition gives the answer. In particular, 1 is not a prime number because the definition says that it is not. From the definition we can work out the first few prime numbers:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots.$$

Note that from Definitions 1.2 and 1.1, if $n \in \mathbb{N}$ with $n \neq 1$ is not prime, then there exist $a, b \in \mathbb{N}$ with $1 < a, b < n$ such that $n = ab$. We often refer to natural numbers greater than 1 that are not prime as composite; this is stated in the following definition.

Definition 1.4. A natural number $n \in \mathbb{N}$ is called a *composite number* if $n \neq 1$ and there exist $a, b \in \mathbb{N}$ with $1 < a, b < n$ such that $n = ab$.

1.2 Mersenne numbers and primes

We are going to look at an interesting sequence of numbers called the Mersenne numbers. They are named after a French monk and scholar, Father Marin Mersenne (1585–1647), who studied them. A *Mersenne number* is a number of the form $2^n - 1$ for some $n \in \mathbb{N}$. In the table below I have listed the first 10 Mersenne numbers.

n	1	2	3	4	5	6	7	8	9	10
$2^n - 1$	1	3	7	15	31	63	127	255	511	1023

Let us observe that for each value of n in the table:

- if n is prime, then $2^n - 1$ is prime; and
- if n is composite, then $2^n - 1$ is composite.

For example, 7 is prime and 127 is prime, and 9 is not prime and $511 = 7 \cdot 73$ is not prime. It is, therefore, tempting to guess that this is always the case. In mathematics, a guess based on some evidence is called a *conjecture*. So we have the following two conjectures.

Conjecture 1.5. *Let $n \in \mathbb{N}$. Suppose that n is prime. Then $2^n - 1$ is prime.*

Conjecture 1.6. *Let $n \in \mathbb{N}$. Suppose that n is composite. Then $2^n - 1$ is composite.*

It turns out that Conjecture 1.5 is not true. To check that it is not true we only need to find one value of n for which n is prime and $2^n - 1$ is not prime. If we consider the prime 11, then we see that

$$2^{11} - 1 = 2047 = 23 \cdot 89$$

is composite. This means that $n = 11$ is a *counterexample* to Conjecture 1.5, so that the conjecture is not true.

We would like to decide if Conjecture 1.6 is true. If we continued the table to include all n up to 30, then we find out that 23 and 29 are other counterexamples to Conjecture 1.5:

$$2^{23} - 1 = 8{,}388{,}607 = 47 \cdot 178{,}481 \qquad 2^{29} - 1 = 536{,}870{,}911 = 2{,}089 \cdot 256{,}999.$$

However, there is no natural number $n \leq 30$ such that n is not prime and $2^n - 1$ is prime. So the evidence suggests that the statement is true. In order to be absolutely sure we need to give a *proof*.

Proof of Conjecture 1.6. Since n is composite, there exists $a, b \in \mathbb{N}$ with $1 < a, b < n$ such that $n = ab$. Consider the identity

$$t^m - 1 = (t - 1)(1 + t + t^2 + \cdots + t^{m-1}),$$

for $m \in \mathbb{N}$. We apply this with $t = 2^b$ and $m = a$ to get

$$\begin{aligned} 2^n - 1 &= 2^{ab} - 1 \\ &= (2^b)^a - 1 \\ &= (2^b - 1)(1 + 2^b + (2^b)^2 + \cdots + (2^b)^{(a-1)}) \\ &= (2^b - 1)(1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}). \end{aligned}$$

Let $x = 2^b - 1$ and $y = 1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}$, then $2^n - 1 = xy$. Since $1 < b < n$, we have $1 < x < 2^n - 1$, which also implies that $1 < y < 2^n - 1$. Hence, $2^n - 1$ is composite. □

Now that we have a proof of Conjecture 1.6, we know beyond any doubt that it is true and we can call it a *theorem*; which we state below.

Theorem 1.6. *Let $n \in \mathbb{N}$. Suppose that n is composite. Then $2^n - 1$ is composite.*

We can use the proof of Conjecture 1.6 to find factors of large numbers of the form $2^n - 1$.

Example 1.7. To find factors of $32767 = 2^{15} - 1$ we can write

$$\begin{aligned} 2^{15} - 1 &= (2^5 - 1)(1 + 2^5 + 2^{10}) \\ &= (32 - 1)(1 + 32 + 1024) \\ &= 31 \cdot 1057. \end{aligned}$$

Alternatively, we can write

$$\begin{aligned} 2^{15} - 1 &= (2^3 - 1)(1 + 2^3 + 2^6 + 2^9 + 2^{12}) \\ &= (8 - 1)(1 + 8 + 64 + 512 + 4096) \\ &= 7 \cdot 4681. \end{aligned}$$

From these factorizations, we see that 7 must divide 1057, and we obtain

$$32767 = 7 \cdot 31 \cdot 151.$$

It is straightforward to check that 151 is prime, so we have factorized 32767 as a product of prime numbers.

Although we know that not all numbers of the form $2^p - 1$ with p prime are prime, these numbers are still very interesting. Mersenne numbers that are prime are called *Mersenne primes*. At present, 48 Mersenne primes have been found and it is unknown whether there are infinitely many. The largest known prime number is the Mersenne prime $2^{57,885,161} - 1$. It has 17,425,170 digits and was found by the Great Internet Mersenne Prime Search, see http://en.wikipedia.org/wiki/Great_Internet_Mersenne_Prime_Search. It was discovered in January 2013, and if you were to write out this number on a long reel of paper taking 1cm for each digit, then it would reach from Birmingham to London!

1.3 The infinitude of primes

Above we have seen that there are very large prime numbers, and you may suspect that there are infinitely many prime numbers. Below we state this as a theorem and prove it, so we know beyond any doubt that it is true. This proof was first given by Euclid in around 350BC, and is one of the most famous proofs in mathematics.

Theorem 1.8. *There are infinitely many prime numbers.*

Proof. Suppose for a contradiction that there are not infinitely many prime numbers. Then there are a finite number of primes and we can write down the finite list of prime numbers

$$p_1, p_2, \ldots, p_r.$$

Let

$$n = p_1 p_2 \ldots p_r + 1.$$

Then $n > p_i$ for all $i = 1, 2, \ldots, r$. Since our list gives all of the prime numbers, this means that n is not a prime number. Therefore, there is a factor d of n with $1 < d < n$. We choose $1 < d < n$ to be a factor of n that is as small as possible. Then d must be a prime number, because if c is a factor of d with $1 \leq c < d$, then c is also a factor of n that is smaller than d so must be equal to 1. So $d = p_i$ for some $i = 1, 2, \ldots, r$.

Therefore, we have $\frac{n}{p_i} \in \mathbb{Z}$. But also we have

$$\begin{aligned} \frac{n}{p_i} &= \frac{p_1 p_2 \ldots p_r + 1}{p_i} \\ &= p_1 p_2 \ldots p_{i-1} p_{i+1} \ldots p_r + \frac{1}{p_i} \notin \mathbb{Z}. \end{aligned}$$

Thus, we have a contradiction, because $\frac{n}{p_i}$ cannot both be an integer and not an integer.

The only conclusion that we can draw is that our initial assumption that there are only finitely many prime numbers must be wrong. Therefore, there are infinitely many prime numbers as required. □

This proof above is an example of a *proof by contradiction*. We'll see more of these types of proofs later.

1.4 The distribution of the primes

We now know that there are infinitely many prime numbers, so we can ask how they are distributed in all the natural numbers. Calculations suggest that the prime numbers are more sparsely distributed as we look at larger numbers. For example, there are 25 primes between 1 and 100, 16 primes between 1000 and 1100 and only 6 primes between 1,000,000 and 1,000,100. To demonstrate this thinning out we prove that we can find an arbitrarily large gap between prime numbers. For the proof recall that for $n \in \mathbb{N}$, we define n *factorial* by

$$n! = 1 \cdot 2 \cdot 3 \cdots n.$$

Theorem 1.9. *Let $n \in \mathbb{N}$. There exists a sequence of n consecutive natural numbers containing no primes.*

Proof. Let $m = n + 1$. We will show that none of the n consecutive integers

$$m! + 2, m! + 3, m! + 4, \ldots, m! + m$$

are prime. First consider,

$$\begin{aligned} m! + 2 &= (1 \cdot 2 \cdot 3 \cdot 4 \cdots m) + 2 \\ &= (2 \cdot 1 \cdot 3 \cdot 4 \cdots m) + 2 \\ &= 2 \cdot (1 \cdot 3 \cdot 4 \cdots m + 1). \end{aligned}$$

Therefore, 2 is a factor of $m! + 2$, so $m! + 2$ is not prime. Similarly,

$$\begin{aligned} m! + 3 &= (1 \cdot 2 \cdot 3 \cdot 4 \cdots m) + 3 \\ &= (3 \cdot 1 \cdot 2 \cdot 4 \cdots m) + 3 \\ &= 3 \cdot (1 \cdot 2 \cdot 4 \cdots m + 1). \end{aligned}$$

Therefore, 3 is a factor of $m! + 3$, so $m! + 3$ is not prime. In general, consider $m! + i$ where $2 \leq i \leq m$. We have

$$\begin{aligned} m! + i &= (1 \cdot 2 \cdots (i-1) \cdot i \cdot (i+1) \cdots m) + i \\ &= (i \cdot 1 \cdot 2 \cdots (i-1) \cdot (i+1) \cdots m) + i \\ &= i \cdot (1 \cdot 2 \cdots (i-1) \cdot (i+1) \cdots m + 1). \end{aligned}$$

Therefore, i is a factor of $m! + i$, so $m! + i$ is not prime.

Hence, we have found a sequence of n consecutive natural numbers containing no prime numbers. □

We finish this chapter by briefly mentioning some more advanced statements about the distribution of the primes.

The *prime number theorem* is a remarkable theorem that tells us approximately how frequently primes occur as we look at larger numbers. Roughly it says that if we pick $n \in \mathbb{N}$ at random near a large number N, then the probability that n is prime is about

$$\frac{1}{\log_e(N)}.$$

It was conjectured by Gauss in 1793 and proved by Hadamard and de la Valée Pousin in 1896. You may ask why does e show up here! You can find out more at
http://en.wikipedia.org/wiki/Prime_number_theorem

A much deeper question about the distribution of primes is whether there is a pattern to how prime numbers lie in all the natural numbers. This is perhaps the biggest open problem in number theory. The *Riemann hypothesis*, which was proposed by B. Riemann in 1859, is a conjecture which implies a lot about the distribution of the primes. Proving that the Riemann hypothesis is true is one of the seven Millennium problems proposed by the Clay Mathematics Institute and there is a $1,000,000 prize for its solution. You can find out more at
http://en.wikipedia.org/wiki/Riemann_hypothesis

Another really interesting problem about prime numbers is the *twin primes conjecture*. This states that there are infinitely many pairs $(p, p+2)$ where both p and $p+2$ are prime. So the conjecture is saying that although we know that the primes become more sparsely distributed as we look at larger numbers, there are still infinitely many prime twins that are as close together as possible. You can find out more at
http://en.wikipedia.org/wiki/Twin_prime.

The twin primes conjecture remains an open problem, though there has been some major progress recently. In 2013, Yitang Zhang announced a proof that for some integer N less than 70 million, there are infinitely many pairs of primes that differ by at most N. Subsequently, there has been a flurry of research activity by many mathematicians around the world and the bound for N has been reduced to 246. If you're interested in finding out a bit more, then you could look at
http://arxiv.org/abs/1409.8361.

1.5 Summary of Chapter 1

At the end of each chapter of these notes, I will summarize the material in the chapter by giving a list of learning aims for the chapter. These aims are more specific than the learning outcomes that were given in Section -1.7 and serve the same purpose of informing you what you should be able to do to demonstrate that you have understood the chapter. As the main goal of this chapter is to give some interesting proofs the list here is quite short.

By the end of this chapter you should be able to:

- state the definition of a factor and a prime number; and
- appreciate the proofs given in this chapter.

Don't worry too much if the proofs in this chapter seems a little challenging at the moment, once we've covered a few more you will get better at them.

1.6 Exercises for Chapter 1

Some of the exercises are quite challenging, and it is not expected that you will be able to do them all straightaway, but you'll benefit from attempting them all.

Q1.1. True or false?

(a) $3 \mid 21$
(b) $7 \mid 13$
(c) $4 \mid 2$
(d) $17 \mid 0$

You should justify your answers.

Q1.2. (a) Use the proof of Conjecture 1.6 to factorize $16383 = 2^{14} - 1$ as a product of smaller natural numbers.
(b) Do (a) again in a different way.
(c) Use this to factorize 16383 as a product of prime numbers.

Q1.3. Consider natural numbers of the form

$$\frac{2^n + 1}{3}$$

where $n \in \mathbb{N}$ is odd. Investigate when they are prime.

You can proceed as we did for numbers of the form $2^n - 1$ in Section 1.2. A good starting point is to list the first 10 numbers of the form $\frac{2^n+1}{3}$ for n odd, then make two conjectures analogous to Conjectures 1.5 and 1.6. To determine whether a number is prime it may be worth using a computer programme like the one on
http://primes.utm.edu/curios/includes/primetest.php.
You may find the identity,

$$t^m + 1 = (t + 1)(1 - t + t^2 - t^3 + \cdots + t^{m-1}).$$

for m odd useful.

Q1.4. A natural number $n \in \mathbb{N}$ is called a *perfect number* if n is equal to the sum of its positive factors less than n.

(a) Show that 6 and 28 are perfect numbers.
(b) Suppose that $2^p - 1$ is a Mersenne prime. Show that $2^{p-1}(2^p - 1)$ is perfect.

Hint: Try to write down all the factors of $2^{p-1}(2^p - 1)$. Then use the formula

$$1 + 2 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$$

for any $n \in \mathbb{N}$.

Q1.5. Any odd integer can be written in the form $4k+1$ or $4k+3$ for some $k \in \mathbb{Z}$.

(a) Let $k, l \in \mathbb{Z}$.

(i) Show that $(4k+1)(4l+1)$ is of the form $4m+1$ for some $m \in \mathbb{Z}$.
(ii) Show that $(4k+1)(4l+3)$ is of the form $4m+3$ for some $m \in \mathbb{Z}$.
(iii) Show that $(4k+3)(4l+3)$ is of the form $4m+1$ for some $m \in \mathbb{Z}$.

(b) Give a proof of the theorem below, which is similar to Theorem 1.8 in the lecture notes.

Theorem. *There are infinitely many prime numbers of the form $4k+3$ for $k \in \mathbb{Z}$.*

Hint: First suppose that there are finitely many primes of the form $4k+3$ and list them as in the proof of Theorem 1.8. Then multiply them all together to get $s \in \mathbb{N}$ and consider $2s+1$. Use (a) to show that $2s+1$ is of the form $4l+3$. Then use (a) again to show that $2s+1$ must have a prime factor of the form $4k+3$.

Q1.6. Let $r \in \mathbb{N}$ and let $p_1, p_2, \ldots, p_r \in \mathbb{N}$ denote the first r prime numbers, and consider $n_r = p_1 p_2 \cdots p_r$.

(a) Is n_r prime for every $r \in \mathbb{N}$?
(b) Does your answer to (a) contradict the proof of Theorem 1.8?
(c) Explain your answer to (b).

Q1.7. $3, 5, 7$ is a list of three primes of the form $p, p+2, p+4$. Prove that there are no other "triplet primes"?
Hint: Use the fact that any $n \in \mathbb{N}$ can be written in the form $3k+l$, where $k \in \mathbb{N}_0$ and $l \in \{0, 1, 2\}$.

The last question is here, so you can read up on some really interesting statements about prime numbers. You shouldn't spend much time thinking about these, but rather research them by looking them up.

Q1.8. (a) Let $n \in \mathbb{N}$ be even with $n > 2$. Can n be written at the sum of two prime numbers?
(b) Let $n \in \mathbb{N}$ be odd with $n > 5$. Can n be written at the sum of three prime numbers?
You should look up Goldbach's conjecture and Goldbach's weak conjecture. Wikipedia is probably a good place to start.

Chapter 2

The integers

In this chapter, we make a structured approach to the properties of the integers. One of ours goals is to prove the fundamental theorem of arithmetic, which roughly says that any natural number can be factorized uniquely as a product of prime factors. A precise statement is given in Theorem 2.22. *If you are ever asked to state the fundamental theorem of arithmetic then you should write down the precise statement in Theorem 2.22.*

Before we can state and prove the fundamental theorem of arithmetic we have to cover some material on factors and prime numbers. Of particular importance is the Euclidean algorithm, which is given in Section 2.4. Later in Section 2.9 we give some rather nice consequences of the fundamental theorem of arithmetic.

2.1 Factors of integers

Recall that we defined factors of integers in Definition 1.2. Below we give some elementary lemmas about factors of numbers. (A lemma is a name for a "little theorem"; often lemmas are used in the proofs of theorems.) We demonstrate the first of these lemmas with a couple of examples.

Examples 2.1.

(a) We have that $13 \mid 26$, because $26 = 2 \cdot 13$, and $13 \mid 78$ because $78 = 13 \cdot 6$.
Then $13 \mid 104 = 26 + 78$, because $104 = 13 \cdot 8 = 13 \cdot (2 + 6)$.

(b) We have that $7 \mid (-28)$, because $28 = 7 \cdot (-4)$, and $7 \mid 91$ because $91 = 7 \cdot 13$.
Then $7 \mid 63 = -28 + 91$, because $63 = 7 \cdot 9 = 7 \cdot (-4 + 13)$.

These examples may seem a bit trivial, but the important point is that it gives us an idea of how to prove the following lemma.

Lemma 2.2. *Let $a, b, c \in \mathbb{Z}$. Suppose that $a \mid b$ and $a \mid c$. Then $a \mid (b + c)$.*

Proof. Since $a \mid b$, there exists $x \in \mathbb{Z}$ such that

$$b = ax. \tag{2.1}$$

Since $a \mid c$, there exists $y \in \mathbb{Z}$ such that

$$c = ay. \tag{2.2}$$

Adding (2.1) and (2.2) gives

$$b + c = ax + ay = a(x + y).$$

Let $z = x + y$. Then $z \in \mathbb{Z}$, because $x, y \in \mathbb{Z}$ and

$$b + c = az.$$

Therefore, $a \mid (b + c)$. □

The next lemma collects together some more general properties of factors. The proof of this lemma is Exercises Q2.1, Q2.2 and Q2.3.

Lemma 2.3. *Let $a, b, c, k, l \in \mathbb{Z}$.*

(a) *Suppose that $a \mid b$ and $a \mid c$. Then $a \mid (kb + lc)$.*
(b) *Suppose that $a \mid b$ and $b \mid c$. Then $a \mid c$.*
(c) *Suppose that $a \mid b$ and $b \mid a$. Then $a = \pm b$.*

You should notice how central the definition of being a factor is to the proofs of these lemmas. The proof starts by using the definition to write down what the hypothesis says. Then it ends with a sentence saying that the conclusion holds in terms of the definition. The important point I'm hoping to make here is that we have to use definitions properly in proofs.

2.2 The division theorem

The division theorem should be very familiar it essentially says that when we divide an integer by a positive integer we obtain a quotient and remainder. It may seem that we are being very formal here, but our statement is very clear and the proof confirms beyond doubt something we have believed for a long time.

Theorem 2.4. *Let $a \in \mathbb{Z}$ and $d \in \mathbb{N}$. Then there exist unique $q, r \in \mathbb{Z}$ such that*

$$a = qd + r \text{ and } 0 \leq r < d.$$

Proof. Let q be the largest integer such that $q \leq \frac{a}{d}$, and let $r = a - qd$. Then $a = qd + r$. Clearly $r \geq 0$.
If $r \geq d$, then $\frac{r}{d} \geq 1$, so $q + 1 \leq q + \frac{r}{d} = \frac{a}{d}$. But q was chosen to be maximal such that $q \leq \frac{a}{d}$, so $r < d$.

So far, we have proved that there exist $q, r \in \mathbb{Z}$ such that

$$a = qd + r \text{ and } 0 \leq r < d.$$

We also need to prove uniqueness. Suppose that there are $q, q', r, r' \in \mathbb{Z}$ such that

$$a = qd + r = q'd + r'$$

and $0 \leq r, r' < d$. Then

$$r - r' = d(q' - q).$$

If $q' - q \neq 0$, then $|q - q'| \geq 1$, which implies that $|r - r'| = d|q - q'| \geq d$. But this is not possible, because $0 \leq r, r' < d$. Therefore, we must have $q' - q = 0$, so $q = q'$ and $r = r'$ too. This proves uniqueness. □

We say that q is the *quotient* and r is the *remainder* when a is divided by d.

We note that $r = 0$ precisely when d is a factor of a. In other words d is a factor of a if and only if the remainder when a is divided by d is zero.

Example 2.5. Let $a = 137$ and $d = 11$. Then by performing long division we see that

$$137 = 12 \cdot 11 + 5.$$

So $q = 12$ and $d = 5$ in this case.

2.3 Highest common factors

An important notion for us is the highest common factor of two integers, which is defined below.

Definition 2.6. Let $a, b \in \mathbb{Z}$.

(a) A *common factor* of a and b is an integer c such that $c \mid a$ and $c \mid b$.

(b) The *highest common factor* of a and b is the largest integer h that is a common factor of a and b. We write $h = \mathrm{hcf}(a, b)$.

Note that if $a = b = 0$, then $\mathrm{hcf}(a, b)$ is not defined.

Also we note that clearly, $\mathrm{hcf}(a, b) = \mathrm{hcf}(b, a)$ and $\mathrm{hcf}(a, b) = \mathrm{hcf}(-a, b)$.

The highest common factor of a and b is sometimes called the *greatest common divisor* of a and b, and denoted $\gcd(a, b)$. In these notes we will always say highest common factor.

Examples 2.7.

(a) Let $a = 28$ and $b = 42$. By working out all the factors of a and b we calculate that the common factors of a and b are

$$-14, -7, -2, -1, 1, 2, 7, 14.$$

Therefore, the highest common factor of a and b is

$$\mathrm{hcf}(a, b) = 14.$$

(b) Let $a = 27$ and $b = 20$. By working out all the factors of a and b we calculate that the only common factors of a and b are -1 and 1, so $\mathrm{hcf}(a, b) = 1$.

(c) Let $a, b \in \mathbb{Z}$ and suppose that $b \mid a$. Then $\mathrm{hcf}(a, b) = b$.
In particular, this implies that $\mathrm{hcf}(0, b) = b$.

2.4 The Euclidean algorithm

Example 2.7 shows that it is easy to calculate $\mathrm{hcf}(a, b)$ when a and b are small, by working out all the factors of a and b. When a and b are large this becomes impractical. There is a more efficient way to calculate the highest common factors called the Euclidean algorithm. The following lemma is key for the Euclidean algorithm.

Lemma 2.8. *Let $a, b, q, r \in \mathbb{Z}$. Suppose that $a = qb + r$. Then*

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

Proof. Let c be a common factor of a and b. Then by Lemma 2.3(a), c is a factor of $r = a - qb$, and thus a common factor of b and r.

Now let c be a common factor of b and r. Then by Lemma 2.3(a), c is a factor of $a = qb + r$, and thus a common factor of a and b.

Therefore, a and b have the same common factors as b and r, and hence the same highest common factor. □

Before formally stating the Euclidean algorithm we demonstrate it with an example.

Example 2.9. We are going to use Lemma 2.8 to calculate the highest common factor of 1989 and 1508.

First, using Theorem 2.4, we write

$$1989 = 1 \cdot 1508 + 481$$

and use Lemma 2.8 to deduce that

$$\mathrm{hcf}(1989, 1508) = \mathrm{hcf}(1508, 481).$$

Next we write

$$1508 = 3 \cdot 481 + 65$$

and use Lemma 2.8 again to deduce that

$$\mathrm{hcf}(1508, 481) = \mathrm{hcf}(481, 65).$$

For the third step we write

$$481 = 7 \cdot 65 + 26$$

and deduce that

$$\mathrm{hcf}(481, 65) = \mathrm{hcf}(65, 26).$$

For the fourth step we write

$$65 = 2 \cdot 26 + 13$$

and deduce that

$$\mathrm{hcf}(65, 26) = \mathrm{hcf}(26, 13).$$

In the fifth step we write

$$26 = 2 \cdot 13 + 0,$$

so $13 \mid 26$, and thus

$$\mathrm{hcf}(26, 13) = 13.$$

Putting all this together we obtain

$$\mathrm{hcf}(1989, 1508) = \mathrm{hcf}(1508, 481) = \mathrm{hcf}(481, 65) = \mathrm{hcf}(65, 26) = \mathrm{hcf}(26, 13) = 13.$$

So we have calculated $\mathrm{hcf}(1989, 1508) = 13$.

We now give a formal description of the Euclidean algorithm.

Algorithm 2.10 (Euclidean Algorithm)**.**

Input: $a, b \in \mathbb{N}$ with $a \geq b > 0$, and set $a_0 = a$, $a_1 = b$.

1**st step**: Find $q_1, a_2 \in \mathbb{Z}$ with

$$a_0 = a_1 q_1 + a_2 \quad \text{and} \quad 0 \leq a_2 < a_1.$$

If $a_2 = 0$, then we output $\text{hcf}(a, b) = a_1$ and stop.
If $a_2 \neq 0$, then we proceed to the 2nd step.

2**nd step**: Find $q_2, a_3 \in \mathbb{Z}$ with

$$a_1 = a_2 q_2 + a_3 \quad \text{and} \quad 0 \leq a_3 < a_2.$$

If $a_3 = 0$, then we output $\text{hcf}(a, b) = a_2$ and stop.
If $a_3 \neq 0$, then we proceed to the 3rd step.

$\vdots \qquad \vdots \qquad \vdots$

k**th step**: Find $q_k, a_{k+1} \in \mathbb{Z}$ with

$$a_{k-1} = a_k q_k + a_{k+1} \quad \text{and} \quad 0 \leq a_{k+1} < a_k.$$

If $a_{k+1} = 0$, then we output $\text{hcf}(a, b) = a_k$ and stop.
If $a_{k+1} \neq 0$, then we proceed to the $(k+1)$th step.

We make two comments about this algorithm. First we note that in the kth step we can find the required $q_k, a_{k+1} \in \mathbb{Z}$ using Theorem 2.4. Second, we note that as $a_0 > a_1 > a_2 > \ldots$, we must eventually get $a_{k+1} = 0$ so that the algorithm does terminate.

We give another example of the use of the Euclidean algorithm in Example 2.13. Below we prove that the Euclidean algorithm does correctly calculate highest common factors. The idea of the proof is given by Example 2.9.

Theorem 2.11. *Let $a, b \in \mathbb{N}$ with $a > b$, and let h be the output of Algorithm 2.10. Then $h = \text{hcf}(a, b)$.*

Proof. Suppose the algorithm terminates on the kth step, Then $h = a_k$. We have

$$a_{k-1} = a_k q_k,$$

so $a_k \mid a_{k-1}$ and $\text{hcf}(a_{k-1}, a_k) = a_k$.

Consider the lth step for $1 \leq l < k$. We have

$$a_{l-1} = a_l q_l + a_{l+1},$$

so

$$\text{hcf}(a_{l-1}, a_l) = \text{hcf}(a_l, a_{l+1})$$

by Lemma 2.8.

We obtain the sequence of equalities:

$$\begin{aligned} h = a_k &= \mathrm{hcf}(a_{k-1}, a_k) \\ &= \mathrm{hcf}(a_{k-2}, a_{k-1}) \\ &\;\;\vdots \qquad\quad \vdots \\ &= \mathrm{hcf}(a_1, a_2) \\ &= \mathrm{hcf}(a_0, a_1) \\ &= \mathrm{hcf}(a, b). \end{aligned}$$

□

2.5 Bézout's lemma and reversing the Euclidean algorithm

The next theorem gives an important property of highest common factors; it is often called Bézout's lemma.

Theorem 2.12 (Bézout's lemma)**.** *Let $a, b \in \mathbb{Z}$ with $a \neq 0$ and let $h = \mathrm{hcf}(a, b)$. Then there exist $x, y \in \mathbb{Z}$ such that*

$$h = xa + yb.$$

In other words h can be expressed as an integral linear combination *of a and b*

We note that the condition $a \neq 0$ in the statement of Bézout's lemma is only required so that $\mathrm{hcf}(a, b)$ is defined; we could have also just as well have assumed that $b \neq 0$

Before giving a proof, we demonstrate how we prove it by an example. The idea is to reverse the Euclidean algorithm.

Example 2.13. We use the Euclidean algorithm to calculate $\mathrm{hcf}(2681, 931)$.
In the first step we write

$$2681 = 2 \cdot 931 + 819 \tag{2.3}$$

In the second step we write

$$931 = 1 \cdot 819 + 112. \tag{2.4}$$

In the third step we write

$$819 = 7 \cdot 112 + 35 \tag{2.5}$$

In the fourth step we write

$$112 = 3 \cdot 35 + 7. \tag{2.6}$$

In the fifth step we write

$$35 = 5 \cdot 7.$$

Therefore, the Euclidean algorithm tells us that

$$\mathrm{hcf}(2681, 931) = 7.$$

Now we reverse the algorithm. First we rearrange (2.6) to obtain

$$7 = 112 - 3 \cdot 35.$$

Second we use (2.5) to substitute for 35, and obtain

$$\begin{aligned} 7 &= 112 - 3 \cdot (819 - 7 \cdot 112) \\ &= -3 \cdot 819 + 22 \cdot 112. \end{aligned}$$

Third we use (2.4) to substitute for 112, and obtain

$$\begin{aligned} 7 &= -3 \cdot 819 + 22 \cdot (931 - 819) \\ &= 22 \cdot 931 - 25 \cdot 819 \end{aligned}$$

Fourth we use (2.3) to substitute for 112, and obtain

$$\begin{aligned} 7 &= 22 \cdot 931 - 25 \cdot (2681 - 2 \cdot 931) \\ &= -25 \cdot 2681 + 72 \cdot 931. \end{aligned}$$

So we have written $7 = \mathrm{hcf}(2681, 931)$ in the form $x2681 + y931$ with $x, y \in \mathbb{Z}$, where $x = -25$ and $y = 72$.

Now we extract the method used in the example above to prove Theorem 2.12.

Proof of Theorem 2.12. We first note that it suffices to consider $a, b \geq 0$. Also we assume without loss of generality that $a \geq b$, otherwise we can swap a and b.

Let $a_0, a_1, a_2, \ldots, a_k$ and $q_1, q_2, \ldots, q_k$ be the sequences of natural numbers produced by the Euclidean algorithm, so $a_k = \mathrm{hcf}(a, b)$.

For each l we have the equation

$$a_l = a_{l-2} - q_{l-1}a_{l-1},$$

which we refer to as equation (l).

From equation (k) we have

$$h = a_k = a_{k-2} - q_{k-1}a_{k-1}.$$

Now using equation $(k-1)$ we can substitute for a_{k-1} to obtain

$$\begin{aligned} h &= a_{k-2} - q_{k-1}(a_{k-3} - q_{k-2}a_{k-2}) \\ &= -q_{k-1}a_{k-3} + (q_{k-1}q_{k-2} + 1)a_{k-2}. \end{aligned}$$

So we have written h in the form

$$h = x_{k-3}a_{k-3} + y_{k-2}a_{k-2},$$

where $x_{k-3}, y_{k-2} \in \mathbb{Z}$. Next we can use equation $(k-2)$ and substitute for a_{k-2} to write

$$h = x_{k-4}a_{k-4} + y_{k-3}a_{k-3},$$

where $x_{k-4}, y_{k-3} \in \mathbb{Z}$. Continuing in this way we eventually obtain an expression

$$h = x_0a_0 + y_1a_1,$$

where $x_0, y_1 \in \mathbb{Z}$. But by definition $a_0 = a$ and $a_1 = b$, so for $x = x_0$ and $y = y_1$ we have

$$h = xa + yb,$$

with $x, y \in \mathbb{Z}$ as required. □

Now we give an alternative proof of Theorem 2.12, which is in a sense more direct. It is a little bit complicated, so you may want to omit it on a first reading. For the proof we require the fact that a nonempty subset S of $\mathbb{N}$ contains a least element, i.e. there exists $n \in S$ such that $n \leq m$ for all $m \in S$.

Alternative proof of Theorem 2.12. Consider the set

$$S = \{ua + vb \in \mathbb{Z} \,:\, u, v \in \mathbb{Z}\}$$

of all integral linear combinations of a and b.

First we note that $S \cap \mathbb{N}$ is nonempty, as either a or $-a$ is in $S \cap \mathbb{N}$. Therefore, there is a least element of $S \cap \mathbb{N}$, which we denote by $h = xa + yb$, where $x, y \in \mathbb{Z}$. We are going to show that $h = \mathrm{hcf}(a, b)$.

First we show that $h \mid a$. By Theorem 2.4, there exist $q, r \in \mathbb{Z}$ with $a = qh + r$ and $0 \leq r < h$. Then

$$\begin{aligned} r &= a - qh \\ &= a - q(xa + yb) \\ &= (1 - qx)a - qyb \end{aligned}$$

Therefore, $r \in S$, because $1 - qx, -qy \in \mathbb{Z}$. If $r \neq 0$, then $r \in S \cap \mathbb{N}$, which is not possible as $r < h$ and h is the least element of $S \cap \mathbb{N}$. Hence, $r = 0$ and $a = qh$, so $h \mid a$.

Similarly, we can prove that $h \mid b$. Therefore, $h \mid a$ and $h \mid b$, so h is a common factor of a and b.

Next we prove that $h \geq c$ for any common factor c of a and b. Suppose that c is a common factor of a and b. Then $c \mid h$, by Lemma 2.3(a). In particular, this means that $c \leq h$. □

In Examples 2.7(a) we saw that the common factors of 28 and 42 are $\pm 1, \pm 2, \pm 7$ and ± 14, so $\mathrm{hcf}(28, 42) = 14$. So in this case each common factor is a factor of the highest common factor. This statement is true in general; it is a consequence of Theorem 2.12 and is very useful later. We call it a *corollary*, which is a theorem that follows easily from another theorem.

Corollary 2.14. *Let $a, b, c \in \mathbb{N}$ and let $h = \mathrm{hcf}(a, b)$. Suppose that c is a common factor of a and b. Then $c \mid h$.*

Proof. By Theorem 2.12, there exist $x, y \in \mathbb{Z}$ such that $h = xa + yb$. Then $c \mid h$ by Lemma 2.3(a). □

2.6 Coprime integers

Below we define the notion of integers being coprime to each other.

Definition 2.15. Let $a, b \in \mathbb{Z}$. We say that a *is coprime to* b if $\mathrm{hcf}(a, b) = 1$.

Note that saying a is coprime to b is equivalent to saying b is coprime to a. We sometimes say that a *and* b *are coprime to each other* instead of a is coprime to b; and sometimes just a *and* b *are coprime*.

As usual we demonstrate this definition with some examples.

Examples 2.16.

(a) Let $a = 168$ and $b = 205$. Then a is coprime to b.
(b) Let $p, q \in \mathbb{N}$ be distinct primes. Then p is coprime to q.

The following corollary is just Theorem 2.12 for the case of coprime integers.

Corollary 2.17. *Let $a, b \in \mathbb{Z}$. Suppose that a is coprime to b. Then there exist $x, y \in \mathbb{Z}$ such that*

$$1 = xa + yb.$$

2.7 Primes and products

The next theorem is the key result that we require for our proof of the fundamental theorem of arithmetic in the next section.

Theorem 2.18. *Let $a, b \in \mathbb{Z}$ and $p \in \mathbb{N}$ be prime. Suppose that $p \mid ab$. Then $p \mid a$ or $p \mid b$.*

Proof. Let $h = \text{hcf}(p, b)$. Since p is prime, the only positive factors of p are 1 and p. Therefore, h must be either 1 or p. We consider these two cases separately.
Case 1: $h = p$. Then $p \mid b$.
Case 2: $h = 1$. Then by Theorem 2.12 there exist $x, y \in \mathbb{Z}$ such that

$$1 = xp + yb. \tag{2.7}$$

Multiplying (2.7) by a we obtain

$$a = axp + ayb = (ax)p + y(ab).$$

Now $p \mid p$, and $p \mid ab$. Therefore, $p \mid a$ by Lemma 2.3(a).

In both cases we have shown that $p \mid a$ or $p \mid b$, which proves the theorem. □

The corollary below is proved by repeated use of Theorem 2.18.

Corollary 2.19. *Let $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ and let $p \in \mathbb{N}$ be prime. Suppose that $p \mid a_1 a_2 \ldots a_n$. Then $p \mid a_i$ for some $i = 1, 2, \ldots, n$.*

Proof. We have $p \mid (a_1 a_2 \ldots a_{n-1})a_n$ so by Theorem 2.18, either $p \mid a_1 a_2 \ldots a_{n-1}$ or $p \mid a_n$. If $p \mid a_n$, then we are done. Otherwise, using Theorem 2.18 again we see that $p \mid a_1 a_2 \ldots a_{n-2}$ or $p \mid a_{n-1}$. Continuing in this way we will eventually see that $p \mid a_i$ for some $i = 1, 2, \ldots, n$. □

2.8 The fundamental theorem of arithmetic

The purpose of this section is to state and prove the fundamental theorem of arithmetic, see Theorem 2.22. This theorem is really important even though it may not seem very exciting at the moment. In Section 2.9, we'll prove a couple of interesting consequences, and we'll also see its importance again in Chapter 3.

First we give an example showing how we can calculate a prime factorization of a natural number. This example gives the idea for how we prove Proposition 2.21, which says that any natural number can be factorized as a product of primes. A *proposition* is just another name for a theorem that we don't think is important enough to call a theorem.

Example 2.20. We calculate a prime factorization of 8658. First we take out the factor 2 to obtain $8658 = 2 \cdot 4329$. Next we see that 4329 is divisible by 3 and we have $4329 = 3 \cdot 1443$. Now we see that 3 is still a factor of 1443 and we have $1443 = 3 \cdot 481$. Finally, we see that $481 = 13 \cdot 37$, and 13 and 37 are prime. Hence, we obtain the prime factorization

$$8658 = 2 \cdot 3 \cdot 3 \cdot 13 \cdot 37.$$

The idea of the proof of Proposition 2.21 is that if $n \in \mathbb{N}$ is not prime, then we can find a prime factor p of n and then continue by applying the same process to the quotient $\frac{n}{p}$.

Proposition 2.21. *Let $n \in \mathbb{N}$ with $n \geq 2$. Then there exist prime numbers $p_1, p_2, \ldots, p_k$ such that*

$$n = p_1 p_2 \cdots p_k.$$

Proof. If n is prime, then we take $k = 1$ and $p_1 = n$.
So suppose that n is not prime. Let d be a factor of n with $1 < d < n$ and d as small as possible. Then d must be prime, because if c is a factor of d with $1 \leq c < d$, then c is a factor of n that is smaller than d so must be equal to 1.
We set $p_1 = d$ and let $n_2 \in \mathbb{N}$ be such that $n = p_1 n_2$.
If n_2 is prime, then we can take $k = 2$ and $p_2 = n_2$ and we are done.
Otherwise, we can apply the argument above to n_2 in place of n and find a prime number p_2 and a natural number n_3 such that $n_2 = p_2 n_3$. Then $n = p_1 p_2 n_3$.
Continuing in this way, we get a sequence of prime numbers $p_1, p_2, p_3, \ldots$ and natural numbers $n = n_1 > n_2 > n_3 > \ldots$. Eventually, for some $k \in \mathbb{N}$ we must have that n_k is prime. Then we take $p_k = n_k$ and we have

$$n = p_1 p_2 \cdots p_k$$

is a factorization of n as a product of primes. □

At the start of this chapter, we said that the fundamental theorem of arithmetic roughly says that any natural number can be factorized uniquely as a product of primes. We have just proved that a natural number can be factorized as a product of primes, so now we need to work out what it means for this factorization to be unique. A first guess might be the following statement.

Let $n \in \mathbb{N}$. Then:

(a) there exist prime numbers $p_1, p_2, \ldots, p_k$ *such that*

$$n = p_1 p_2 \cdots p_k.$$

(b) if $q_1, q_2, \ldots, q_l$ *are prime numbers such that* $n = q_1 q_2 \cdots q_l$*, then* $l = k$ *and* $q_i = p_i$ *for all* $i = 1, \ldots, k$*.*

If we think about this a little bit, then we can find a problem with this statement. Namely that there is nothing stopping us from reordering the prime factors. For example, consider the case $n = 6$. We have $6 = 2 \cdot 3$ and $6 = 3 \cdot 2$, so we can take $r = s = 2$, $p_1 = 2$, $p_2 = 3$, $q_1 = 3$ and $q_2 = 2$. Then $p_1 \neq q_1$. So $n = 6$ gives counterexample to the statement above.

To deal with this problem we have to make sure we order the prime factors. This is done in our statement of the fundamental theorem of arithmetic below.

Theorem 2.22 (Fundamental theorem of arithmetic)**.** *Let* $n \in \mathbb{N}$ *with* $n > 1$*. Then:*

(a) *there exist prime numbers* $p_1 \leq p_2 \leq \cdots \leq p_k$ *such that*

$$n = p_1 p_2 \cdots p_k.$$

(b) *if* $q_1 \leq q_2 \leq \cdots \leq q_l$ *are prime numbers such that* $n = q_1 q_2 \cdots q_l$*, then*

$$k = l \quad \textit{and} \quad q_i = p_i \textit{ for all } i = 1, 2, \ldots, k.$$

Proof. (a) This is just Proposition 2.21.
(b) We have $p_1 \mid n$ and $n = q_1 q_2 \cdots q_l$, so $p_1 \mid q_i$ for some $i = 1, 2, \ldots, l$ by Corollary 2.19. Since q_i is prime, the only factors of q_i are 1 and q_i, and thus we must have $p_1 = q_i$. In particular, $q_1 \leq p_1$. Similarly, we can show that $q_1 = p_j$ for some $j = 1, 2, \ldots, k$, so in particular $p_1 \leq q_1$. Hence $p_1 = q_1$, and so $p_2 p_3 \cdots p_k = q_2 q_3 \cdots q_l$.

Suppose that $k \leq l$. Continuing to argue as above we can show that

$$p_1 = q_1, \quad p_2 = q_2, \quad \ldots, \quad p_k = q_k.$$

Therefore,

$$n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_k.$$

Now suppose that $k < l$. Then we have $1 = q_{k+1} q_{k+2} \cdots q_l$, which is impossible. Therefore, we must have $k = l$. So we have proved that $l = k$ and $q_i = p_i$ for all $i = 1, 2, \ldots, k$, as required.

If $k \geq l$, then we can prove $l = k$ and $q_i = p_i$ for all $i = 1, 2, \ldots, k$ similarly. □

We note that in a prime factorization of $n \in \mathbb{N}$ given by Theorem 2.22, some of the p_is may be equal. If we collect these equal primes together we get a factorization of n of the form

$$n = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k}$$

where $p_1 < p_2 < \cdots < p_k$ are primes and $s_1, s_2, \ldots, s_k \in \mathbb{N}$. Part (b) of the fundamental theorem of arithmetic, then tells us that if $q_1 < q_2 < \cdots < q_l$ are primes and $t_1, t_2, \ldots, t_l \in \mathbb{N}$ such that $n = q_1^{t_1} q_2^{t_2} \cdots q_l^{t_l}$, then

$$k = l \quad \text{and} \quad q_i = p_i, s_i = t_i \text{ for all } i = 1, \ldots, k.$$

Sometimes it is more convenient for us to use this formulation.

We end this section by remarking that the fundamental theorem of arithmetic is not an obvious statement. You may think that it is, because you have probably believed it for a long time, though never seen a proof. Let me try to convince you that it is really not obvious. It may take a bit of time for this to sink in, and you're encouraged to ask if you doesn't make sense straightaway.

I'll let you know that 1487, 1559, 1789 and 1873 are all prime numbers. Now consider the question: Is $1559 \cdot 1789 = 1487 \cdot 1873$? Suppose that you're not allowed to use the fundamental theorem of arithmetic and you have to answer this – how would you do this? I imagine that you would calculate that $1559 \cdot 1789 = 2789051 \neq 2785151 = 1487 \cdot 1873$.

Now let p_1, p_2, q_1, q_2 be primes with $p_1 \leq p_2$, $q_1 \leq q_2$ and $\{p_1, p_2\} \neq \{q_1, q_2\}$ and consider the question: Is $p_1 p_2 = q_1 q_2$? If you are given specific values, then you would do this by calculating the value of both products – but you can't do this for general p_1, p_2, q_1, q_2 and it seems that you're a bit stuck. So we need to have proved the fundamental theorem of arithmetic to know that $p_1 p_2 \neq q_1 q_2$.

You can also look at the exercise Q2.13 to see another reason why the fundamental theorem of arithmetic is not obvious.

2.9 Some consequences of the fundamental theorem

As mentioned at the start of Section 2.8 we are going to demonstrate the importance of Theorem 2.22 with a couple of nice consequences.

Our first consequence of Theorem 2.22 is Theorem 2.24, which says that square root of a natural number that is not a perfect square is irrational. First we prove a special case of this, which gives us an idea how to prove the general theorem. For the statement, we recall that $x \in \mathbb{R}$ is *irrational* if $x \notin \mathbb{Q}$.

Proposition 2.23. *$\sqrt{2}$ is irrational.*

Proof. Suppose for a contradiction that $\sqrt{2} \in \mathbb{Q}$. Then there exists $a, b \in \mathbb{N}$ such that

$$\sqrt{2} = \frac{a}{b}.$$

By Theorem 2.22 there are factorizations

$$a = q_1^{t_1} q_2^{t_2} \cdots q_l^{t_l}$$

and

$$b = r_1^{u_1} r_2^{u_2} \cdots r_m^{u_m}$$

where $q_1 < q_2 < \cdots < q_l, r_1 < r_2 < \cdots < r_m$ are primes and $t_1, t_2, \ldots, t_l, u_1, u_2, \ldots, u_m \in \mathbb{N}$. By cancelling common factors we assume that $q_i \neq r_j$ for any $i = 1, 2, \ldots, l$ and $j = 1, 2, \ldots, m$.

We have

$$2 = \frac{a^2}{b^2} \quad \text{so} \quad 2b^2 = a^2.$$

Therefore,

$$2 r_1^{2u_1} r_2^{2u_2} \cdots r_m^{2u_m} = q_1^{2t_1} q_2^{2t_2}, \cdots q_l^{2t_l}.$$

which gives two prime factorizations of $2b^2$.

Suppose $m \geq 1$. By Theorem 2.22, we have $r_1 = q_j$ for some $j = 1, 2, \ldots, m$. But we assumed that $q_i \neq r_j$ for any $i = 1, 2, \ldots, l$, $j = 1, 2, \ldots, m$. Thus we must have $m = 0$, which means that $b = 1$, so that $2 = a^2$. But then a must be even and 4 is a factor of a^2. This is not possible so we have the required contradiction. □

For the statement of Theorem 2.24, we recall that $n \in \mathbb{N}$ is a *perfect square* if there exists $m \in \mathbb{N}$ such that $n = m^2$. For the proof we just need to add some extra bits to the proof of Proposition 2.23.

Theorem 2.24. *Let $n \in \mathbb{N}$. Suppose that n is not a perfect square. Then $\sqrt{n}$ is irrational.*

Proof. Suppose for a contradiction that $\sqrt{n}$ is rational. Then

$$\sqrt{n} = \frac{a}{b}$$

for some $a, b \in \mathbb{N}$.

By Theorem 2.22 there are factorizations

$$n = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k},$$

$$a = q_1^{t_1} q_2^{t_2} \cdots q_l^{t_l}$$

and

$$b = r_1^{u_1} r_2^{u_2} \cdots r_m^{u_m}$$

where $p_1 < p_2 < \cdots < p_k, q_1 < q_2 < \cdots < q_l, r_1 < r_2 < \cdots < r_m$ are primes and $s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_l, u_1, u_2, \ldots, u_m \in \mathbb{N}$. By cancelling common factors we assume that $q_i \neq r_j$ for any $i = 1, 2, \ldots, l$ and $j = 1, 2, \ldots, m$.

We have

$$n = \frac{a^2}{b^2} \quad \text{so} \quad nb^2 = a^2.$$

Therefore,

$$p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k} r_1^{2u_1} r_2^{2u_2} \cdots r_m^{2u_m} = q_1^{2t_1} q_2^{2t_2} \cdots q_l^{2t_l},$$

which gives two prime factorizations of nb^2.

Suppose $m \geq 1$. By Theorem 2.22, we have $r_1 = q_j$ for some $j = 1, 2, \ldots, m$. But we know that $q_i \neq r_j$ for any $i = 1, 2, \ldots, l$, $j = 1, 2, \ldots, m$. Thus we must have $m = 0$, which means that $b = 1$, so that $n = a^2$ is a perfect square. But we know this is not the case, so we have the required contradiction. □

An alternative proof of Proposition 2.23 is given below. In a sense this proof is more elementary but it does not generalize as easily.

Alternative proof of Proposition 2.23. Suppose that $\sqrt{2} \in \mathbb{Q}$. Then there exists $a \in \mathbb{Z}$ and $b \in \mathbb{N}$ such that

$$\sqrt{2} = \frac{a}{b}.$$

If a and b have any common factors, then we can cancel them. So we can assume that a and b have no common factors.

Now

$$2 = \frac{a^2}{b^2} \quad \text{so} \quad 2b^2 = a^2.$$

Therefore, a^2 is even, which in turn means that a must be even. So we can write $a = 2c$ for some $c \in \mathbb{Z}$. From this we see that

$$2 = \frac{a^2}{b^2} = \frac{4c^2}{b^2} \quad \text{so} \quad 2 = \frac{b^2}{c^2}.$$

Arguing exactly as before, we see that b must be even. But this means that 2 is a factor of both a and b, and we assumed that a and b do not have any common factors, which is a contradiction. □

Now we give our second consequence of Theorem 2.22. For the statement we require the definition of a perfect nth power for $n \in \mathbb{N}$, which generalizes that of a perfect square. We say that $a \in \mathbb{N}$ is a *perfect nth power* if there exists $b \in \mathbb{N}$ such that $a = b^n$.

Theorem 2.25. *Let $a, b, n \in \mathbb{N}$. Suppose that a is coprime to b and ab is a perfect nth power. Then both a and b are perfect nth powers.*

Proof. Let $c \in \mathbb{N}$ be such that $ab = c^n$. By Theorem 2.22, we have factorizations

$$c = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k},$$

$$a = q_1^{t_1} q_2^{t_2} \cdots q_l^{t_l}$$

and

$$b = r_1^{u_1} r_2^{u_2} \cdots r_m^{u_m}$$

where $p_1 < p_2 < \cdots < p_k, q_1 < q_2 < \cdots < q_l, r_1 < r_2 < \cdots < r_m$ are primes and $s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_l, u_1, u_2, \ldots, u_m \in \mathbb{N}$. Since a is coprime to b we have that $q_i \neq r_j$ for any $i = 1, 2, \ldots, l$ and $j = 1, 2, \ldots, m$.

The equation $ab = c^n$ gives

$$q_1^{t_1} q_2^{t_2} \cdots q_l^{t_l} r_1^{u_1} r_2^{u_2} \cdots r_m^{u_m} = p_1^{ns_1} p_2^{ns_2} \cdots p_k^{ns_k}.$$

By Theorem 2.22, each q_i is equal to some p_j, and the corresponding powers t_i and ns_j must be equal. Similarly, each r_i is equal to some p_j and then u_i is equal to ns_j. We conclude that n is a factor of each of the powers t_i and u_i, say $t_i = nv_i$ and $u_i = nw_i$. Therefore,

$$a = (q_1^{v_1} q_2^{v_2} \cdots q_l^{v_l})^n$$

and

$$b = (r_1^{w_1} r_2^{w_2} \cdots r_m^{w_m})^n$$

are perfect nth powers. □

Theorem 2.25 may not seem that exciting, but in the example below, we give a quite spectacular consequence.

Example 2.26. We are going to show that

There is no nonzero even square that is one more than a cube.

In other words we are going to show that the equation

$$4x^2 = y^3 + 1 \tag{2.8}$$

has no solutions with $x, y \in \mathbb{Z}$ and $x \neq 0$.

First we rewrite the equation as $y^3 = 4x^2 - 1$, and then factorize to get

$$y^3 = (2x + 1)(2x - 1).$$

Both $2x+1$ and $2x-1$ are both odd, so $\mathrm{hcf}(2x+1, 2x-1)$ is odd. Also, by Lemma 2.3(a), $\mathrm{hcf}(2x+1, 2x-1)$ is a factor of $2 = (2x+1) - (2x-1)$. It follows that $\mathrm{hcf}(2x+1, 2x-1)$ must be equal to 1, so $2x + 1$ is coprime to $2x - 1$.

Now using Theorem 2.25, we see that both $2x + 1$ and $2x - 1$ are perfect cubes. However, from the list of cubes

$$0, \pm 1, \pm 8, \pm 27, \pm 64, \pm 125, \ldots$$

we see that the only cubes that differ by 2 are 1 and -1. Therefore, we must have $x = 0$, which shows that (2.8) has no solutions with $x, y \in \mathbb{Z}$ and $x \neq 0$.

The equation (2.8) is an example of a *Diophantine equation*. In general a Diophantine equation is a polynomial equation in which the solutions are required to be integers. Solving Diophantine equations is a fascinating area of mathematics, and in general they are very difficult to solve. For example, the Diophantine equation

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

has not been completely solved for all values of $k \in \mathbb{N}$. A particularly famous example of a Diophantine equation is

$$x^n + y^n = z^n,$$

for $n \in \mathbb{N}$, which is the subject of *Fermat's last theorem*. It was proved by Andrew Wiles in 1995 that it has no nonzero integer solutions for $n \geq 3$.

2.10 Summary of Chapter 2

By the end of this chapter you should be able to:

- prove elementary lemmas about factors;
- state and apply the division theorem;
- define common factors, highest common factors and coprime integers;
- state and apply Bézout's lemma that "there exist $x, y \in \mathbb{Z}$ such that $\text{hcf}(a, b) = xa + yb$";
- apply the Euclidean algorithm to find the highest common factor of $a, b \in \mathbb{Z}$, and reverse it to find $x, y \in \mathbb{Z}$ such that $\text{hcf}(a, b) = xa + yb$;
- state, prove and apply the theorem that "if $p \mid ab$, then $p \mid a$ or $p \mid b$"; and
- state and apply the fundamental theorem of arithmetic.

In these learning aims I have given short informal statements of some theorems in order to keep the aims fairly short. If you are ever asked to state these theorems you should always give the full statement from earlier in the notes.

2.11 Exercises for Chapter 2

As mentioned for the Exercises for Chapter 1, some of the exercises are quite challenging, and it is not expected that you will be able to do them all straightaway.
You should not use the fundamental theorem of arithmetic in all questions up to Q2.13, but you should use it in the later questions.

Q2.1. Prove Lemma 2.3(a):

Lemma. *Let $a, b, c, k, l \in \mathbb{Z}$. Suppose that $a \mid b$ and $a \mid c$. Then $a \mid (kb + lc)$.*

Q2.2. Prove Lemma 2.3(b):

Lemma. *Let $a, b, c \in \mathbb{Z}$. Suppose that $a \mid b$ and $b \mid c$. Then $a \mid c$.*

Q2.3. Prove Lemma 2.3(c):

Lemma. *Let $a, b \in \mathbb{Z}$. Suppose that $a \mid b$ and $b \mid a$. Then $a = \pm b$.*

Q2.4. Let $n \in \mathbb{N}$ with $n > 1$. Suppose that n is a factor of $(n-1)! + 1$. Prove that n is prime.

Q2.5. (a) Use the Euclidean algorithm to find $\mathrm{hcf}(931, 210)$.
(b) Use your working to find $x, y \in \mathbb{Z}$ such that

$$\mathrm{hcf}(931, 210) = 931x + 210y.$$

Q2.6. Let $a, b, c \in \mathbb{N}$. Suppose that a is coprime to b and $a \mid bc$. Prove that $a \mid c$.
Hint: We have seen a theorem, which is similar and you should be able to adapt part of its proof.

Q2.7. Let $a, b, c \in \mathbb{Z}$. Suppose that a is coprime to b, and that $a \mid c$ and $b \mid c$. Prove that $ab \mid c$.

Q2.8. Let $a, b, c \in \mathbb{Z}$. Suppose that a is coprime to c and b is coprime to c. Prove that ab is coprime to c.

Q2.9. The Fibonacci sequence is the sequence

$$f_0, f_1, f_2, f_3, \ldots$$

defined by

- $f_0 = 1$ and $f_1 = 1$; and
- $f_{n+2} = f_n + f_{n+1}$ for $n \geq 0$.

So the sequence starts

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

Investigate the highest common factor of consecutive elements of the Fibonacci sequence.

First work out $\mathrm{hcf}(f_n, f_{n+1})$ *for* $n \leq 6$*, then make a conjecture, then prove it.*

Q 2.10. Determine whether each of the following statements is true and justify your answer.

(a) Let $a, b, h \in \mathbb{N}$. Suppose that there exist $x, y \in \mathbb{Z}$ such that $h = xa + yb$. Then $h = \mathrm{hcf}(a, b)$.
(b) Let $a, b \in \mathbb{Z}$. Suppose that there exist $x, y \in \mathbb{Z}$ such that $1 = xa + yb$. Then a is coprime to b.
(c) Let $a, b, c \in \mathbb{Z}$. Suppose that $a \nmid b$ and $a \nmid c$. Then $a \nmid b + c$.

When you are asked to justify your answer it means you have to prove it if it is true and give a counterexample if it is not true.

In the next exercise and also in Q2.17, we use the definition of the least common multiple of two natural numbers, which we give next.

Definition. Let $a, b \in \mathbb{N}$.

(a) A *common multiple* of a and b is an integer m such that $a \mid m$ and $b \mid m$.
(b) The *least common multiple* of a and b is the smallest $l \in \mathbb{N}$ that is a common multiple of a and b. We write $l = \mathrm{lcm}(a, b)$.

Q2.11. Let $a, b \in \mathbb{N}$. Prove that $\mathrm{lcm}(a, b) = \frac{ab}{\mathrm{hcf}(a,b)}$.

Q2.12. For $n \in \mathbb{N}$, let $P(n)$ be the statement that each $m \in \mathbb{N}$ with $2 \leq m \leq n$ can be written as a product of prime numbers. Use this statement to give an alternative proof of Proposition 2.23 using the principle of mathematical induction.

Q2.13. Let $\mathbb{E} \subseteq \mathbb{N}$ be the set of even natural numbers. We say that $n \in \mathbb{E}$ is *prima* if n cannot be expressed in the form $n = ab$, where $a, b \in \mathbb{E}$.

(a) Show that 6 is prima, but 4 is not prima.
(b) Write down the general form of a prima in $\mathbb{E}$.
(c) Give a counterexample to the statement below, i.e. find an element of $\mathbb{E}$ that has two different prima factorizations.
Let $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_l$ be prima elements of $\mathbb{E}$. Suppose that

$$p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l.$$

Then

$$l = k \quad \textit{and} \quad p_i = q_i \quad \textit{for all} \quad i = 1, 2, \ldots, k.$$

Q2.14. Prove that the cube root of 2 is irrational.

Q2.15. Find all solutions $x, y \in \mathbb{Z}$ to the following Diophantine equations.

(a) $x^2 - x = y^3$
(b) $x^4 = 9y^2 + 3y - 2$

Hint: In (a) first factorize the left hand side, and in (b) first factorize the right hand side.

Q2.16. Let $n \in \mathbb{N}$. Suppose that n is a perfect square and that n is a perfect cube. Prove that n is a perfect 6th power.

Q2.17. Let $a, b \in \mathbb{N}$ with prime factorization.

$$a = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k},$$

and

$$b = p_1^{t_1} p_2^{t_2} \cdots p_k^{t_k}$$

where $p_1 < p_2 < \cdots < p_k$ are primes and $s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_l \in \mathbb{N}_0$ (we allow some of the s_i and t_i to be 0). Prove that

(a) $\mathrm{hcf}(a, b) = p_1^{\min(s_1,t_1)} p_2^{\min(s_2,t_2)} \cdots p_k^{\min(s_k,t_k)}$.
(b) $\mathrm{lcm}(a, b) = p_1^{\max(s_1,t_1)} p_2^{\max(s_2,t_2)} \cdots p_k^{\max(s_k,t_k)}$.

Deduce that $\mathrm{lcm}(a, b) = \frac{ab}{\mathrm{hcf}(a,b)}$.

Chapter 3

Modular arithmetic

In this chapter we introduce the notion of congruence of integers modulo a fixed natural number, and use this to develop the theory of modular arithmetic. This is an important area of algebra, which is a very useful method for studying the integers. Later in the chapter, we give a particularly striking application of to the theory of cryptography, which we depend on all the time, when making financial transactions on the internet.

3.1 Congruence modulo n

We start by giving the main definition for this chapter.

Definition 3.1. Let $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. We write

$$a \equiv b \bmod n$$

and say that *a is congruent to b modulo n* if

$$n \mid a - b.$$

We write $a \not\equiv b \bmod n$ if a is not congruent to b modulo n.

Note that $a \equiv b \bmod n$ is equivalent to saying that there exists $x \in \mathbb{Z}$ such that

$$a = b + nx.$$

As usual some examples will help us to understand the definition.

Examples 3.2.

(a) $43 \equiv 7 \bmod 9$, because $9 \mid 36 = 43 - 7$.
(b) $11 \equiv -28 \bmod 13$, because $13 \mid 39 = 11 - (-28)$.
(c) Let $a \in \mathbb{Z}$. Then:
  - a is even if and only if $a \equiv 0 \bmod 2$; and
  - a is odd if and only if $a \equiv 1 \bmod 2$.
(d) It is 4pm now, so in 269 hours it will be 9pm.
This is because $269 \equiv 5 \bmod 24$.

(e) It is Tuesday today, so in 100 days time it will be Thursday.
This is because $100 \equiv 2 \bmod 7$ and Thursday is two days after Tuesday.

The examples (c), (d) and (e) show that we are already familiar with certain cases of congruences.

We begin our study of congruences by relating them to remainders in the following lemma and corollary.

Lemma 3.3. *Let $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. Then $a \equiv b \bmod n$ if and only if a and b leave the same remainder when divided by n.*

Before we give the proof, we explain the statement briefly. The phrase "if and only if" means that we have to prove two things:

- if $a \equiv b \bmod n$, then a and b leave the same remainder when divided by n; and
- if a and b leave the same remainder when divided by n, then $a \equiv b \bmod n$.

Proof. Using Theorem 2.4, we can write $a = qn + r$ and $b = q'n + r'$, where $q, q', r, r' \in \mathbb{Z}$ and $0 \leq r, r' < n$.

Suppose that $a \equiv b \bmod n$. Then $n \mid a - b$, so by Lemma 2.3(a)

$$n \mid (a - b) - (q - q')n = r - r'.$$

Also $-n < r - r' < n$. This forces $r - r' = 0$, so that $r = r'$, which says that a and b leave the same remainder when divided by n.

Now suppose that a and b leave the same remainder when divided by n. Then $r = r'$, and $a - b = (q - q')n$. Thus $n \mid a - b$ and $a \equiv b \bmod n$. □

Corollary 3.4. *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then there exists unique $b \in \mathbb{Z}$ with $0 \leq b < n$ such that $a \equiv b \bmod n$.*

Proof. By Lemma 3.3, we can take b to be the remainder when a is divided by n. □

Next we give some elementary properties of congruences.

Lemma 3.5. *Let $n \in \mathbb{N}$ and $a, b, c \in \mathbb{Z}$. Then:*

(a) $a \equiv a \bmod n$; *(Reflexive property)*
(b) *if* $a \equiv b \bmod n$, *then* $b \equiv a \bmod n$; *and* *(Symmetric property)*
(c) *if* $a \equiv b \bmod n$ *and* $b \equiv c \bmod n$, *then* $a \equiv c \bmod n$. *(Transitive property)*

Proof. (a) We have $n \mid 0 = a - a$, so $a \equiv a \bmod n$.
(b) Since, $a \equiv b \bmod n$ we have $n \mid a - b$. Then $n \mid b - a = -(a - b)$. Hence, $b \equiv a \bmod n$.
(c) Since, $a \equiv b \bmod n$ and $b \equiv c \bmod n$ we have $n \mid a - b$ and $n \mid b - c$. Then $n \mid a - c = (a - b) + (b - c)$. Hence, $a \equiv c \bmod n$. □

On the right in the statement of Lemma 3.5, we have given names to the properties satisfied by congruence modulo n. These are the properties required for a relation to be an equivalence relation – you learnt about equivalence relations in 1ACa. We can neatly summarize this is the corollary below.

Corollary 3.6. *Let $\sim$ be the relation on $\mathbb{Z}$ defined by $a \sim b$ means $a \equiv b \bmod n$. Then $\sim$ is an equivalence relation.*

We will use the theory of equivalence relations later in this chapter. There is a recap on equivalence relations in Appendix A.

3.2 Arithmetic with congruences

In Lemma 3.7 below we show that congruence interacts well with arithmetic operations. The proof of (b) is exercise Q3.2.

Lemma 3.7. *Let $n \in \mathbb{N}$ and $a, b, a', b' \in \mathbb{Z}$. Suppose that $a \equiv b \bmod n$ and $a' \equiv b' \bmod n$. Then:*

(a) $a + a' \equiv b + b' \bmod n$; *and*

(b) $aa' \equiv bb' \bmod n$.

Proof. (a) Since $a \equiv b \bmod n$ and $a' \equiv b' \bmod n$, there exist $x, x' \in \mathbb{Z}$ such that

$$a = b + nx \tag{3.1}$$

and

$$a' = b' + nx'. \tag{3.2}$$

Adding (3.1) and (3.2) gives

$$a + a' = b + b' + n(x + x').$$

We have $x + x' \in \mathbb{Z}$, because $x, x' \in \mathbb{Z}$. Therefore, $a + a' \equiv b + b' \bmod n$.

(b) is exercise Q3.2. □

The properties given Lemmas 3.5 and 3.7 allow us to manipulate expressions with congruences in a similar way to how we manipulate expressions with equals signs, as we'll see when we work with them.

We can use Lemma 3.7 to work out remainders when we do divisions of large numbers, as demonstrated in the next example.

Example 3.8. We are going to find the remainder when $107 \cdot 122 + 73$ is divided by 11. So by Lemma 3.3 we have to find $r \in \mathbb{Z}$ with $0 \leq r < 11$ such that $107 \cdot 122 + 73 \equiv r \bmod 11$.

First we see that $107 \equiv 8 \bmod 11$, and $122 \equiv 1 \bmod 11$. Therefore, by Lemma 3.7(b),

$$\begin{aligned} 107 \cdot 122 &\equiv 8 \cdot 1 \bmod 11 \\ &\equiv 8 \bmod 11. \end{aligned}$$

Next we see that $73 \equiv 7 \bmod 11$, so, by Lemma 3.7(a),

$$\begin{aligned} 107 \cdot 254 + 73 &\equiv 8 + 7 \bmod 11 \\ &\equiv 15 \bmod 11 \\ &\equiv 4 \bmod 11. \end{aligned}$$

Therefore, the remainder when $107 \cdot 122 + 73$ is divided by 11 is 4.

Note that we could have also worked this out by first calculating $107 \cdot 122 + 73 = 13127$ and then working out the remainder when 13127 is divided by 11, but this would have been much more work. Actually with a calculator we can do this very quickly. However, if we want to multiply very large numbers together and work out remainders, then it is infeasible to use a calculator in this way. We will see in Section 3.9 that on occasions such calculations need to be carried out.

The following lemma allows us to "take powers of congruences". It is proved by repeatedly using Lemma 3.7(b).

Lemma 3.9. *Let $m, n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. Suppose that $a \equiv b \bmod n$. Then*

$$a^m \equiv b^m \bmod n.$$

Proof. Using Lemma 3.7(b), for the case $a' = a$ and $b' = b$ we obtain $a^2 \equiv b^2 \bmod n$.

Now we can apply Lemma 3.7(b), for the case $a' = a^2$ and $b' = b^2$, and we obtain $a^3 \equiv b^3 \bmod n$.

Continuing in this way we eventually get $a^m \equiv b^m \bmod n$. □

We now use Lemma 3.9 to do a couple of calculations.

Examples 3.10. (a) We are going to find the remainder when 14^{24} is divided by 9. First we note that $14 \equiv 5 \bmod 9$, so $14^{24} \equiv 5^{24} \bmod 9$, by Lemma 3.9. Next we calculate

$$\begin{aligned} 5^2 &\equiv 25 \bmod 9 \\ &\equiv 7 \bmod 9 \end{aligned}$$

$$\begin{aligned} 5^4 &\equiv 7^2 \bmod 9 \\ &\equiv 49 \bmod 9 \\ &\equiv 4 \bmod 9 \end{aligned}$$

$$\begin{aligned} 5^8 &\equiv 4^2 \bmod 9 \\ &\equiv 16 \bmod 9 \\ &\equiv 7 \bmod 9 \end{aligned}$$

$$\begin{aligned} 5^{16} &\equiv 7^2 \bmod 9 \\ &\equiv 4 \bmod 9. \end{aligned}$$

Therefore, using Lemma 3.7, we get

$$\begin{aligned} 5^{24} &\equiv 5^{16} \cdot 5^8 \bmod 9 \\ &\equiv 7 \cdot 4 \bmod 9 \\ &\equiv 28 \bmod 9 \\ &\equiv 1 \bmod 9 \end{aligned}$$

Thus $14^{24} \equiv 1 \bmod 9$, so the remainder when 14^{24} is divided by 9 is 1.

(b) We are going to find the remainder when 27^{67} is divided by 7. We start with the congruence $27 \equiv -1 \bmod 7$. Then we can calculate

$$\begin{aligned} 27^{67} &\equiv (-1)^{67} \bmod 7 \\ &\equiv -1 \bmod 7 \\ &\equiv 6 \bmod 7. \end{aligned}$$

So the remainder is 6. The trick here is to use the negative number -1, it would have been more work if we had started by writing $27 \equiv 6 \bmod 7$.

In the examples below we some nice applications of congruences. First we give a test to see whether an integer can be a perfect square, and then we give an easy way to determine if an integer is divisible by 3.

Examples 3.11. (a) We are going to show that 59778 is not a perfect square.

Let $a \in \mathbb{Z}$. By Corollary 3.4, there exists $b \in \{0, 1, 2, \ldots, 9\}$ such that $a \equiv b \bmod 10$. Then by Lemma 3.9, we have $a^2 \equiv b^2 \bmod 10$. Now we can make the following table, where the third row gives $c \in \{0, 1, 2, \ldots, 9\}$ such that $b^2 \equiv c \bmod 10$, so that we have $a^2 \equiv c \bmod 10$.

b	0	1	2	3	4	5	6	7	8	9
b^2	0	1	4	9	16	25	36	49	64	81
c	0	1	4	9	6	5	6	9	4	1

Therefore, a^2 is congruent to one of 0, 1, 4, 5, 6 or 9 modulo 10. Since $59778 \equiv 8 \bmod 10$ it cannot be a perfect square.

In fact we have shown that any integer whose last digit is 2, 3, 7 or 8 is not a perfect square.

(b) Let $a \in \mathbb{N}$ with digits $a_r a_{r-1} \ldots a_2 a_1 a_0$. So

$$a = a_0 + 10a_1 + 10^2 a_2 + \cdots + 10^{r-1} a_{r-1} + 10^r a_r.$$

We are going to show that $3 \mid a$ if and only if $3 \mid a_0 + a_1 + a_2 + \cdots + a_{r-1} + a_r$.

First we see that $10 \equiv 1 \bmod 3$, so by Lemma 3.9 we have $10^s \equiv 1 \bmod 3$ for all $s \in \mathbb{N}$. Therefore, using Lemma 3.7, we get

$$a \equiv a_0 + a_1 + a_2 + \cdots + a_{r-1} + a_r \bmod 3.$$

We have $3 \mid a$ if and only if $a \equiv 0 \bmod 3$. Thus $3 \mid a$ if and only if

$$a_0 + a_1 + a_2 + \cdots + a_{r-1} + a_r \equiv 0 \bmod 3.$$

if and only if

$$3 \mid a_0 + a_1 + a_2 + \cdots + a_{r-1} + a_r.$$

3.3 Linear congruence equations

A *linear congruence equation* is an equation of the form

$$ax \equiv b \bmod n$$

where $n \in \mathbb{N}$, $a, b \in \mathbb{Z}$ and we are trying to solve for x. We begin by looking at some examples.

Examples 3.12. (a) We are going to find all $x \in \mathbb{Z}$ such that

$$3x \equiv 6 \bmod 12. \tag{3.3}$$

If $x \equiv 2 \bmod 12$, then x is a solution to (3.3). We need to check whether there are anymore solutions.

To do this we can we use the fact that any $x \in \mathbb{Z}$ is congruent modulo 12 to an element of the set $\{0, 1, 2, \ldots, 11\}$, so it suffices to consider only elements of this set. Then we can form the following table, where the last row gives $y \in \{0, 1, 2, \ldots, 11\}$ such that $3x \equiv y \bmod 12$.

x	0	1	2	3	4	5	6	7	8	9	10	11
$3x$	0	3	6	9	12	15	18	21	24	27	30	33
y	0	3	6	9	0	3	6	9	0	3	6	9

Thus $x = 2$, $x = 6$ and $x = 10$ are solutions to (3.3).

Hence, the solutions to (3.3) are given by

$$x \equiv 2 \bmod 12, \quad x \equiv 6 \bmod 12, \quad \text{or} \quad x \equiv 10 \bmod 12.$$

This shows that we cannot just cancel the 3 in (3.3).

We note that an alternative way to approach this example is to say that $3x \equiv 6 \bmod 12$ if an only if there exists $k \in \mathbb{Z}$ such that $3x = 6 + 12k$. Now we can divide by 3 to say that this occurs if and only if $x = 2 + 4k$. Therefore, the solutions of $3x \equiv 6 \bmod 12$ are given by $x \equiv 2 \mod 4$, which is the same as $x \equiv 2 \bmod 12$, $x \equiv 6 \bmod 12$ or $x \equiv 10 \bmod 12$. Don't worry if you don't understand this alternative method straightaway, as the first method is fine to use.

(b) We are going to find all $x \in \mathbb{Z}$ such that

$$2x \equiv 8 \bmod 9. \tag{3.4}$$

If $x \equiv 4 \bmod 9$, then x is a solution to (3.4). Again, we need to check whether there are anymore solutions.

To do this we can we use the fact that any $x \in \mathbb{Z}$ is congruent modulo 9 to an element of the set $\{0, 1, 2, \ldots, 8\}$, so it suffices to consider only elements of this set. Then we can form the following table, where the last row gives $y \in \{0, 1, 2, \ldots, 8\}$ such that $2x \equiv y \bmod 9$.

x	0	1	2	3	4	5	6	7	8
$2x$	0	2	4	6	8	10	12	14	16
y	0	2	4	6	8	1	3	5	7

Hence, the solutions to (3.4) are given by $x \equiv 4 \bmod 9$. So in this case we can cancel the 2 in (3.4).

The key difference to notice in the examples is that we can make a cancellation when a is coprime to n. In Corollary 3.14, we prove that this is the case in general, and give a condition for a linear congruence equation to have a unique solution modulo n. First we prove the following important theorem about when we can find "multiplicative inverses" for congruences.

Theorem 3.13. *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. Suppose that a is coprime to n. Then there exists $z \in \mathbb{Z}$ such that*

$$az \equiv 1 \bmod n.$$

Proof. Since a is coprime to n, there exist $z, y \in \mathbb{Z}$ such that

$$1 = az + ny,$$

by Corollary 2.17. This equation tells us that

$$az \equiv 1 \bmod n. \qquad \square$$

The statement of the next corollary may look a little complicated to start with, but all it is saying that the linear congruence equation (3.5) has "a unique solution modulo n".

Corollary 3.14. *Let $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. Suppose that a is coprime to n. Consider the linear congruence equation*

$$ax \equiv b \bmod n. \tag{3.5}$$

(a) *There is a solution $x = s \in \mathbb{Z}$ (3.5).*
(b) *Let $r \in \mathbb{Z}$. Then $x = r$ is a solution of (3.5) if and only if $r \equiv s \bmod n$.*

Hence, the solutions of (3.5) are given by $x \equiv s \bmod n$

Proof. Since a is coprime to n, there exists $z \in \mathbb{Z}$ such that

$$az \equiv 1 \bmod n,$$

by Theorem 3.13. Let $s \in \mathbb{Z}$ with $s \equiv zb \bmod n$. Then

$$\begin{aligned} as &\equiv a(zb) \bmod n \\ &\equiv (az)b \bmod n \\ &\equiv b \bmod n. \end{aligned}$$

Therefore, $x = s$ is a solution of (3.5), which proves (a).

Let $r \in \mathbb{Z}$. Suppose that $r \equiv s \bmod n$. Then $ar \equiv as \bmod n$, so $ar \equiv b \bmod n$ and $x = r$ is a solution of (3.5).

Now suppose $x = r$ is a solution of (3.5). Then

$$\begin{aligned} r &\equiv (az)r \bmod n \\ &\equiv z(ar) \bmod n \\ &\equiv zb \bmod n \\ &\equiv s \bmod n. \end{aligned}$$

This proves (b). $\square$

Summing up Corollary 3.14 tells us that we can solve the linear congruence equations (3.5) for a unique x modulo n when a is coprime to n. If a is not coprime to n, then we may have more than one solution as in Examples 3.12(a). In exercise Q3.9, we see that linear congruence equations sometime have no solutions, and we will give a necessary and sufficient condition for a solution to exist.

The proof of Corollary 3.14 gives us a method to solve linear congruence equations of the form (3.5), when a is coprime to n. We demonstrate this method in the next example, and there are some more for you to try in exercise Q3.8.

Example 3.15. Consider the linear congruence equation

$$5x \equiv 7 \bmod 11. \tag{3.6}$$

To solve this equation we look for $z \in \mathbb{Z}$ such that $5z \equiv 1 \bmod 11$. To do this we just try all values of $z \in \{0, 1, 2, \ldots, 10\}$ and we find $z = 9$ does the job, because $5 \cdot 9 = 45 \equiv 1 \bmod 11$. (We could also use the Euclidean algorithm to find z, and note that when we are looking at larger numbers this would be a much more efficient way.) Therefore, we can multiply (3.6) by 9 to obtain

$$\begin{aligned} 45x &\equiv 63 \bmod 11 \\ x &\equiv 8 \bmod 11. \end{aligned}$$

This gives the solutions to (3.6).

We end this section with a corollary about cancelling in congruences, which will be useful later.

Corollary 3.16. *Let $n \in \mathbb{N}$ and $a, b, c \in \mathbb{Z}$. Suppose that c is coprime to n and*

$$ac \equiv bc \bmod n.$$

Then

$$a \equiv b \bmod n.$$

Proof. Since c is coprime to n, there exists $z \in \mathbb{Z}$ such that

$$cz \equiv 1 \bmod n$$

by Theorem 3.13. Then, using Lemma 3.7, we have

$$\begin{aligned} a &\equiv acz \bmod n \\ &\equiv bcz \bmod n \\ &\equiv b \bmod n. \end{aligned}$$

Therefore,

$$a \equiv b \bmod n. \qquad \square$$

3.4 Systems of simultaneous congruences and the Chinese remainder theorem

Think of a natural number x less than 30? Work out

- the remainder a when x is divided by 2;
- the remainder b when x is divided by 3; and
- the remainder c when x is divided by 5.

It may seem surprising at first that we can determine x uniquely from a, b and c. This is sometimes called the "30 riddle", and is based on an ancient Chinese puzzle. We will see a case of this in Examples 3.20(a) below.

More generally, in this section, we look at the theory of systems of simultaneous congruences. The important result is the Chinese remainder theorem, which is Theorem 3.19. As a special case, this explains why x is uniquely determined by a, b and c, as above.

We'll need the following lemma later on, so we state it now for convenience; it was already proved in exercises Q2.7 and Q2.8.

Lemma 3.17. *Let $a, b, c \in \mathbb{Z}$.*

(a) *Suppose that a is coprime to b, and that $a \mid c$ and $b \mid c$. Then $ab \mid c$.*
(b) *Suppose that a is coprime to c and that b is coprime to c. Then ab is coprime to c.*

We consider a pair of simultaneous congruences

$$\begin{aligned} x &\equiv a \bmod n \\ x &\equiv b \bmod m, \end{aligned} \tag{3.7}$$

where $a, b \in \mathbb{Z}$, $n, m \in \mathbb{N}$ and we are trying to solve for x. We look at some examples.

Examples 3.18. (a) We are going to look for $x \in \mathbb{Z}$ such that

$$\begin{aligned} x &\equiv 1 \bmod 3 \\ x &\equiv 2 \bmod 5. \end{aligned} \tag{3.8}$$

To do this we first list all integers x with $0 \leq x \leq 14$ and $x \equiv 1 \bmod 3$. These are:

$$1, 4, 7, 10, 13$$

Next we list x with $0 \leq x \leq 14$ and $x \equiv 2 \bmod 5$. These are:

$$2, 7, 12.$$

We observe that 7 is the only number on both lists, so $x = 7$ is a solution to the simultaneous congruences above.

Now let $y \in \mathbb{Z}$ and let z be the unique element of $\{0, 1, \ldots, 14\}$ such that $y \equiv z \bmod n$. Then we have $y \equiv z \bmod 3$ because $3 \mid 15$; and similarly $y \equiv z \bmod 5$, because $5 \mid 15$. Hence, $x = y$ is a solution of (3.8) if and only if $x = z$ is a solution of (3.8).

It follows that the solutions of (3.8) are given by $x \equiv 7 \bmod 15$.

Now we give an alternative method to solve these simultaneous congruences, this time using a more structured approach.

Let $x \in \mathbb{Z}$ be a solution of (3.8). Then $x \equiv 1 \bmod 3$ so $x = 1 + 3y$ for some $y \in \mathbb{Z}$. Since $x \equiv 2 \bmod 5$, we obtain

$$\begin{aligned} 1 + 3y &\equiv 2 \bmod 5 \\ 3y &\equiv 1 \bmod 5. \end{aligned} \tag{3.9}$$

We solve this linear congruence equation for y using the method used in Example 3.15. So we look for $k \in \mathbb{Z}$ such that $3k \equiv 1 \bmod 5$. From the equation $2 \cdot 3 + (-1) \cdot 5 = 1$, we see that $2 \cdot 3 \equiv 1 \bmod 5$, so that we can take $k = 2$. Multiplying (3.9) by 2, we obtain

$$y \equiv 2 \bmod 5.$$

Therefore, $y = 2 + 5z$ for some $z \in \mathbb{Z}$. Thus,

$$\begin{aligned} x &= 1 + 3(2 + 5z) \\ &= 7 + 15z. \end{aligned}$$

So $x \equiv 7 \bmod 15$.

Now let $x \in \mathbb{Z}$ with $x \equiv 7 \bmod 15$. Then by reversing the arguments above shows that x is a solution of (3.8). (Alternatively we can check this directly, by saying: since $3 \mid 15$, we have $x \equiv 7 \bmod 3$, so that $x \equiv 1 \bmod 3$, and similarly, we can show that $x \equiv 2 \bmod 5$.)

It follows that the solutions of (3.8) are given by $x \equiv 7 \bmod 15$.

(b) We are going to find all $x \in \mathbb{Z}$ such that

$$\begin{aligned} x &\equiv 4 \bmod 9 \\ x &\equiv 7 \bmod 11. \end{aligned} \tag{3.10}$$

Listing all integers x with $0 \leq x \leq 98$ such that $x \equiv 4 \bmod 9$ and those for which $x \equiv 7 \bmod 11$ would be pretty time consuming. So we proceed using the second method used in (a) above.

Let $x \in \mathbb{Z}$ be a solution of (3.10). Then $x \equiv 4 \bmod 9$ so $x = 4 + 9y$ for some $y \in \mathbb{Z}$. Since $x \equiv 7 \bmod 11$, we obtain

$$\begin{aligned} 4 + 9y &\equiv 7 \bmod 11 \\ 9y &\equiv 3 \bmod 11. \end{aligned} \tag{3.11}$$

We solve this linear congruence equation for y using the method used in Example 3.15. So we look for $k \in \mathbb{Z}$ such that $9k \equiv 1 \bmod 11$. Since 9 and 11 are coprime, we can do this by finding $k, l \in \mathbb{Z}$ such that $9k + 11l = 1$ using the Euclidean algorithm. In this case, we obtain $5 \cdot 9 + (-4) \cdot 11 = 1$. (For small numbers like these it is possible to obtain this equation by guesswork.) Therefore, we have

$$5 \cdot 9 \equiv 1 \bmod 11,$$

so we take $k = 5$.

Now multiplying (3.11) by 5 we obtain

$$\begin{aligned} y &\equiv 15 \bmod 11 \\ &\equiv 4 \bmod 11. \end{aligned}$$

Therefore, $y = 4 + 11z$ for some $z \in \mathbb{Z}$. Thus,

$$\begin{aligned} x &= 4 + 9(4 + 11z) \\ &= 40 + 99z. \end{aligned}$$

So $x \equiv 40 \bmod 99$.

Let $x \in \mathbb{Z}$ with $x \equiv 40 \bmod 99$. Then reversing the arguments above, we can deduce that x is a solution of (3.10).

It follows that the solutions of (3.10) are given by $x \equiv 40 \bmod 99$.

Now we are going to give an alternative way to solve (3.10). First we observe from the equation $5 \cdot 9 + (-4) \cdot 11 = 1$ that

$$\begin{aligned} 45 = 5 \cdot 9 &\equiv 1 \bmod 11 \\ -44 = (-4) \cdot 11 &\equiv 1 \bmod 9 \end{aligned}$$

Also we have

$$\begin{aligned} -44 &\equiv 0 \bmod 11 \\ 45 &\equiv 0 \bmod 9. \end{aligned}$$

Now consider $x = 45 \cdot 7 + (-44) \cdot 4 = 139$. From the congruences above we obtain

$$\begin{aligned} 45 \cdot 7 + (-44) \cdot 4 &\equiv 0 \cdot 7 + 1 \cdot 4 \bmod 9 \\ &\equiv 4 \bmod 9. \end{aligned}$$

and

$$\begin{aligned} 45 \cdot 7 + (-44) \cdot 4 &\equiv 1 \cdot 7 + 0 \cdot 4 \bmod 11 \\ &\equiv 7 \bmod 11. \end{aligned}$$

Hence, $x = 139$ is a solution of (3.10).

This method of solution does not yet guarantee that all other solutions of (3.10) are given by $x \equiv 139 \bmod 99$. This does, however, follow from the Chinese remainder theorem below. Therefore, since $139 \equiv 40 \bmod 99$, the solutions of (3.10) are given by $x \equiv 40 \bmod 99$.

In both of these examples above, we have considered pairs of simultaneous congruences of the form (3.7), where n is coprime to m. Generalizing what we did in these examples above, we can obtain methods for solving such pairs of simultaneous congruences. There are different methods here and you can choose which one you prefer when you have to solve simultaneous congruences.

We move on to consider general systems of simultaneous congruences of the form (3.12) in the Chinese remainder theorem (Theorem 3.19) below. The Chinese remainder theorem tells us about solutions to systems of simultaneous congruences, under a coprimeness assumption. The proof is a bit more advanced than most of the proofs in the course and is a little brief in places, but is included in these printed notes for completeness. This proof is not part of the syllabus and certainly is not examinable, so you may want to omit reading it carefully at first. The idea of the proof is to solve the congruences two at a time using the second method from (b) in Examples 3.18.

Theorem 3.19 (The Chinese remainder theorem)**.**
Let $n_1, n_2, \ldots, n_k \in \mathbb{N}$ *and* $a_1, a_2, \ldots, a_k \in \mathbb{Z}$. *Suppose that* $\mathrm{hcf}(n_i, n_j) = 1$ *for* $i \neq j$. *Consider the system of simultaneous congruences*

$$\begin{aligned} x &\equiv a_1 \bmod n_1 \\ x &\equiv a_2 \bmod n_2 \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \\ x &\equiv a_k \bmod n_k. \end{aligned} \tag{3.12}$$

(a) *There is a solution* $x = s \in \mathbb{Z}$ *of* (3.12).
(b) *Let* $r \in \mathbb{Z}$. *Then* $x = r$ *is a solution of* (3.12) *if and only if* $r \equiv s \bmod n_1 n_2 \cdots n_k$.

Hence, the solutions of (3.12) *are given by* $x \equiv s \bmod n_1 n_2 \cdots n_k$.

Proof. We begin by considering the case $k = 2$ and we let $n_1 = n$, $n_2 = m$, $a_1 = a$ and $a_2 = b$. Since n is coprime to m there exists $k, l \in \mathbb{Z}$ such that $kn + ml = 1$ by Corollary 2.17. From this equation we obtain the congruences

$$\begin{aligned} kn &\equiv 1 \bmod m \\ lm &\equiv 1 \bmod n. \end{aligned}$$

We also clearly have the congruences

$$\begin{aligned} lm &\equiv 0 \bmod m \\ kn &\equiv 0 \bmod n. \end{aligned}$$

Let $x = knb + lma$. We see that

$$\begin{aligned} x = knb + lma &\equiv 0b + 1a \bmod n \\ &\equiv a \bmod n, \end{aligned}$$

and

$$\begin{aligned} x = knb + lma &\equiv 1b + 0a \bmod m \\ &\equiv b \bmod m. \end{aligned}$$

Hence, x is a solution of $x \equiv a \bmod n$ and $x \equiv b \bmod m$, which proves (a).

Now let $y \in \mathbb{Z}$. Suppose that $x \equiv y \bmod nm$. Then $x \equiv y \bmod n$ and $x \equiv y \bmod m$, and thus y is also a solution of (3.12).

Now suppose that y is also a solution of (3.12). Then $x \equiv y \bmod n$ and $x \equiv y \bmod m$. So $n \mid x - y$ and $m \mid x - y$. Therefore, $nm \mid x - y$ by Lemma 3.17(a), and hence $x \equiv y \bmod nm$. This completes the proof of (b), for the case $k = 2$.

Thus we have proved that the pair of simultaneous congruences

$$\begin{aligned} x &\equiv a_1 \bmod n_1 \\ x &\equiv a_2 \bmod n_2. \end{aligned}$$

is equivalent to the single congruence

$$x \equiv c \bmod n_1 n_2,$$

where $c = knb + lma$. Therefore, solving (3.12) is equivalent to solving

$$\begin{aligned} x &\equiv c \bmod n_1 n_2 \\ x &\equiv a_3 \bmod n_3 \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \\ x &\equiv a_k \bmod n_k \end{aligned}$$

so we have reduced the number of congruence equations by one. Also we have that $\mathrm{hcf}(n_1 n_2, n_j) = 1$ for all j by Lemma 3.17(b).

Continuing in this way we can reduce to having a single congruence of the form

$$x \equiv d \bmod n_1 n_2 \cdots n_k,$$

for some $d \in \mathbb{Z}$. This proves the theorem. □

We note that the proof of the Chinese remainder theorem gives a method for solving a system of simultaneous congruences. This method is involves repeatedly solving pairs of simultaneous congruences. You can solve these pairs of congruences in different ways as we saw in Examples 3.18.

We give a couples of examples, where we solve systems of three simultaneous congruences by considering pairs of congruences in turn. The first example is a case of the 30 riddle from the start of this section.

Examples 3.20. (a) We are going to look for $x \in \mathbb{Z}$ such that

$$\begin{aligned} x &\equiv 1 \bmod 2 \\ x &\equiv 2 \bmod 3 \\ x &\equiv 3 \bmod 5. \end{aligned} \tag{3.13}$$

It is straightforward to solve the first two congruences and we proceed using the second method from Examples 3.18(b). To do this we can use the equation $1 = -2+3$ to observe that

$$\begin{aligned} 3 &\equiv 1 \bmod 2 \\ -2 &\equiv 1 \bmod 3. \end{aligned}$$

Also we have

$$\begin{aligned} -2 &\equiv 0 \bmod 2 \\ 3 &\equiv 0 \bmod 3. \end{aligned}$$

Thus we have that $3 \cdot 1 + (-2)\dot{2} = -1$ is a solution of the first pair of congruences, and so is 5, because $5 \equiv -1 \bmod 6$. Therefore, by the Chinese remainder theorem, the solutions are given by

$$x \equiv 5 \bmod 6.$$

Hence, solving (3.13) is equivalent to solving the pair of congruences

$$\begin{aligned} x &\equiv 5 \bmod 6 \\ x &\equiv 3 \bmod 5. \end{aligned}$$

To solve this pair of congruences we write $1 = 6 - 5$, and use this to observe that

$$\begin{aligned} -5 &\equiv 1 \bmod 6 \\ 6 &\equiv 1 \bmod 5. \end{aligned}$$

and also we have

$$\begin{aligned} 6 &\equiv 0 \bmod 6 \\ -5 &\equiv 0 \bmod 5. \end{aligned}$$

Let $x = (-5) \cdot 5 + 6 \cdot 3 = -7$. We have

$$\begin{aligned} (-5) \cdot 5 + 6 \cdot 3 &\equiv 1 \cdot 5 + 0 \cdot 3 \bmod 6 \\ &\equiv 5 \bmod 6. \end{aligned}$$

and

$$\begin{aligned} (-5) \cdot 5 + 6 \cdot 3 &\equiv 0 \cdot 5 + 1 \cdot 3 \bmod 5 \\ &\equiv 3 \bmod 5, \end{aligned}$$

so x is a solution of (3.13). Then, using the Chinese remainder theorem, we deduce that the solutions of the systems of simultaneous congruences (3.13) are given by $x \equiv -7 \bmod 30$, or equivalently

$$x \equiv 23 \bmod 30.$$

(b) We are going to solve of simultaneous congruences

$$\begin{aligned} x &\equiv 3 \bmod 5 \\ x &\equiv 4 \bmod 7 \\ x &\equiv 8 \bmod 11. \end{aligned} \tag{3.14}$$

We first solve the first pair using same method as the first method we used in Examples 3.18(b). So we let x be a solution of the first pair, and say that $x = 3 + 5y$ for some $y \in \mathbb{Z}$. Then we substitute this in to the second congruence to obtain

$$\begin{aligned} 3 + 5y &\equiv 4 \bmod 7 \\ 5y &\equiv 1 \bmod 7. \end{aligned}$$

Next we solve for y by finding $k \in \mathbb{Z}$ such that $5k \equiv 1 \bmod 7$. A bit of trial and error shows that $k = 3$ does the job. Now multiplying by 3 gives

$$\begin{aligned} 15y &\equiv 3 \bmod 7 \\ y &\equiv 3 \bmod 7. \end{aligned}$$

Thus we have $y = 3 + 7z$ for some $z \in \mathbb{Z}$, and so $x = 3 + 5(3 + 7z) = 18 + 35z$. Hence, any solution of the first pair congruences satisfies

$$x \equiv 18 \bmod 35,$$

and we can check that any such x is indeed a solution, so that this gives the solutions.

Now we have to solve the pair of congruences

$$\begin{aligned} x &\equiv 18 \bmod 35 \\ x &\equiv 8 \bmod 11. \end{aligned}$$

We let x be a solution and say that $x = 18 + 35u$ for some $u \in \mathbb{Z}$, and thus

$$\begin{aligned} 18 + 35u &\equiv 8 \bmod 11 \\ 2u &\equiv -10 \bmod 11 \\ 2u &\equiv 1 \bmod 11. \end{aligned}$$

We observe that $6 \cdot 2 = 12 \equiv 1 \bmod 11$. Thus we obtain

$$\begin{aligned} 12u &\equiv 6 \bmod 11 \\ u &\equiv 6 \bmod 11. \end{aligned}$$

Thus we have $u = 6 + 11v$ for some $v \in \mathbb{Z}$, and so $x = 18 + 35(6 + 11v) = 228 + 385v$. Therefore, any solution of (3.14) satisfies

$$x \equiv 228 \bmod 385.$$

Further, we can check that any such x is indeed a solution, so that $x \equiv 228 \bmod 385$ gives all the solutions of (3.14).

Let's finish this section by looking at the "210 riddle", which is a step up from the 30 riddle that we saw at the start of the section.

Think of a natural number x less than 210? Work out

- the remainder a when x is divided by 2;
- the remainder b when x is divided by 3;
- the remainder c when x is divided by 5; and
- the remainder d when x is divided by 7.

It may seem surprising at first that we can determine x uniquely from a, b, c and d. But now we know that this is the case, thanks to the Chinese remainder theorem. In fact if we work through the proof we see that x is the natural number less than 210, which is congruent to

$$105a + 70b + 126c + 120d.$$

You should think through why this works, and then maybe you want to try it out on your friends.

3.5 Congruence classes

In the next section, we're going to define the ring of integers modulo n, which is a "number system" a bit like the integers. First we need to introduce congruence classes, so the next definition is key to our development of modular arithmetic.

Definition 3.21. Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. We define *the congruence class of a modulo n* to be

$$[a]_n = \{x \in \mathbb{Z} \,:\, x \equiv a \bmod n\}.$$

In words, $[a]_n$ is the set of integers that are congruent to a modulo n.

We demonstrate this definition with some examples.

Examples 3.22.

(a) $[2]_6 = \{\ldots, -10, -4, 2, 8, 14, \ldots\}$ and
$[13]_6 = \{\ldots, 1, 7, 13, 19, 25, \ldots\}$.

(b) $[5]_{11} = \{\ldots, -17, -6, 5, 16, 27, \ldots\}$ and
$[-17]_{11} = \{\ldots, -39, -28, -17, -6, 5, \ldots\}$.
So $[5]_{11} = [-17]_{11}$.

(c) $[0]_2$ is the set of even integers, and
$[1]_2$ is the set of odd integers.

(d) Let $a \in \mathbb{Z}$. Then $[a]_1 = \mathbb{Z}$.

The next proposition gives an alternative description of congruence classes.

Proposition 3.23. *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then $[a]_n$ is the set of $x \in \mathbb{Z}$ such that a and x leave the same remainder when divided by n.*

Proof. Let $x \in \mathbb{Z}$. By Lemma 3.3, $x \in [a]_n$ if and only if a and x leave the same remainder when divided by n. □

By Corollary 3.6, the relation $\sim$ defined on $\mathbb{Z}$ by $a \sim b$ means $a \equiv b \bmod n$ is an equivalence relation. Comparing the definitions of equivalence classes from Definition A.11 and the definition of congruence classes from Definition 3.21, we see that

$$[a]_\sim = [a]_n$$

for any $a \in \mathbb{Z}$. Thus we can use the theory of equivalence relations to study congruence classes.

Below we recall Theorem A.15, which is in Appendix A, and is the important result about equivalence relations that we want to use.

Theorem. *Let A be a set, $\sim$ an equivalence relation on A, and $a, b \in A$. Then the following hold:*

(a) $a \in [a]_\sim$;
(b) $[a]_\sim = [b]_\sim$ *if and only if* $a \sim b$;
(c) $[a]_\sim = [b]_\sim$ *or* $[a]_\sim \cap [b]_\sim = \varnothing$.
(d) $A/\sim$ *is a partition of* A.

Now we can apply this theorem to deduce the corollary below. Parts (a) and (b) of the corollary follow immediately and part (c) can be deduced from Corollary 3.4. You may also find it useful to look back at the examples above to help you understand (a) and (b).

Corollary 3.24. *Let $n \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. Then*

(a) $[a]_n = [b]_n$ *if and only if* $a \equiv b \bmod n$.
(b) $[a]_n = [b]_n$ *or* $[a]_n \cap [b]_n = \varnothing$.
(c) *there are exactly n congruence classes modulo n, namely*

$$[0]_n, [1]_n, [2]_n, \ldots, [n-2]_n \text{ and } [n-1]_n.$$

Alternatively, we note that it is fairly easy to deduce the Corollary 3.24 directly from Proposition 3.23.

3.6 The ring of integers modulo n

The properties of congruences that we saw earlier can be put together nicely by defining an addition and multiplication on the set of congruence classes modulo n, which we denote by $\mathbb{Z}_n$. In Definition 3.25 we define addition and multiplication on $\mathbb{Z}_n$, so it is a "number system" a bit like the integers $\mathbb{Z}$. We call $\mathbb{Z}_n$ with this addition and multiplication the ring of integers modulo n. In Section 3.7, we'll see that this $\mathbb{Z}_n$ shares a lot of properties with $\mathbb{Z}$.

Lets's dive in with the definition of the ring of integers modulo n. Then we'll have some examples to help us to understand it.

Definition 3.25. Let $n \in \mathbb{N}$. We define *the set of congruence classes modulo n* to be

$$\mathbb{Z}_n = \{[a]_n \ : \ a \in \mathbb{Z}\}.$$

We define an addition $+$ and multiplication $\cdot$ on $\mathbb{Z}_n$ as follows. Let $x, y \in \mathbb{Z}_n$ and choose $x_0, y_0 \in \mathbb{Z}$ such that

$$x = [x_0]_n \quad \text{and} \quad y = [y_0]_n.$$

Define

$$x + y = [x_0 + y_0]_n$$

and

$$x \cdot y = [x_0 y_0]_n.$$

The set $\mathbb{Z}_n$ with the addition $+$ and multiplication $\cdot$ is called *the ring of integers modulo* n.

Note that $\mathbb{Z}_n$ is a set of subsets of $\mathbb{Z}$, which may seem a bit weird to get your head around at first, but once we've worked with it for a bit, it will get better. We have

$$\mathbb{Z}_n = \{[0]_n, [1]_n, [2]_n, \ldots, [n-1]_n\},$$

and, in practice, we can think of the elements $[a]_n$ of $\mathbb{Z}_n$ just to be some symbols and we have rules for adding and multiplying them. Let's look at a couple of examples to help us understand the definition of $\mathbb{Z}_n$.

Examples 3.26. (a) We consider the case $n = 2$. We have

$$\mathbb{Z}_2 = \{[0]_2, [1]_2\}.$$

So $\mathbb{Z}_2$ is the set containing the set of even numbers and the set of even numbers. For now we denote $[0]_2 = \text{even}$ and $[1]_2 = \text{odd}$, so $\mathbb{Z}_2 = \{\text{even}, \text{odd}\}$.

The addition on $\mathbb{Z}_2$ is given by the addition table below.

$+$	even	odd
even	even	odd
odd	odd	even

So one thing this table is saying is

$$\text{even} + \text{odd} = \text{odd},$$

which is just the familiar fact that if we add an even number to an odd number, then we get an odd number.

The multiplication on $\mathbb{Z}_2$ is given by the multiplication table below.

$\cdot$	even	odd
even	even	even
odd	even	odd

One thing this table is saying is that

$$\text{even} \cdot \text{odd} = \text{even},$$

which is just saying that if we multiply an even and odd number then, as we know well, we get an even number.

Now putting the addition and multiplication table in our original notation for $\mathbb{Z}_2$ we have.

$+$	$[0]_2$	$[1]_2$
$[0]_2$	$[0]_2$	$[1]_2$
$[1]_2$	$[1]_2$	$[0]_2$

$\cdot$	$[0]_2$	$[1]_2$
$[0]_2$	$[0]_2$	$[0]_2$
$[1]_2$	$[0]_2$	$[1]_2$

(b) Now we consider the case $n = 6$, which is large enough to give us a better feeling about the definition on $\mathbb{Z}_n$. We have

$$\mathbb{Z}_6 = \{[0]_6, [1]_6, [2]_6, [3]_6, [4]_6, [5]_6\}.$$

We can work out the addition table below. A couple of examples of the calculations required are:

- $[4]_6 + [1]_6 = [5]_6$; and
- $[4]_6 + [5]_6 = [9]_6 = [3]_6$.

For the second sum above, we can have the equality $[9]_6 = [3]_6$, because $9 \equiv 3 \bmod 6$. In general for $a, b \in \{0, 1, 2, 3, 4, 5\}$ we work out $[a]_6 + [b]_6 = [c]_6$, where $c \in \{0, 1, 2, 3, 4, 5\}$ with $c \equiv a + b \bmod 6$ to get the table below.

$+$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[0]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[1]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$
$[2]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$
$[3]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$
$[4]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$
$[5]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$

Similarly, we can work out the multiplication table below. A couple of examples of the calculations required are:

- $[2]_6 \cdot [4]_6 = [8]_6 = [2]_6$; and
- $[5]_6 \cdot [3]_6 = [15]_6 = [3]_6$.

Above we have the equalities: $[8]_6 = [2]_6$, because $8 \equiv 2 \bmod 6$; and $[15]_6 = [3]_6$, because $15 \equiv 3 \bmod 6$. In general for $a, b \in \{0, 1, 2, 3, 4, 5\}$ we work out $[a]_6 \cdot [b]_6 = [c]_6$, where $c \in \{0, 1, 2, 3, 4, 5\}$ with $c \equiv ab \bmod 6$ to get the table below.

$\cdot$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[0]_6$	$[0]_6$	$[0]_6$	$[0]_6$	$[0]_6$	$[0]_6$	$[0]_6$
$[1]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[2]_6$	$[0]_6$	$[2]_6$	$[4]_6$	$[0]_6$	$[2]_6$	$[4]_6$
$[3]_6$	$[0]_6$	$[3]_6$	$[0]_6$	$[3]_6$	$[0]_6$	$[3]_6$
$[4]_6$	$[0]_6$	$[4]_6$	$[2]_6$	$[0]_6$	$[4]_6$	$[2]_6$
$[5]_6$	$[0]_6$	$[5]_6$	$[4]_6$	$[3]_6$	$[2]_6$	$[1]_6$

We note that in the $n = 6$ example above, we worked out $[a]_6 + [b]_6 = [c]_6$, where $c \in \{0, 1, 2, 3, 4, 5\}$ with $c \equiv a + b \bmod 6$. In fact we could define addition on $\mathbb{Z}_n = \{[0]_n, [1]_n, [2]_n, \ldots, [n-1]_n\}$ by saying that for $a, b \in \{0, 1, 2, \ldots, n-1\}$, we define $[a]_n + [b]_n = [c]_n$, where $c \in \{0, 1, 2, \ldots, n-1\}$ with $c \equiv a + b \bmod n$; and we could define multiplication similarly. This may seem simpler and it is not difficult for us to show that the definition is equivalent, using Lemma 3.7. However, it turns out to be more convenient to work with Definition 3.25 as we'll see in the next section when we prove some properties of $\mathbb{Z}_n$. We are left with a potential problem to think about though, which we explain next.

Are $+$ and $\cdot$ well defined on $\mathbb{Z}_n$?

You may have noticed that there are many different ways to work out a sum or products in $\mathbb{Z}_6$ in the example above.

For instance, let $x = [3]_6$ and $y = [4]_6$. It is also possible write $x = [15]_6$ and $y = [-2]_6$. Then to work out $x + y$ we have the choice of working out either $[3]_6 + [4]_6 = [7]_6$, or working out $[15]_6 + [-2]_6 = [13]_6$. We have $[7]_6 = [1]_6 = [13]_6$, so in the end the calculation didn't depend on the choice.

Let's consider another instance, let $x = [13]_6$ and $y = [5]_6$, where it is possible to write $x = [1]_6$ and $y = [-1]_6$. Then to calculate $x \cdot y$ we can either calculate $[13]_6 \cdot [5]_6 = [65]_6$, or $[1]_6 \cdot [-1]_6 = [-1]_6$. We have $[65]_6 = [5]_6 = [-1]_6$, so in the end the calculation didn't depend on the choice.

Now let's consider this idea generally. Let $n \in \mathbb{N}$. There is a potential ambiguity in the definition of the binary operation $+$ on $\mathbb{Z}_n$. Let $x, y \in \mathbb{Z}_n$ and suppose that we wish to calculate $x + y$. Using the rule in Definition 3.25, we choose $x_0, y_0 \in \mathbb{Z}$ such that $x = [x_0]_n$ and $y = [y_0]_n$ and then get the answer

$$x + y = [x_0 + y_0]_n.$$

But what would happen if instead we picked different $x_0', y_0' \in \mathbb{Z}$ such that $x = [x_0']_n$ and $y = [y_0']_n$ then we would get the answer

$$x + y = [x_0' + y_0']_n.$$

Obviously, there would be a problem if

$$[x_0 + y_0]_n \neq [x_0' + y_0']_n.$$

It turns out that this cannot happen, and we explain why below.

Since, $[x_0]_n = [x_0']_n$, we have $x_0 \equiv x_0' \bmod n$, and similarly $y_0 \equiv y_0' \bmod n$. Therefore, by Lemma 3.7, we have $x_0 + y_0 \equiv x_0' + y_0' \bmod n$, so that $[x_0 + y_0]_n = [x_0' + y_0']_n$. So the two possible definitions of $x + y$ are equal. We express this by saying that $+$ is *well defined* on $\mathbb{Z}_n$.

In general if we define something, which involves some choices, then we say that it is *well defined*, if it does not depend on those choices. We can show that $\cdot$ is well defined using a similar argument to above, and is exercise Q3.16. You can also look at exercise Q3.17 to see another example of a function that is well defined.

3.7 Properties of $\mathbb{Z}_n$

Below we will see that addition and multiplication in $\mathbb{Z}_n$ satisfy a number of familiar properties of addition and multiplication in $\mathbb{Z}$. Before we do this, we give a list of some properties for $\mathbb{Z}$: all of these properties should be very familiar to you. On the right are the names for these properties.

(A0) For all $x, y \in \mathbb{Z}$, $x + y \in \mathbb{Z}$ (closure under addition)

(A1) For all $x, y, z \in \mathbb{Z}$, $(x + y) + z = x + (y + z)$. (associative law of addition)

(A2) There exists $0 \in \mathbb{Z}$ such that for all $x \in \mathbb{Z}$, $x + 0 = x = 0 + x$. (existence of additive identity)

(A3) For all $x \in \mathbb{Z}$, there exists $-x \in \mathbb{Z}$ such that $x + (-x) = 0 = (-x) + x$. (existence of additive inverses)

(A4) For all $x, y \in \mathbb{Z}$, $x + y = y + x$. (commutative law of addition)

(M0) For all $x, y \in \mathbb{Z}$, $x \cdot y \in \mathbb{Z}$ (closure under multiplication)

(M1) For all $x, y, z \in \mathbb{Z}$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$. (associative law of multiplication)

(M2) There exists $1 \in \mathbb{Z}$ such that for all $x \in \mathbb{Z}$, $x \cdot 1 = x = 1 \cdot x$. (existence of multiplicative identity)

(M4) For all $x, y \in \mathbb{Z}$, $x \cdot y = y \cdot x$. (commutative law of multiplication)

(D) For all $x, y, z \in \mathbb{Z}$, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$. (distributive law)

The way that (A2) and (M2) are phrased may seem a little odd at first, they are just saying that there are special elements in $\mathbb{Z}$, which we denote by 0 and 1; and these are of course just the integers 0 and 1. Similarly, the element $-x \in \mathbb{Z}$ in (A3) is the integer that the notation suggests. Also don't worry that (M3) is missing, this is not a typo and there is a reason for this, which you'll see if your study rings in Algebra 2.

We'll see that the addition and multiplication on $\mathbb{Z}_n$ satisfy (essentially) the same list of properties as those above for $\mathbb{Z}$. In the following lemma we show that addition on $\mathbb{Z}_n$ is associative and also that the distributive law holds in $\mathbb{Z}_n$.

Lemma 3.27. *Let $x, y, z \in \mathbb{Z}_n$. Then:*

(a) $(x + y) + z = x + (y + z)$. *In other words $+$ is associative.*
(b) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$. *In other words $+$ is distributive over $\cdot$.*

Proof. (a) Let $x_0, y_0, z_0 \in \mathbb{Z}$ such that $x = [x_0]_n$, $y = [y_0]_n$ and $z = [z_0]_n$. By the rule for $+$, we have

$$x + y = [x_0 + y_0]_n.$$

Applying the rule again gives

$$(x + y) + z = [(x_0 + y_0) + z_0]_n. \tag{3.15}$$

Similarly, we get

$$x + (y + z) = [x_0 + (y_0 + z_0)]_n. \tag{3.16}$$

We know that addition of integers is associative, so $(x_0 + y_0) + z_0 = x_0 + (y_0 + z_0)$. Therefore, (3.15) and (3.16) give

$$(x + y) + z = x + (y + z).$$

(b) Let $x_0, y_0, z_0 \in \mathbb{Z}$ such that $x = [x_0]_n$, $y = [y_0]_n$ and $z = [z_0]_n$. By the rule for $+$, we have

$$y + z = [y_0 + z_0]_n.$$

Applying the rule for $\cdot$ gives

$$x + (y + z) = [x_0(y_0 + z_0)]_n. \tag{3.17}$$

Similarly we can show that

$$(x \cdot y) + (x \cdot z) = [x_0y_0 + x_0z_0]_n. \tag{3.18}$$

We know that the distributive law holds for $\mathbb{Z}$, so $x_0(y_0 + z_0) = x_0y_0 + x_0z_0$. Therefore, (3.17) and (3.18) give

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

□

We can prove in a similar way that all of the properties in the following list are satisfied.

(A0) For all $x, y \in \mathbb{Z}_n$, $x + y \in \mathbb{Z}_n$ (closure under addition)

(A1) For all $x, y, z \in \mathbb{Z}_n$, $(x + y) + z = x + (y + z)$.
(associative law of addition)

(A2) There exists $[0]_n \in \mathbb{Z}_n$ such that for all $x \in \mathbb{Z}_n$, $x + [0]_n = x = [0]_n + x$.
(existence of additive identity)

(A3) For all $x \in \mathbb{Z}_n$, there exists $-x \in \mathbb{Z}_n$ such that $x + (-x) = [0]_n = (-x) + x$.
(existence of additive inverses)

(A4) For all $x, y \in \mathbb{Z}_n$, $x + y = y + x$.
(commutative law of addition)

(M0) For all $x, y \in \mathbb{Z}_n$, $x \cdot y \in \mathbb{Z}_n$
(closure under multiplication)

(M1) For all $x, y, z \in \mathbb{Z}_n$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
(associative law of multiplication)

(M2) There exists $[1]_n \in \mathbb{Z}_n$ such that for all $x \in \mathbb{Z}_n$, $x \cdot [1]_n = x = [1]_n \cdot x$.
(existence of multiplicative identity)

(M4) For all $x, y \in \mathbb{Z}_n$, $x \cdot y = y \cdot x$.
(commutative law of multiplication)

(D) For all $x, y, z \in \mathbb{Z}_n$, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.
(distributive law)

In mathematical language this list of properties tells us that $\mathbb{Z}_n$ (with the addition and multiplication defined in Definition 3.25) is a *commutative ring with one*. Also as these properties are satisfied by $\mathbb{Z}$, we have that $\mathbb{Z}$ is another example of a commutative ring with one.

There are many other important examples of rings in mathematics, and you'll be able to learn more about rings in Algebra 2. In that course we'll see how many familiar properties of the integers hold more generally for other rings. The theory of rings is central to mathematics, with motivation and applications throughout mathematics and the physical sciences. The use of rings in number theory and algebraic geometry led to a major development of their theory throughout the 20th century. Nowadays ring theory and applications of ring theory remain amongst the most important areas of mathematics research. Indeed much of my own research regards the structure and representation theory of certain rings.

Don't worry if this last section seems a bit abstract at the moment. For now you should just get an idea of what terms like "associative", "commutative", "additive inverse" and "multiplicative identity" mean. Later on it will all seem more understandable.

3.8 Fermat's little theorem

To end this chapter build on the material we've built up and cover a couple of very nice applications of the theory, namely Fermat's little theorem (in this section) and RSA

cryptography (in the next section). We will use the notation given in the following definition.

Definition 3.28. Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. We write $a \pmod n$ to denote the unique element b of $\{0, 1, 2, \ldots, n-1\}$ such that $a \equiv b \bmod n$.

For example, $19 \pmod 4 = 3$ and $23 \pmod 6 = 5$.

We going to prove a cool theorem called Fermat's little theorem. Before, we prove the theorem we demonstrate it with an example.

Example 3.29. Let $p = 7$ and let $a = 3$. In the table below we look at the values of $3b \pmod 7$ for all $b = 1, 2, \ldots, 6$.

b	1	2	3	4	5	6
$3b$	3	6	9	12	15	18
$3b \pmod 7$	3	6	2	5	1	4

We can see that the bottom row gives a rearrangement of $1, 2, 3, 4, 5, 6$. Therefore, we see that

$$(1 \cdot 3)(2 \cdot 3)(3 \cdot 3)(4 \cdot 3)(5 \cdot 3)(6 \cdot 3) \equiv 6! \bmod 7,$$

by Lemma 3.7. Therefore, we have

$$6! \cdot 3^6 \equiv 6! \bmod 7.$$

Now we see that $7 \nmid 6! = 720$. Therefore, as 7 is prime, it is coprime to 6!. Thus by Corollary 3.16, we obtain

$$3^6 \equiv 1 \bmod 7.$$

We now use the idea in the example above to prove Fermat's little theorem.

Theorem 3.30. *Let $p \in \mathbb{N}$ be prime and let $a \in \mathbb{Z}$. Suppose that $p \nmid a$. Then*

$$a^{p-1} \equiv 1 \bmod p.$$

Proof. Since p is prime and $p \nmid a$, we have that a is coprime to p.

For $b = 1, 2, \ldots, p-1$, let $d_b = ab \pmod p$.
Suppose that $d_b = d_c$ for $b, c \in \{0, 1, 2, \ldots, p-1\}$. Then we have we have $ab \equiv ac \bmod p$ and thus $b \equiv c \bmod p$ by Corollary 3.16. Therefore, $b = c$, because $b, c \in \{0, 1, \ldots, p-1\}$. It follows that $d_1, d_2, \ldots, d_{p-1}$ is a rearrangement of $1, 2, \ldots, p-1$. Thus $d_1 d_2 \ldots d_{p-1} = (p-1)!$. Also using Lemma 3.7, we have

$$\begin{aligned}(p-1)!\, a^{p-1} &= (1a)(2a)\ldots((p-1)a) \\ &\equiv d_1 d_2 \ldots d_{p-1} \bmod p\end{aligned}$$

Hence,

$$(p-1)!\, a^{p-1} \equiv (p-1)! \bmod p.$$

Now $p \nmid (p-1)!$ and thus p is coprime to $(p-1)!$. Therefore, by Corollary 3.16, we obtain

$$a^{p-1} \equiv 1 \bmod p. \qquad \square$$

We now give the following corollary of Theorem 3.30.

Corollary 3.31. *Let $p \in \mathbb{N}$ be prime and let $a \in \mathbb{Z}$. Then*

$$a^p \equiv a \bmod p.$$

Proof. We consider two cases.
Case 1: $a \equiv 0 \bmod p$. Then $a^p \equiv 0 \bmod p$, so $a^p \equiv a \bmod p$.
Case 2: $a \not\equiv 0 \bmod p$. Then $a^{p-1} \equiv 1 \bmod p$ by Theorem 3.30. So $a^p \equiv a \bmod p$. □

Another way of stating the corollary above is to say that for any integer a and a prime p, we have that

$$p \mid a^p - a.$$

This is a really striking statement!

Next we prove a theorem that is similar to Fermat's last theorem. We'll need this in the next section when we look at RSA public key cryptography.

Theorem 3.32. *Let $p, q \in \mathbb{N}$ be distinct primes, $k \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then*

$$a^{k(p-1)(q-1)+1} \equiv a \bmod pq.$$

Proof. First we show that $a^{k(p-1)(q-1)+1} \equiv a \bmod p$.
If $a \equiv 0 \bmod p$, then this is clear.
If $a \not\equiv 0 \bmod p$, then $p \nmid a$, so $a^{p-1} \equiv 1 \bmod p$, by Theorem 3.30 and, therefore, $a^{k(p-1)(q-1)} \equiv 1 \bmod p$. Hence,

$$a^{k(p-1)(q-1)+1} \equiv a \bmod p.$$

Similarly, we can show that

$$a^{k(p-1)(q-1)+1} \equiv a \bmod q.$$

Therefore,

$$p \mid a^{k(p-1)(q-1)+1} - a \quad \text{and} \quad q \mid a^{k(p-1)(q-1)+1} - a.$$

Since, $p \neq q$, we have that p is coprime to q. Therefore,

$$pq \mid a^{k(p-1)(q-1)+1} - a,$$

by Lemma 3.17(a). Hence,

$$a^{k(p-1)(q-1)+1} \equiv a \bmod pq.$$ □

3.9 The RSA cryptosystem

People have always had the need to communicate in a secret way, so that their enemies are not able to understand what they are saying. Consequently many *cryptosystems* for encoding communications have been designed. Typically a cryptosystem works as explained below.

Alice wants to send a secret message to Bob. They proceed as follows.

- Alice converts the message into a sequence of numbers $\mathbf{m} = (m_1, m_2, \ldots, m_r)$ called the *plaintext*.
- Alice enciphers the plaintext by performing some operation on the string of numbers to obtain a different sequence of natural numbers $\mathbf{c} = (c_1, c_2, \ldots, c_r)$ called the *ciphertext* and sends it to Bob.
- Bob knows how to invert the operation that Alice performed, so he is able decipher the ciphertext to calculate $\mathbf{m}$ from $\mathbf{c}$.

Often the process of enciphering and deciphering involves knowledge of a *key*.

The general description of a cryptosystem given above is unlikely to make that much sense, until we have seen an example.

A symmetric key cryptosystem

Alice wants to send Bob a message using a simple cryptosystem known as a *Caesar shift*. In advance they have agreed on a key, which is used to encrypt and decrypt the message. The key k is an integer between 0 and 25; in this example we take $k = 18$. Alice wants to send the message

EAT MY SHOES

She first converts each letter in the alphabet to a natural number between 0 and 25, where $A \mapsto 1, B \mapsto 2, \ldots Y \mapsto 25, Z \mapsto 0$ to obtain the plaintext

$$\mathbf{m} = (5, 1, 20, 13, 25, 8, 15, 5, 19).$$

Next for each of the entries m_i in $\mathbf{m}$ she calculates

$$c_i = m_i + 18 \,(\mathrm{mod}\, 26)$$

to obtain the ciphertext

$$\mathbf{c} = (23, 19, 12, 5, 17, 0, 7, 23, 1).$$

Then Alice sends $\mathbf{c}$ to Bob. Since Bob knows the key is 18, he is able to calculate

$$m_i = c_i - 18 \,(\mathrm{mod}\, 26)$$

and recover the plaintext $\mathbf{m}$.

This Caesar shift is not very secure because if someone is able to guess what the key is, then they can break the code. In fact it would be very easy to guess the key using some "frequency analysis" if the message was longer. Also an enemy would only need to try 26 possible keys before managing to decipher the message.

The Caesar shift is an example of a *symmetric key cryptosystem*. This is one in which two parties agree on a secret key in advance of communication. There are symmetric key cryptosystems that are secure if Alice and Bob are able to secretly communicate the key between themselves. However, this is likely to be problem, because they don't yet have a way to communicate securely. This difficulty makes symmetric key cryptosystems impractical for the amount of information that needs to be encoded nowadays for secure internet transactions.

Below we describe the RSA cryptosystem, which is a *public key cryptosystem*. Public key cryptosystems involve a *public key* used to encode, and a *private key* used to decode. Therefore, they avoid the problem of having to communicate the key used for encryption and decryption.

The RSA cryptosystem

The RSA public key cryptosystem is used for many of the secure transactions that we make on the internet, so we are utterly dependent on it. The security is based on the belief that it is very difficult to factorize large numbers into a product of primes, which we discuss before moving on to the RSA cryptosystem.

Suppose you wanted to factorize 6557, then you could get a calculator out, and you would work out quite quickly that $6557 = 79 \cdot 83$. However, if you wanted to factorize 9,088,109 then it would take you quite a long time to work out that $9,088,109 = 2969 \cdot 3061$. As we see below the security of the RSA cryptosystem depends on factorizing a number with about 400 digits into the product of two primes. It is estimated that this would take thousands of years using the most powerful computers. So for practical purposes it is completely infeasible.

The RSA cryptosystem works as follow, when Alice wants to send a message to Bob.

Encryption

First Bob needs a *public key*. To get a public key Bob finds two large prime numbers p and q with $p \neq q$ and sets $N = pq$, and he also chooses $e \in \mathbb{N}$ such that $0 < e < (p-1)(q-1)$ and e is coprime to $(p-1)(q-1)$. Bob's public key is the pair (N, e). He makes his public key available to everyone.

Alice wants to send a message to Bob. First Alice converts her message so that the plaintext is a sequence of natural numbers

$$\mathbf{m} = (m_1, m_2, \ldots, m_r),$$

where $0 \leq m_i < N$ for $i = 1, 2, \ldots, r$. We don't go into details, but this can be done in a similar way to the assignment $A \mapsto 1, B \mapsto 2, \ldots Y \mapsto 25, Z \mapsto 0$, except that each m_i corresponds to a sequence of letters. To encode this Alice calculates

$$c_i = m_i^e \pmod{N}$$

for $i = 1, 2, \ldots, r$. Then the ciphertext is

$$\mathbf{c} = (c_1, c_2, \ldots, c_r).$$

Decryption

Bob needs his *private key* to decrypt the ciphertext. To calculate the private key, he uses the Euclidean algorithm to find $x, y \in \mathbb{Z}$ such that

$$x(p-1)(q-1) + ye = 1.$$

Then the private key is $d = y \pmod{(p-1)(q-1)}$. Thus $d \in \mathbb{N}$ is the unique natural number that satisfies $0 < d < (p-1)(q-1)$ and $ed \equiv 1 \bmod (p-1)(q-1)$.

When Bob receives the ciphertext he calculates $c_i^d \pmod{N}$ for $i = 1, 2, \ldots, r$. Now $ed = k(p-1)(q-1)+1$ for some $k \in \mathbb{N}$ and $m_i^{k(p-1)(q-1)+1} \equiv m_i \bmod N$, by Theorem 3.32.

Therefore,

$$\begin{aligned} c_i^d \pmod N &= m_i^{ed} \pmod N \\ &= m_i^{k(p-1)(q-1)+1} \pmod N \\ &= m_i. \end{aligned}$$

So Bob has recovered the plaintext.

Security

The security of the communication using the RSA cryptosystem depends on the fact that an enemy who intercepts the message is not able to decode it. Suppose an enemy, called Eve, wants to intercept and decode the message. Eve knows what $N = pq$ and e are, but in order to be able to decode the message, she has to know what d is. In order to work out d, Eve needs to know what $(p-1)(q-1)$ is. Now

$$(p-1)(q-1) = pq - p - q + 1$$

so if Eve knows $(p-1)(q-1)$, then she can work out what $p+q$ is. Then from knowing $p+q$ and pq she can work out what p and q are.

It follows that to break the RSA cryptosystem with public key (N, d), Eve needs to be able to find the prime numbers p and q such that $N = pq$. So being able to decode a message is equivalent to being able to factorize a large number into a product of primes. At present the prime numbers p and q used for an RSA public key typically have about 200 digits. As discussed above it is infeasible to factorize a number with 400 digits into a product of primes, so it is infeasible to break the RSA cryptosystem. Thus it is effectively impossible for Eve to decode Alice's message to Bob.

Summary

We summarize the protocol for a secret message to be sent by Alice to Bob using the RSA cryptosystem.

- Bob creates a public key (N, e), where $N = pq$ is the product of primes p and q and $e \in \mathbb{N}$ such that $0 < e < (p-1)(q-1)$ and e is coprime to $(p-1)(q-1)$.
- Alice encodes the plaintext $\mathbf{m} = (m_1, m_2, \ldots, m_r)$ by setting $c_i = m_i^e \pmod N$ to obtain the ciphertext $\mathbf{c} = (c_1, c_2, \ldots, c_r)$.
- Bob calculates the private key d, which is the unique natural number such that $0 < d < (p-1)(q-1)$ and $ed \equiv 1 \bmod (p-1)(q-1)$ using the Euclidean algorithm.
- Bob calculates $c_i^d \pmod N = m_i$, to decrypt the ciphertext and recover the plaintext.

To end this section we give an example of using the RSA cryptosystem. We use much smaller primes than those used in practice.

Example 3.33. Let $p = 29$ and $q = 37$, so we have $N = 1073$, and we let $e = 11$. So the public key is $(1073, 11)$.

Next we find the private key. First we calculate $(p-1)(q-1) = 1008$. Then we use the Euclidean algorithm to find $x, y \in \mathbb{Z}$ such that $1008x + 11y = 1$. First we write

$$1008 = 91 \cdot 11 + 7$$

Second we write

$$11 = 7 + 4.$$

Third we write

$$7 = 4 + 3.$$

Fourth we write

$$4 = 3 + 1.$$

Then we reverse these steps to get

$$\begin{aligned} 1 &= 4 - 3 \\ &= 4 - (7 - 4) \\ &= -7 + 2 \cdot 4 \\ &= -7 + 2 \cdot (11 - 7) \\ &= 2 \cdot 11 - 3 \cdot 7 \\ &= 2 \cdot 11 - 3 \cdot (1008 - 91 \cdot 11) \\ &= -3 \cdot 1008 + 275 \cdot 11. \end{aligned}$$

Therefore, we have $275 \cdot 11 \equiv 1 \bmod 1008$. Thus the private key is $d = 275$.

Now suppose we want to encode the plaintext

$$\mathbf{m} = (134, 529, 406).$$

We calculate

$$\begin{aligned} 134^{11} \pmod{1073} &= 251, \\ 529^{11} \pmod{1073} &= 545, \\ 406^{11} \pmod{1073} &= 406. \end{aligned}$$

So we obtain the ciphertext

$$\mathbf{c} = (251, 545, 406).$$

Decoding involves the calculations

$$\begin{aligned} 251^{275} \pmod{1073} &= 134, \\ 545^{275} \pmod{1073} &= 529, \\ 406^{275} \pmod{1073} &= 406. \end{aligned}$$

Note that I did these calculations using a modular arithmetic calculator like the one you can find on http://users.wpi.edu/~martin/mod.html.

The theory of cryptography is a really interesting branch of pure mathematics. An excellent book that you can read to find out more is:

- S. Singh, *The Code Book: The Secret History of Codes and Code-breaking*, Fourth Estate Ltd., 2002.

Cryptography is also discussed in Chapter 17 of Liebeck's book, which is the recommended book for this course. There is also a lot of information on wikipedia and there are many other references.

Another interesting problem related to the RSA cryptosystem, is the need to find very large primes. There is some really nice mathematics behind this, and you can read more about it in Chapter 16 of Liebeck's book. Possibly some information will be added here in an updated version of these notes in future.

3.10 Summary of Chapter 3

By the end of this chapter you should be able to:

- state the definition of congruence modulo n;
- prove elementary lemmas about congruences and arithmetic of congruences;
- perform calculations with congruences;
- solve linear congruence equations;
- prove and apply the theorem that "if a is coprime to n, then there exists z such that $az \equiv 1 \bmod n$";
- solve systems of simultaneous congruences and understand the Chinese remainder theorem;
- explain the construction of $\mathbb{Z}_n$ and make calculations in $\mathbb{Z}_n$;
- state the definition of $a \,(\mathrm{mod}\, n)$;
- state and apply Fermat's little theorem; and
- explain and justify how the RSA cryptosystem works, and calculate examples.

3.11 Exercises for Chapter 3

Q3.1. True or false:

(a) $4 \equiv 28 \bmod 6$.
(b) $7 \equiv 33 \bmod 5$.
(c) $12 \equiv 6 \bmod 4$.
(d) $-5 \equiv 72 \bmod 11$.

Q3.2. Prove Lemma 3.7(b):

Lemma. *Let $n \in \mathbb{N}$ and $a, b, a', b' \in \mathbb{Z}$. Suppose that $a \equiv b \bmod n$ and $a' \equiv b' \bmod n$. Then $aa' \equiv bb' \bmod n$.*

Q3.3. Find the remainder when:

(a) 7^{16} is divided by 5.
(b) $15^{43} - 3^{23}$ is divided by 14.

Q3.4. Determine whether each of the following statements is true and justify your answer.

(a) Let $n \in \mathbb{N}$ and $a, b, a', b' \in \mathbb{Z}$. Suppose that $a+a' \equiv b+b' \bmod n$. Then $a \equiv b \bmod n$ and $a' \equiv b' \bmod n$.
(b) Let $n \in \mathbb{N}$ and $a, b, c \in \mathbb{Z}$. Suppose that $ac \equiv bc \bmod n$. Then $a \equiv b \bmod n$.
(c) Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then $a^n \equiv a \bmod n$.

When you are asked to justify your answer it means you have to prove it if it is true and give a counterexample if it is not true.

Q3.5. Let $a \in \mathbb{Z}$.

(a) Prove that a^2 is congruent to 0 or 1 modulo 4.
(b) Prove that a^4 is congruent to 0 or 1 modulo 5.

Q3.6. Let $n \in \mathbb{N}$ with digits $a_r a_{r-1} \dots a_2 a_1 a_0$. So

$$n = a_0 + 10a_1 + 10^2 a_2 + \cdots + 10^{r-1} a_{r-1} + 10^r a_r.$$

(a) Prove that $9 \mid n$ if and only if $9 \mid a_0 + a_1 + a_2 + \cdots + a_{r-1} + a_r$.
(b) Prove that $11 \mid n$ if and only if $11 \mid a_0 - a_1 + a_2 - \cdots + (-1)^{r-1} a_{r-1} + (-1)^r a_r$.

Q3.7. Solve the following linear congruence equations.

(a) $4x \equiv 6 \bmod 8$
(b) $2x \equiv 8 \bmod 10$

Q3.8. Solve the following linear congruence equations.

(a) $3x \equiv 4 \bmod 11$
(b) $7x \equiv 2 \bmod 13$

Q3.9. Let $a, b, n \in \mathbb{N}$. Consider the linear congruence equation.

$$ax \equiv b \bmod n \tag{3.19}$$

(a) Suppose that (3.19) has a solution. Prove that $\mathrm{hcf}(a, n) \mid b$.
(b) Suppose that $\mathrm{hcf}(a, n) \mid b$.
Let $h = \mathrm{hcf}(a, n)$, $a' = \frac{a}{h}$, $b' = \frac{b}{h}$ and $n' = \frac{n}{h}$.
Prove that x is a solution to (3.19) if and only if it is a solution to

$$a'x \equiv b' \bmod n'.$$

Deduce that (3.19) has a solution $x = s \in \mathbb{Z}$ and that the solutions of (3.19) are given $x \equiv s \bmod n'$.

Q3.10. Solve the following pairs of simultaneous congruences.

(a)

$$x \equiv 4 \bmod 7$$
$$x \equiv 8 \bmod 13$$

(b)

$$x \equiv 7 \bmod 9$$
$$x \equiv 4 \bmod 14$$

Q3.11. Solve the following system of congruences

$$x \equiv 2 \bmod 3$$
$$x \equiv 4 \bmod 5$$
$$x \equiv 3 \bmod 8.$$

Q3.12. Show that the following pair of congruences does not have a solution

$$x \equiv 2 \bmod 4$$
$$x \equiv 3 \bmod 6.$$

Why does this not contradict the Chinese remainder theorem?

Q3.13. In this question we outline an alternative method to solve a pair of simultaneous congruences, this is in a sense equivalent to some of the other methods that we have seen.

Consider the pair of simultaneous congruences.

$$x \equiv 7 \bmod 9$$
$$x \equiv 4 \bmod 14$$

(a) Say why a solution x of these simultaneous congruences can be written both in the form $x = 7 + 9k$ for some $k \in \mathbb{Z}$ and in the $x = 4 + 14l$ for some $l \in \mathbb{Z}$.

(b) Equate the two equations in (a) and use this to obtain $3 = 14l - 9k$.

(c) Now find $u, v \in \mathbb{Z}$ such that $14u - 9v = 1$.
You can do this either using the Euclidean algorithm or just by guessing and trying a few possibilities.

(d) Verify that $l = 3u$ and $k = 3v$ give a solution to the equation in (b), and deduce that $x = 7 + 27v$ is a solution of the pair of simultaneous congruences.

(e) Without using the Chinese remainder theorem, can you explain why all solutions of the pair of simultaneous congruences are given by $x \equiv 7 + 27v \bmod 126$.

Q3.14. Prove the following lemma about existence of additive inverses in $\mathbb{Z}_n$.

Lemma. *Let $x \in \mathbb{Z}_n$. Then there exists $y \in \mathbb{Z}_n$ such that $x + y = [0]_n = y + x$.*

Q3.15. (a) Calculate the addition and multiplication tables of $\mathbb{Z}_5$

(b) Verify that for all $x \in \mathbb{Z}_5 \setminus \{[0]_5\}$, there exists $y \in \mathbb{Z}_5 \setminus \{[0]_5\}$ such that $x \cdot y = [1]_5$. In other words that every element of $\mathbb{Z}_5 \setminus \{[0]_5\}$ has a multiplicative inverse.
It is quite cumbersome to write out the notation $[a]_5$ all the time, so you can write $\overline{a}$ for $[a]_5$.

Q3.16. Let $n \in \mathbb{N}$. Prove that multiplication on $\mathbb{Z}_n$ is well defined.

Q3.17. Let $n \in \mathbb{N}$. Define $f : \mathbb{Z}_n \to \mathbb{R}$ as follows.
Let $x \in \mathbb{Z}_n$ and choose $x_0 \in \mathbb{Z}$ such that $x = [x_0]_n$. Define

$$f(x) = \sin(2\pi x_0/n).$$

Prove that f is well defined.

Q3.18. Let $a \in \mathbb{Z}$. Prove that $42 \mid a^7 - a$.

Hint: Consider Theorem 3.30 for $p = 2, 3, 7$. Also Lemma 3.17(a) may be helpful.

Q 3.19. Let $n = 561$ and let $a \in \mathbb{Z}$. Suppose that a is coprime to n. Prove that $a^{n-1} \equiv 1 \bmod n$.

Hint: First factorize 561 *as a product of primes.*
For each of the primes p in this factorization show that $a^{n-1} \equiv 1 \bmod p$ (you may want to use Theorem 3.30 (Fermat's little theorem) to do this).
Finally apply Lemma 3.17(a).

Q3.20. Let $a, b, c \in \mathbb{Z}$ and $p \in \mathbb{N}$ be prime. Suppose that $p \nmid a$ and $[a]_p \cdot [c]_p = [b]_p$. Then $[c]_p = [a^{p-2}]_p \cdot [b]_p$.

Q3.21. (a) Let $p, q \in \mathbb{N}$ be prime numbers, $N = pq$ and let $e \in \mathbb{N}$ such that e is coprime to $(p-1)(q-1)$. Let $(m_1, m_2, \ldots, m_r)$ be a sequence of natural numbers with $0 \leq m_i < N$ for $i = 1, 2, \ldots, r$.

(i) Explain how to encode the plaintext $(m_1, m_2, \ldots, m_r)$ using the RSA public key cryptosystem with public key (N, e) to get the ciphertext $(c_1, c_2, \ldots, c_r)$.

(ii) What is the private key d for the RSA public key cryptosystem with public key (N, e)? How can we calculate it?

(b) Let $p = 31$ and $q = 37$, $N = pq = 1147$, and we let $e = 463$. Consider the RSA cryptosystem with public key (N, e).

(i) Calculate the private key d for the cryptosystem.

(ii) You are sent the ciphertext $\mathbf{c} = (166, 53, 759)$. Decipher it.

It will help to use a modular arithmetic calculator for this question. Like the one on http://users.wpi.edu/~martin/mod.html

The remaining exercises are a bit more challenging, but they are very interesting.

Q3.22. Let $p \in \mathbb{N}$ be prime.
Investigate the value of $(p-1)! \pmod p$.

Work out the value of $(p-1)! \pmod p$ for small values of p. Make a conjecture, and then try to see why it is true for $p = 7$ and $p = 11$. Then try to prove it in general.
Hint: Theorem 3.13 and Corollary 3.16 will be helpful.

Q3.23. Let $n \in \mathbb{N}$. Let $a \in \mathbb{Z}_n$, we define $a^2 = a \cdot a$, also if $a = [a_0]_n \in \mathbb{Z}_n$, then we define $-a = [-a_0]_n$. We say that $a \in \mathbb{Z}_n$ is a square if there exists $b \in \mathbb{Z}_n$ such that $a = b^2$. For example, for $n = 4$, we see that $[0]_4 = ([0]_4)^2$ and $[1]_4 = ([1]_4)^2$ are squares, and from the multiplication table for $\mathbb{Z}_4$ from the lectures, we see that $[2]_4$ and $[3]_4$ are not squares. So there are 2 squares in $\mathbb{Z}_4$.

(a) Work out the number of squares in $\mathbb{Z}_n$, when $n = 3$, 5, 7 and 11.
(b) Make a conjecture about the number of squares in $\mathbb{Z}_n$ when n is an odd prime.
(c) Prove your conjecture.

Hint for (c): First note that $[0]_n$ is always a square. Next prove that $b^2 \neq [0]_n$ if $b \neq [0]_n$. Next show for $b, c \in \mathbb{Z}_n \setminus \{[0]_n\}$, we have $b^2 = c^2$ if and only $b = c$ or $b = -c$. Then finish the proof.

Q3.24. A finite sequence of natural numbers of the form

$$n, n + d, n + 2d, \ldots, n + (m - 1)d,$$

where $n, d, m \in \mathbb{N}$, is called a finite arithmetic progression, or FAP for short, of length m. A prime FAP is a FAP consisting of prime numbers. For example, $3, 5, 7$ is a prime FAP of length 3, and $5, 11, 17, 23, 29$ is a prime FAP of length 5.

(a) Let $n, n + d, n + 2d$ be a prime FAP of length 3 with $n \neq 3$. Prove that $6 \mid d$.
(b) Let $n, n + d, n + 2d, n + 3d$ be a prime FAP of length 4. Prove that $6 \mid d$.
(c) Let $n, n + d, n + 2d, n + 3d, n + 4d$ be a prime FAP of length 5 with $n \neq 5$. Prove that $30 \mid d$.

Hint: For (a) you should prove that both $2 \mid d$ and $3 \mid d$, then you can use Lemma 3.17(a). To show that $3 \mid d$, you can work as follows:
Suppose for a contradiction that $3 \nmid d$. Note that n is congruent modulo 3 to 0, 1 or 2 and d congruent modulo 3 to 1 or 2. Then, by considering all these possibilities, show that $n \equiv 0 \bmod 3$, $n + d \equiv 0 \bmod 3$ or $n + 2d \equiv 0 \bmod 3$. Then get a contradiction.
For part (c) you should try to think of a quicker way to prove that $5 \mid d$, as considering 20 cases will take a long time. Perhaps Theorem 3.13 may be helpful.

Q3.25. We use the notation from the previous question. Let $n, n + d, n + 2d, \ldots, n + (m - 1)d$ be a prime FAP of length m. Make a conjecture about which prime numbers must be factors of d, and then prove it.

Chapter 4

Permutations

In this chapter we are going to study permutations, which are bijective functions of sets. These are particularly nice and useful functions in mathematics.

We will use some results about functions that you have covered in 1ACa. There is a brief recap on functions, which includes everything that we'll need in Appendix B. Before we go on we recall here some of the key things about functions that we'll want. We start off with the definitions of identity functions and inverse functions, which are Definitions B.10 and B.12.

Definition. Let A be a set. The *identity function on* A is the function $\mathrm{id}_A : A \to A$ defined by $\mathrm{id}_A(x) = x$.

We note that id_A is clearly a bijection.

Definition. Let A and B be sets and let $f : A \to B$ be a bijection. The *inverse of* f is the function $f^{-1} : B \to A$ defined by

$$f^{-1}(x) \text{ is the unique element } y \in A \text{ such that } f(y) = x.$$

Next we state a proposition about functions that collects a few useful properties of functions. All parts of the proposition are covered within Lemmas B.9, B.6, B.11 and B.14.

Proposition 4.1. *Let A, B, C and D be sets and let $f : A \to B$, $g : B \to C$ and $h : C \to D$ be functions.*

(a) *Suppose that f and g are bijections. Then $g \circ f : A \to C$ is a bijection.*
(b) $(h \circ g) \circ f = h \circ (g \circ f)$.
(c) $f \circ \mathrm{id}_A = f$ *and* $\mathrm{id}_B \circ f = f$.
(d) *Suppose that f is a bijection then $f^{-1} : B \to A$ is a bijection, and $f^{-1} \circ f = \mathrm{id}_A$ and $f \circ f^{-1} = \mathrm{id}_B$.*

4.1 Permutations

We begin with the main definition of this chapter. It is conventional to use the Greek letter Ω (pronounced omega) for a set when working with permutations.

Definition 4.2. Let Ω be a set. A *permutation* of Ω is a bijection $\Omega \to \Omega$.
We define

$$\mathrm{Sym}(\Omega) = \{f \ : \ f \text{ is a permutation of } \Omega\}.$$

So $\mathrm{Sym}(\Omega)$ is the set of all permutations of Ω.

We give some examples.

Examples 4.3. Let $\Omega = \{1, 2, 3, 4\}$.

(a) Define $f : \Omega \to \Omega$ by

$$f(x) = \begin{cases} x+1 & \text{if } x \neq 4 \\ 1 & \text{if } x = 4. \end{cases}$$

Then f is a permutation of Ω.

(b) Define $g : \Omega \to \Omega$ by

$$g(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ x+1 & \text{if } x \text{ is odd.} \end{cases}$$

Then g is *not* a permutation of Ω, because g is not injective (or surjective).

(c) Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$. Then f is a permutation of $\mathbb{R}$.

4.2 Two-row notation

For the rest of this chapter we are only interested in permutations of finite sets. In fact we only consider sets of the form $\Omega = \{1, 2, \ldots, n\}$, where $n \in \mathbb{N}$. In this case we just write S_n for $\mathrm{Sym}(\Omega)$, and we just write id rather than id_Ω.

Below we give a convenient notation for representing permutations.

Definition 4.4. Let $n \in \mathbb{N}$ and $f \in S_n = \mathrm{Sym}(\{1, 2, \ldots, n\})$.
The *two-row notation* for f is the symbol

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}.$$

To help us understand this we give some examples.

Examples 4.5.

(a) Let $f \in S_4$ be as in Examples 4.3(a). Then the two-row notation for f is

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

(b) Define $g \in S_5$ by

$$g(1) = 3,\ g(2) = 2,\ g(3) = 5,\ g(4) = 4,\ g(5) = 1.$$

Then the two-row notation for g is

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 4 & 1 \end{pmatrix}.$$

(b) We can list all 6 elements of S_3:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

As we can see in the examples the second row in the two-row notation for $f \in S_n$ is a rearrangement of $1, 2, \ldots, n$. From this we can work out that the number of permutations of $\{1, 2, \ldots, n\}$ is $n!$. Therefore, we have $|S_n| = n!$.

4.3 Composition

By Proposition 4.1(a) we know that the composition of two permutations is a permutation. Therefore, if $f, g \in \mathrm{Sym}(\Omega)$, then $g \circ f \in \mathrm{Sym}(\Omega)$, where Ω is a a set. An alternative way of saying this is to say that $\mathrm{Sym}(\Omega)$ is *closed under composition*.

In the example below we show how to work out the composition of two permutations using the two-row notation.

Example 4.6. Let $f, g \in S_5$ with two-row notation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}.$$

and

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}.$$

Then we have

$$g \circ f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$$

To work this out we can write

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ f(1) & f(2) & f(3) & f(4) & f(5) \\ g(f(1)) & g(f(2)) & g(f(3)) & g(f(4)) & g(f(5)) \end{pmatrix}.$$

We obtain the bottom two rows by rearranging the columns of g so that the top row of g is the same as the bottom row of f. Then we remove the middle row. It's not necessary for you to write the middle row, if you can do the calculation without it.

Also we can calculate

$$f \circ g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix}.$$

We observe that $g \circ f \neq f \circ g$, so composition of permutations is not commutative.

4.4 Inversion

Let f be a permutation of a set Ω. Then f is a bijection of Ω, so f has an inverse f^{-1} and f^{-1} is a bijection, by Proposition 4.1(d). Therefore, $f^{-1} \in \mathrm{Sym}(\Omega)$. We demonstrate how to work out the inverse of a permutation in two-row notation in the example below.

Example 4.7. Let $f \in S_4$ with two-row notation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

Then we have

$$f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

To work this out we can swap the rows of f and write

$$\begin{pmatrix} 2 & 4 & 3 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} f(1) & f(2) & f(3) & f(4) \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

then rearrange the columns so that the top row is 1234; though you can, and may prefer, to do it directly.

4.5 Powers of a permutation

We now define powers of permutations in a very similar way to how we define powers of real numbers, just that we are using composition rather than multiplication.

Definition 4.8. Let Ω be a set, $f \in \mathrm{Sym}(\Omega)$, and $r \in \mathbb{Z}$.
We define f^r as follows.

- For $r = 0$, we set $f^0 = \mathrm{id}_\Omega$.
- For $r > 0$, we set $f^r = f \circ f \circ \cdots \circ f$, where there are r factors all equal to f.
- For $r < 0$, we let $s = -r$, so $s > 0$ and then set $f^r = (f^{-1})^s$.

So we have

$$f^1 = f,\ f^2 = f \circ f,\ f^3 = f \circ f \circ f,\ f^4 = f \circ f \circ f \circ f, \ldots,$$

and

$$f^{-2} = f^{-1} \circ f^{-1},\ f^{-3} = f^{-1} \circ f^{-1} \circ f^{-1}, \ldots.$$

We give an example of taking powers.

Example 4.9. Let $f \in S_4$ with two-row notation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}.$$

Then we have

$$f^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$

$$f^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix},$$

and

$$f^4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

So $f^4 = \mathrm{id}$.

Also we can work out that $f^{-1} = f^3$. Can you explain why?

The next lemma says that powers of permutation have similar properties to powers of real numbers. We omit the proof, as this can be done similarly to how it would be proved for powers of numbers.

Lemma 4.10. *Let Ω be a set, $f \in \mathrm{Sym}(\Omega)$, and $r, s \in \mathbb{Z}$. Then*

(a) $f^{r+s} = f^r \circ f^s$; *and*
(b) $f^{rs} = (f^r)^s$.

4.6 Cycles

In this section we define cycles. This leads to an alternative convenient way of thinking about permutations that we develop in the next section.

Definition 4.11. Let Ω be a set, and $f \in \mathrm{Sym}(\Omega)$.
We say that f is a *cycle* of length k, if there exist distinct elements $a_1, a_2, \ldots, a_k \in \Omega$ such that

$$f(a_1) = a_2,\ f(a_2) = a_3,\ \ldots,\ f(a_{k-1}) = a_k,\ f(a_k) = a_1,$$

and $f(a) = a$ for all $a \in \Omega \setminus \{a_1, a_2, \ldots, a_k\}$.

We use the notation

$$f = (a_1\ a_2\ \ldots\ a_k)$$

It is best to understand this definition through some examples.

Example 4.12. .

(a) Let $f \in S_5$ with two-row notation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}.$$

Then f is a cycle of length 4, because

$$f(1) = 2,\ f(2) = 5,\ f(5) = 4,\ f(4) = 1,$$

and $f(3) = 3$. So

$$f = (1\,2\,5\,4).$$

(b) Let $g \in S_5$ with two-row notation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 2 & 4 & 3 \end{pmatrix}.$$

Then g is a cycle of length 3 and

$$g = (2\,5\,3).$$

4.7 Cycle decomposition and cycle notation

In Theorem 4.15 below we state and then sketch a proof that any permutation can be written as a product of disjoint cycles. Before this we demonstrate it with an example.

Example 4.13. We are going to express

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 8 & 3 & 2 & 5 & 1 & 6 & 9 & 4 & 7 \end{pmatrix}$$

as a product of disjoint cycles.

First we look at the sequence

$$1,\ f(1),\ f^2(1), \cdots = 1, 8, 4, 5, 1, 8, \ldots.$$

This give us our first cycle

$$(1\,8\,4\,5).$$

Next we look at

$$2,\ f(2),\ f^2(2), \cdots = 2, 3, 2, 3, 2, 3, \ldots.$$

This gives our second cycle

$$(2\,3).$$

Next we look at

$$6,\ f(6),\ f^2(6), \cdots = 6, 6, 6, \ldots.$$

So 6 is a fixed point of 6, and we view it as a cycle of length 1. So we have our third cycle

$$(6).$$

Looking at

$$7,\ f(7),\ f^2(7), \cdots = 7, 9, 7, 9, 7, 9 \ldots.$$

This gives our last cycle

$$(7\,9).$$

Thus we have decomposed f as a product of cycles:

$$f = (1\,8\,4\,5) \circ (2\,3) \circ (6) \circ (7\,9).$$

The cycle type of f is the symbol $1^1 2^2 4^1$, which tells us the length of the cycles in this decomposition of f.

Before stating Theorem 4.15 we need to say what a "product of disjoint cycles" means.

Definition 4.14. Let $n \in \mathbb{N}$, and let $c_1, c_2, \ldots, c_r \in S_n$ be cycles:

$$c_i = (a_{i,1}\ a_{i,2}\ \ldots\ a_{i,k_r}).$$

(a) The *product of the cycles* $c_1, c_2, \ldots, c_r$ is just their composition

$$c_1 \circ c_2 \circ \cdots \circ c_r.$$

(b) We say that the cycles $c_1, c_2, \ldots, c_r$ are *disjoint* if

$$a_{i,j} \neq a_{k,l}$$

whenever $i \neq k$. So this means that no two cycles contain an entry in common.

We do not include all of the details of the proof of Theorem 4.15 below, so we only call it a sketch proof; in particular, we don't explain why the cycles are disjoint in the sketch proof.

Theorem 4.15. *Let $n \in \mathbb{N}$ and $f \in S_n$. Then f can be written a product of disjoint cycles.*

Sketch proof. We construct the cycle decomposition as follows.
First we form the sequence

$$a_1 = 1,\ a_2 = f(a_1)\ a_3 = f(a_2),\ \ldots.$$

Since Ω is finite, we must have $a_{k+1} = a_1$ for some $k \in \mathbb{N}$, and we can choose k be to be minimal. Then we let c_1 be the cycle of length k:

$$c_1 = (a_1\ a_2\ \ldots\ a_k).$$

If $k = n$, then we see that $f = c_1$. So we have written f as a product of disjoint cycles.
So suppose $k \neq n$, then we can pick $i \in \{1, 2, \ldots, n\} \setminus \{a_1, a_2, \ldots, a_k\}$ to be minimal. We define another cycle

$$c_2 = (a'_1\ a'_2\ \ldots\ a'_{k'}),$$

where

$$a'_1 = i,\ a'_2 = f(a'_1)\ a'_3 = f(a'_2),\ \ldots.$$

We can check that c_1 and c_2 are disjoint.
If $k + k' = n$, then we see that $f = c_1 \circ c_2$. So we have written f as a product of disjoint cycles.
Continuing in this way, we will eventually have written f as a product of disjoint cycles. □

Armed with Theorem 4.15, we can now define the cycle notation and cycle type of a permutation.

Definition 4.16. Let $n \in \mathbb{N}$, and $f \in S_n$.

(a) The *cycle notation* of f is the decomposition of f as a product of disjoint cycles:

$$f = c_1 \circ c_2 \circ \cdots \circ c_r.$$

(b) The *cycle type* of f is the symbol

$$1^{m_1} 2^{m_2} \ldots n^{m_n},$$

where m_i is the number of cycles of length i, and we do not write i^{m_i} if $m_i = 0$.

We give some more examples of cycle decompositions.

Examples 4.17.

(a) Let

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 3 & 6 & 1 & 4 \end{pmatrix}.$$

Then the cycle notation of f is

$$f = (1\,2\,5) \circ (3) \circ (4\,6).$$

So the cycle type of f is $1^1 2^1 3^1$.

(b) Let

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 1 & 5 & 3 & 6 \end{pmatrix}.$$

Then the cycle notation of g is

$$g = (1\,4\,5\,3) \circ (2) \circ (6).$$

So the cycle type of g is $1^2 4^1$.

(c) We can write down all $6 = 3!$ the elements of S_3 in cycle notation. They are id, (12)(3), (13)(2), (1)(23), (123) and (132).

The cycle notation for a permutation is not unique. First any of the cycles can begin with any element in it and secondly the disjoint cycles can be rearranged. For instance, with f as in Example 4.17, we could write the cycle notation as

$$f = (3) \circ (6\,4) \circ (5\,1\,2).$$

The order of the disjoint cycles can be changed because *disjoint cycles commute*; but beware that in general cycles do not commute.

4.8 Calculating in cycle notation

Here we look at calculating compositions and inverses of permutations in cycle notation. We just do this by looking at one example. The main idea of how to do this is by talking to yourself, as we'll see. There are more examples for you to try in exercise Q4.3.

Example 4.18. Let

$$f = (124) \circ (365), \quad g = (1326) \circ (45), \quad h = (153) \circ (2) \circ (46) \quad \in S_6.$$

We're going to work out $g \circ f$, which we do by saying.
Well f sends 1 to 2 and g sends 2 to 6, so $g \circ f$ sends 1 to 6.
Next we consider 6, and say that f sends 6 to 5 and g sends 5 to 4, so $g \circ f$ sends 6 to 4.
Next we consider 4, and say that f sends 4 to 1 and g sends 1 to 3, so $g \circ f$ sends 4 to 3.
Next we consider 3, and say that f sends 3 to 6 and g sends 6 to 1, so $g \circ f$ sends 3 to 1.
Thus we get that (1643) is a cycle in $g \circ f$.

Now we consider 2, and say that f sends 2 to 4 and g sends 4 to 5, so $g \circ f$ sends 2 to 5.
Next we consider 5, and say that f sends 5 to 3 and g sends 3 to 2, so $g \circ f$ sends 5 to 2.
Thus we get that (25) is a cycle in $g \circ f$.
Hence,

$$g \circ f = (1643) \circ (25).$$

Let's do $h \circ g$ too, to help us to get used to this. First we can write out the cycle notation of h and g next to each other to denote their composition.

$$h \circ g = (153) \circ (2) \circ (46) \circ (1326) \circ (45)$$

Then going along the cycles from right to left we say:
1 goes to 3 goes to 1.
So (1) is a cycle in $h \circ g$.
2 goes to 6 goes to 4.
4 goes to 5 goes to 3.
3 goes to 2 goes to 2.
So (243) is a cycle in $h \circ g$.
5 goes to 4 goes to 6.
6 goes to 1 goes to 5.
So (56) is a cycle in $h \circ g$.
Hence,

$$h \circ g = (1) \circ (243) \circ (56).$$

As a last example on composing in cycle notation we'll do $g \circ h$. First write them next to each other to denote their composition.

$$g \circ h = (1326) \circ (45) \circ (153) \circ (2) \circ (46).$$

Then going along the cycles from right to left we say:
1 goes to 5 goes to 4.
4 goes to 6 goes to 1.
So (14) is a cycle in $g \circ h$.
2 goes to 2 goes to 6.
6 goes to 4 goes to 5.
5 goes to 3 goes to 2.
So (265) is a cycle in $g \circ h$.
3 goes to 1 goes to 3.
So (3) is a cycle in $g \circ h$.
Hence,

$$g \circ h = (14) \circ (265) \circ (3).$$

In these calculations you may find it a bit unnatural that we have to read the cycles from right to left. This is because when we write a composition like $f \circ g$ it means do g and then f, so we are going from right to left. Sometimes functions are "written on the right" to make this more natural, but we choose not to do that here, though you'll possibly see this in some books and in future courses.

The last thing we'll do in this example is to work out f^{-1}, g^{-1} and h^{-1}.
To work out f^{-1}. We say:

Well 1 is the image of 4 under f, so f^{-1} sends 1 to 4.
Next we say that 4 is the image of 2 under f, so f^{-1} sends 4 to 2.
Next we say that 2 is the image of 1 under f, so f^{-1} sends 2 to 1.
Thus (142) is a cycle in f^{-1}. Similarly we obtain that (356) is a cycle in f^{-1}. Hence,

$$f^{-1} = (142) \circ (356).$$

Note that $(142) = (421)$, because we can change which element we write first in the cycle, and similarly $(356) = (563)$. Therefore, $f^{-1} = (421)(563)$. So that we obtain f^{-1} by reversing the order of the elements in the cycles.
In fact this method of reversing the order of the elements in the cycles work for finding the inverse of any permutation, you should convince yourself of this. In particular, we obtain

$$g^{-1} = (6231) \circ (54) = (1623) \circ (45) \quad \text{and} \quad h^{-1} = (351) \circ (2) \circ (64) = (135) \circ (2) \circ (46).$$

4.9 The sign of a permutation

Please note that this section has recently been typed so is more likely to contain typos.

We're going to cover a more subtle aspect of theory of permutations in this section, which may take a bit more time to grasp. This is the sign of a permutation, which is defined in Definition 4.20 below. Before we get on to the definition it will help to demonstrate the idea with an example.

Example 4.19. Let X_1, X_2 and X_3 be indeterminates. We can let permutations in S_3 act on these indeterminates in the same way that they act on the numbers 1, 2 and 3. By this we mean for $f \in S_3$ and $i \in \{1, 2, 3\}$, we say that f sends X_i to $X_{f(i)}$. This can be extended to polynomials in X_1, X_2 and X_3 and given a polynomial $M = m(X_1, X_2, X_3)$ we define $f(M) = m(X_{f(1)}, X_{f(2)}, X_{f(3)})$. There's quite a lot being defined here, so lets see a couple of examples: for $f = (123)$ and $M = X_1X_2 - X_2X_3^2$, we have $f(M) = X_2X_3 - X_3X_1^2 = X_2X_3 + X_1^2X_3$; and for $g = (23)$ and $N = X_1^2X_3 + X_1X_2X_3$, we have $g(N) = X_1^2X_2 + X_1X_3X_2 = X_1^2X_2 + X_1X_2X_3$.

We are particularly interested in the polynomial

$$\Delta_3 = (X_1 - X_2)(X_1 - X_3)(X_2 - X_3).$$

Let's see what happens when we apply each of the elements of S_3 to Δ_3. Clearly we have

$$\mathrm{id}(\Delta_3) = \Delta_3.$$

Next we consider $f = (12)$, and we calculate

$$\begin{aligned} f(\Delta_3) &= (X_2 - X_1)(X_2 - X_3)(X_1 - X_3) \\ &= (X_2 - X_1)(X_1 - X_3)(X_2 - X_3) \\ &= -(X_1 - X_2)(X_2 - X_3)(X_1 - X_3) \\ &= -\Delta_3. \end{aligned}$$

We got from the first line to the second line, by rearranging the factors and from the second line to the third by using $X_2 - X_1 = -(X_1 - X_2)$.
Let's also consider $f = (123)$, and we calculate

$$\begin{aligned} f(\Delta_3) &= (X_2 - X_3)(X_2 - X_1)(X_3 - X_1) \\ &= (X_2 - X_1)(X_3 - X_1)(X_2 - X_3) \\ &= (X_1 - X_2)(X_2 - X_3)(X_1 - X_3) \\ &= \Delta_3. \end{aligned}$$

We can do all of the permutations in a similar way and we can summarize what we find in the table below.

f	id	$(12)(3)$	$(13)(2)$	$(1)(23)$	(123)	(132)
$f(\Delta_3)$	Δ_3	$-\Delta_3$	$-\Delta_3$	$-\Delta_3$	Δ_3	Δ_3.

So we see that $f(\Delta_3)$ is always equal to either Δ_3 or $-\Delta_3$. To see why this occurs note that when apply $f \in S_3$ to Δ_3, we obtain

$$f(\Delta_3) = (X_{f(1)} - X_{f(2)})(X_{f(1)} - X_{f(3)})(X_{f(2)} - X_{f(3)}).$$

Then we observe that the factors $X_i - X_j$, for $(i, j) \in \{(1, 2), (1, 3), (2, 3)\}$ have been permuted, but some of the factors have been reversed from $X_i - X_j$ to $X_j - X_i = -(X_i - X_j)$.

Now let's consider general $n \in \mathbb{N}$. We can consider the polynomial

$$\Delta_n = \prod_{1 \leq i < j \leq n} (X_i - X_j) \tag{4.1}$$

in the indeterminates $X_1, X_2, \ldots, X_n$. The symbol $\prod$ here means the product of all the terms; similarly to how we use the symbol $\sum$ to denote a sum. Then for $f \in S_n$ we can define

$$f(\Delta_n) = \prod_{1 \leq i < j \leq n} (X_{f(i)} - X_{f(j)}) \tag{4.2}$$

For similar reason to those given in Example 4.19 (for the case $n = 3$), we always have $f(\Delta_n) = \pm\Delta_n$. This allows us define the sign and parity of an element of S_n.

Definition 4.20. Let $n \in \mathbb{N}$ and let $f \in S_n$.

We define the *sign* of f to be the number $\operatorname{sgn}(f) \in \{1, -1\}$ such that $f(\Delta_n) = \operatorname{sgn}(f)\Delta_n$, where Δ_n is defined in (4.1) and $f(\Delta_n)$ is defined in (4.2).

We define the *parity* of f by saying that f is *even* if $\operatorname{sgn}(f) = 1$ and f is *odd* if $\operatorname{sgn}(f) = -1$.

Now that we have the definition of the sign of a permutation, we move on to consider how to calculate it examples. The first step is the next lemma about the sign of the composition of permutations.

Lemma 4.21. *Let $n \in \mathbb{N}$ and let $f, g \in S_n$. Then $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f)\operatorname{sgn}(g)$.*

Proof. We calculate $(f \circ g)(\Delta_n)$. On the one hand we get

$$(f \circ g)(\Delta_n) = \operatorname{sgn}(f \circ g)\Delta_n,$$

and other the other hand we get

$$\begin{aligned}(f \circ g)(\Delta_n) &= f(g(\Delta_n)) \\ &= f(\operatorname{sgn}(g)(\Delta_n)) \\ &= \operatorname{sgn}(g) f(\Delta_n)) \\ &= \operatorname{sgn}(g)\operatorname{sgn}(f)\Delta(n).\end{aligned}$$

Hence, $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f)\operatorname{sgn}(g)$. □

We move on to determine the sign of a 2-cycle; often we refer to a 2-cycle as a *transposition*.

Lemma 4.22. *Let $n \in \mathbb{N}$ and let $f = (kl) \in S_n$ be a transposition, where $k, l \in \{1, 2, \ldots, n\}$. Then* $\operatorname{sgn}(f) = -1$.

Proof. We may assume that $k < l$. We consider $f(\Delta_n) = \prod_{1 \leq i < j \leq n}(X_{f(i)} - X_{f(j)})$. The factors $X_i - X_j$ for $i < j$ all occur in $f(\Delta_n)$ with some of them reversed. The factors that are reversed are

$$X_k - X_{k+1}, X_k - X_{k+2}, \ldots, X_k - X_l,$$

$$X_{k+1} - X_l, X_{k+1} - X_l, X_{l-1} - X_l.$$

So there are $(l-k) + (l-k) - 1 = 2(l-k) - 1$ such factors, which is an odd number. Hence, we obtain that $f(\Delta_n) = -\Delta_n$, so that $\operatorname{sgn}(f) = -1$. □

Using the previous two lemmas, we are now in a position to determine the sign of any cycle.

Lemma 4.23. *Let $n, k \in \mathbb{N}$ with $k \leq n$, and let $f \in S_n$ be a k-cycle. Then* $\operatorname{sgn}(f) = (-1)^{k-1}$.

Proof. We have $f = (a_1 a_2 \ldots a_k)$ for some a_i, and we observe that we can write f as a product of $k - 1$ transpositions:

$$f = (a_1 a_2) \circ (a_2 a_3) \circ \cdots \circ (a_{k-2} a_{k-1}) \circ (a_{k-1} a_k).$$

Now the lemma follows from Lemmas 4.21 and 4.22. □

We are now in a position to assemble the ingredients from the previous three lemmas to obtain a general formula for the sign of a permutation in terms of its cycle type.

Proposition 4.24. *Let $n \in \mathbb{N}$ and let $f \in S_n$ with cycle type $1^{m_1} 2^{m_2} \ldots n^{m_n}$. Let $m = m_2 + m_4 + \cdots + m_{2l}$ where $2l = n$ is n if n is even and $2l = n - 1$ if n is odd; so m is the number of cycles in the cycle notation of f with even length. Then* $\operatorname{sgn}(f) = (-1)^m$.

Proof. This follows immediately from Lemma 4.21 and 4.23. □

We mention an alternative interpretation of the parity of a permutation. It follows from the proofs given above that any permutation $f \in S_n$ can be written as a product of the transpositions. Then the parity of f is even if there are an even number of transpositions in this product and odd if there are an odd number of transpositions. We need to use the proofs above to check that this does not depend on the way we write f as a product of transpositions.

We end this section with some examples where we work out the signs of some permutations.

Examples 4.25. (a) Let $f = (16) \circ (279) \circ (3845) \in S_9$. Then f is even.
(b) Let $g = (13) \circ (27) \circ (4865) \in S_8$. Then g is odd.

You may wonder why we have made quite a lot of fuss about the sign of a permutation, as it may not seem that useful straightaway. However, you should rest assured that this is something important that you're likely to encounter again in your further studies. For instance it is needed to work with determinants. Also it is important in group theory, see at by Q4.5, and there is a fun application of it in Q4.6.

4.10 Summary of Chapter 4

By the end of this chapter you should be able to:

- calculate the two-row notation of a permutation;
- calculate compositions, inverses and powers of permutations in two-row notation;
- calculate the cycle notation and cycle type of a permutation; and
- calculate compositions, inverses and powers of permutations in cycle notation.

4.11 Exercises for Chapter 4

Q4.1. Let

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 1 & 4 & 2 & 3 \end{pmatrix} \quad \text{and} \quad g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 6 & 1 & 4 \end{pmatrix}$$

be elements of S_6 in two-row notation.

Calculate each of the following permutations giving your answer in two-row notation.

(a) $f \circ g$
(b) $g \circ f$
(c) f^{-1}
(d) g^2
(e) $f^{-2} \circ g^3$
(f) f^5
(g) $g^2 \circ f^2$
(h) $(g \circ f)^2$

Q4.2. Determine the cycle notation and cycle type for

(a)
$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 2 & 7 & 8 & 4 & 1 & 3 & 9 & 6 \end{pmatrix}$$

(b)
$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 4 & 7 & 5 & 2 & 6 & 3 & 1 & 8 \end{pmatrix}$$

(c)
$$h = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 7 & 8 & 3 & 6 & 1 & 5 & 9 & 2 \end{pmatrix}$$

(d)
$$k = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 6 & 3 & 9 & 2 & 1 & 8 & 7 & 5 \end{pmatrix}$$

Q4.3. Let

$$f = (14)(253) \quad \text{and} \quad g = (1534)(2) \quad \text{and} \quad h = (142)(35)$$

be elements of S_5 in cycle notation.

Calculate the following permutations giving your solution in cycle notation.

(a) fg
(b) fh
(c) gh
(d) h^2
(e) g^3
(f) f^{-1}
(g) g^{-1}
(h) h^{-3}

Q4.4. For each of the permutations in Q4.2 determine whether they are even or odd.

Q4.5. Let $n \in \mathbb{N}$ with $n \geq 2$. Let $A_n = \{f \in S_n \ : \ f \text{ is even}\}$.

(a) Show that the number elements of A_n is $\frac{n!}{2}$.
Hint: Show that if $f \in A_n$, then $(12)f$ is odd. Then show that the function $f \mapsto (12)$ is a bijection from A_n to $S_n \setminus A_n = \{f \in S_n \ : \ f \text{ is odd}\}$.

(b) Prove that A_n is a group.
You may want to ask for a hint about how to do this.

Q4.6. The "fifteen puzzle" consists of 15 square blocks labelled $1, 2, \ldots, 15$ arranged in a 4×4 frame, with one space. We can

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & \square \end{array}$$

where the square denotes the space.
We can move the pieces around by sliding them in to the empty space.
Therefore, the possible moves can be viewed as elements of $\mathrm{Sym}(\{1, 2, \ldots, 15, \square\})$, which we can think of as the same as S_{16} (where we write $\square$ instead of 16).

(a) Consider a sequence of slides after which the empty space is back in its initial position at the bottom right. Show that this corresponds to an even permutation in S_{16}.

(b) Deduce that it is not possible to rearrange the puzzle to the configuration

$$\begin{array}{cccc} 15 & 14 & 13 & 12 \\ 11 & 10 & 9 & 8 \\ 7 & 6 & 5 & 4 \\ 3 & 2 & 1 & \square \end{array}$$

Hint: For (a) observe that any single slide corresponds to a transposition.
Then think about the total number of horizontal slides, and the total number of vertical slides that would be used if the space ends up in its initial place.

Chapter 5

Groups

The last chapter of these notes will be an introduction to group theory. It is being written at the moment and will be added here some time soon.

Appendix A

Equivalence relations

As we have used the theory of equivalence relations in the construction of the ring of integers modulo n in Sections 3.5 and 3.6, we include a recap on equivalence relations in this appendix. You learned about equivalence relations in 1ACa, and covered the material here. It is convenient to have it here, as we can refer to it in Sections 3.5 and 3.6.

Recall that equivalence relations are a special type of relation that occur in lots of areas of mathematics. Roughly the idea behind equivalence relations is that sometimes we want to consider certain elements of a set to be "equivalent" – even though they are not actually equal – and equivalence relations give us a way to do this. The material in this appendix should help to make more sense of the previous sentence.

A.1 Relations

We begin with the definition of a relation. This definition may seem quite abstract at first, and we can work with the more informal description of a relation given after the formal definition. In Definition A.1 we use the notation $A \times A$ for the set of ordered pairs of elements of a set A, i.e. $A \times A = \{(a, b) \,:\, a, b \in A\}$.

Definition A.1. Let A be a set. A *relation* on A is a subset $R \subseteq A \times A$. For $a, b \in A$, we write aRb to mean $(a, b) \in R$.

Informally a relation R on a set A is a way of comparing two elements of A, where there is a "rule" to determine if two elements are related. They are best understood through examples, so we give some examples of relations below.

Examples A.2.

(a) $<$ (is less than) is a relation on $\mathbb{R}$.
(b) $=$ (is equal to) is a relation on any set A.
(c) Define the relation R on $\mathbb{Z}$ by xRy means xy is a perfect square.
So for example $2R2$, $12R3$ are true, but $2R3$ and $4R7$ are not true.
(d) Define the relation R on $\mathbb{R}$ by xRy means $xy \in \mathbb{Z}$.
So for example $\sqrt{2}R\sqrt{2}$, but $eR\pi$ is not true.
(e) Let A be the set of all triangles in $\mathbb{R}^2$.
Define the relation R on A by xRy means x and y are similar triangles.

(f) Let A be the set of students at the University of Birmingham.
Define the relation R on A by xRy means x and y were born in the same month.

(g) Let A be the set of all people on Earth.
Define the relation R on A by xRy means that x and y have a common ancestor.
We note that this example connects the mathematical definition of a relation to the word we use in English for family members.

In these examples, we see that there is a general form for defining a relation on a set. We define a relation R on a set A by writing

xRy means *some statement involving* x *and* y

So given $a, b \in A$, we can decide whether aRb by substituting $x = a$ and $y = b$ into the *statement involving* x *and* y and deciding if the resulting statement is true. For instance in Examples A.2(c), we can decide that $18R2$ is true, by checking that $18 \cdot 2 = 36 = 6^2$ is a perfect square.

This is similar to how we define a function $f : A \to B$, by writing

$f(x) =$ *some expression involving* x

To work out $f(a)$ for $a \in A$, we then substitute $x = a$ into the *expression involving* x and work out what we get. For example, we can define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 5x^2 \cos(x)$ and then we have $f(\pi) = -5\pi^2$.

When we define the function f, there is nothing special about the dummy variable x that we use. We could equally well use another letter, so for example we could define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = 5t^2 \cos(t)$. Similarly, we could use different dummy variables to define a relation. For example, in Examples A.2(d), we could define R by saying sRt means $st \in \mathbb{Z}$.

Also there is nothing special about the letter R for a relation, we could equally well use another letter. Later we will use the symbol $\sim$ for a relation.

A.2 Special types of relation

In this section we consider some properties satisfied by certain relations. First we define what it means for a relation to be reflexive.

Definition A.3. Let R be a relation on a set A. We say that R is *reflexive* if:

for all $a \in A$, aRa.

Examples A.4.

(a) The relation $\leq$ on $\mathbb{R}$ is reflexive.

(b) Define the relation R on $\mathbb{Z}$ by xRy means $|x - y| = 1$. Then R is not reflexive.
Counterexample. $|0 - 0| = 0 \neq 1$, so $0R0$ is not true.

(c) Define the relation R on $\mathbb{R}$ by xRy means $x + y \in \mathbb{Z}$. Then R is not reflexive.
Counterexample. $\pi + \pi \notin \mathbb{Z}$, so $\pi R \pi$ is not true.

(d) Let A be the set of all people living in the UK. Define the relation R on A by xRy means x and y have the same father. Then R is reflexive.

Observe that in the examples above when we have shown that a relation is not reflexive we have given a counterexample. This is the way you should **always** show that a relation does not satisfy a property.

In the examples above, where the relation is reflexive, we have not given a justification as it is obvious in these cases. If you have to show that a relation satisfies a property that is not obvious, then you have to give a proof as in Examples A.6(a) below.

Next we define what it means for a relation to be symmetric.

Definition A.5. Let R be a relation on a set A. We say that R is *symmetric* if:
for all $a, b \in A$, if aRb, then bRa.

Examples A.6.

(a) Define the relation R on $\mathbb{Z}$ by xRy means $x + y = 7$. Then R is symmetric.
Proof. Suppose that $a, b \in \mathbb{Z}$ with aRb. Then $a + b = 7$.
Thus $b + a = 7$. Hence bRa. □

(b) The relation $|$ (is a factor of) on $\mathbb{Z}$ is not symmetric.
Counterexample. $1 \mid 2$, but $2 \nmid 1$.

(c) The relation $\geq$ on $\mathbb{R}$ is not symmetric.
Counterexample. $1 \geq 0$, but $0 \not\geq 1$.

(d) Let A be the set of all differentiable functions on $\mathbb{R}$. Define the relation R by fRg means $f' = g'$, i.e. f and g have the same derivative. Then R is symmetric.

Next we define what it means for a relation to be transitive.

Definition A.7. Let R be a relation on a set A. We say that R is *transitive* if:
for all $a, b, c \in A$, if aRb and bRc, then aRc.

Examples A.8.

(a) The relation $\leq$ on $\mathbb{R}$ is transitive.
Proof. Suppose that $a, b, c \in \mathbb{R}$ with $a \leq b$ and $b \leq c$. Then $a \leq c$. □

(b) The relation $|$ on $\mathbb{Z}$ is transitive by Lemma 2.3(b).

(c) Define the relation R on $\mathbb{R}$ by xRy means $|x - y| < 2$. Then R is not transitive.
Counterexample. We have $|2 - 1| = 1 \leq 2$ and $|1 - 0| = 1 \leq 2$, so $2R1$ and $1R0$.
However, $|2 - 0| = 2 \not< 2$, so $2R0$ is not true.

(d) Let A be the set of all people alive.
Define the relation R on A by xRy means that x is a child of y.
Then R is not transitive.

(e) Let A be as in (d).
Define the relation R on A by xRy means that x is a descendant of y.
Then R is transitive.

A.3 Equivalence relations

Next we define equivalence relations. We usually use the symbol $\sim$ rather than R to denote an equivalence relation.

Definition A.9. Let $\sim$ be a relation on a set A. We say that $\sim$ is an *equivalence relation* if it satisfies the three properties:

(a) for all $a \in A$, $a \sim a$. (Reflexive property)
(b) for all $a, b \in A$, if $a \sim b$, then $b \sim a$. (Symmetric property)
(c) for all $a, b, c \in A$, if $a \sim b$ and $b \sim c$, then $a \sim c$. (Transitive property)

So an equivalence relation is a relation that is reflexive, symmetric and transitive.

We now give some examples of equivalence relations. In (b), (d) and (e) it is easy to show that the examples are equivalence relations, so we do not include a proof.

Examples A.10.

(a) Let $n \in \mathbb{N}$.
Define the relation $\sim$ on $\mathbb{Z}$ by $x \sim y$ means $x \equiv y \bmod n$.
Then $\sim$ is an equivalence relation, by Lemma 3.5.
(b) The relation $\sim$ on $\mathbb{R}$ defined by $x \sim y$ means $x^2 = y^2$ is an equivalence relation.
(c) Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
Define the relation $\sim$ on $\mathbb{N}_0 \times \mathbb{N}_0$ by $(x, y) \sim (z, w)$ means $x + w = z + y$.
Then $\sim$ is an equivalence relation.
Proof. We have to show that $\sim$ is reflexive, symmetric and transitive.
Reflexive: Let $(a, b) \in \mathbb{N}_0 \times \mathbb{N}_0$, then $a + b = a + b$, so $(a, b) \sim (a, b)$.
Symmetric: Let $(a, b), (c, d) \in \mathbb{N}_0 \times \mathbb{N}_0$ such that $(a, b) \sim (c, d)$. Then $a + d = c + b$, so $c + b = a + d$, and $(c, d) \sim (a, b)$.
Transitive: Let $(a, b), (c, d), (e, f) \in \mathbb{N}_0 \times \mathbb{N}_0$ such that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then $a + d = c + b$ and $c + f = e + d$. Adding these equations we get

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

Cancelling the c and d, gives $a + f = e + b$, so that $(a, b) \sim (e, f)$. □
(d) Let A be the set of students at the University of Birmingham.
Define the relation $\sim$ on A by $x \sim y$ means x and y were born in the same month.
Then $\sim$ is an equivalence relation.
(e) The relation $=$ on any set A is an equivalence relation.

A.4 Equivalence classes

We now define equivalence classes, which give a notation for considering elements related under an equivalence relation as being equivalent.

Definition A.11. Let A be a set, $\sim$ an equivalence relation on A and $a \in A$.

(a) The *equivalence class of* a is defined to be

$$[a]_\sim = \{x \in A \,:\, x \sim a\}.$$

(b) The *set of equivalence classes of* $\sim$ is defined to be

$$A/\sim = \{[a]_\sim \ : \ a \in A\}.$$

Note that $A/\sim$ *is a set of subsets of* A.

The notation for an equivalence class allows to have one symbol $[a]_\sim$ to represent all elements that we are viewing as equivalent, though it is often necessary for us to remember that $[a]_\sim$ is actually a set.

In the examples below we work out the equivalence classes and the set of equivalence classes for each of the equivalence relations in Examples A.10.

Examples A.12.

(a) Let $a \in \mathbb{Z}$. Then

$$[a]_\sim = \{x \in \mathbb{Z} \ : \ x \equiv a \bmod n\}.$$

In this case, $[a]_\sim$ is the same as the congruence class of a modulo n denoted $[a]_n$, which is defined in Definition 3.21. The we see that

$$\mathbb{Z}/\sim = \mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$

(b) Let $a \in \mathbb{R}$. Then

$$[a]_\sim = \{x \in \mathbb{R} \ : \ x^2 = a^2\} = \{a, -a\}.$$

Therefore,

$$\mathbb{R}/\sim = \{[a]_\sim \ : \ a \in \mathbb{R}, a \geq 0\}.$$

(c) Let $(a, b) \in \mathbb{N}_0 \times \mathbb{N}_0$. Then

$$[(a,b)]_\sim = \{(x,y) \in \mathbb{N}_0 \times \mathbb{N}_0 \ : \ x + b = a + y\},$$

so $[(a,b)]_\sim$ is the set of pairs (x, y) such that $x - y = a - b$. For $n \in \mathbb{Z}$, we define

$$C_n = \{(x,y) \in \mathbb{N}_0 \times \mathbb{N}_0 \ : \ x - y = n\}.$$

Then C_n is an equivalence class and we have

$$(\mathbb{N}_0 \times \mathbb{N}_0)/\sim = \{C_n \ : \ n \in \mathbb{Z}\}.$$

This example actually gives the way that the integers can be defined from the natural numbers.

(d) Let $a \in A$. Then $[a]_\sim$ is the set of all students who have their birthday in the same month as a. For m a month of the year we define

$$C_m = \{x \in A \ : \ x \text{ was born in } m\}.$$

Then C_m is an equivalence class and we have

$$A/\sim = \{C_m \ : \ m \text{ is a month of the year}\}.$$

(e) Let $a \in A$. Then $[a]_\sim = \{a\}$ and $A/\sim = \{[a]_\sim \ : \ a \in A\}$.

A.5 Partitions

We introduce partitions in this section. Later in Theorem A.15 we prove a connection between equivalence relations and partitions.

Definition A.13. Let A be a set. A *partition of* A is a set P of non-empty subsets of A such that:

(a) for all $x \in A$ there exists $B \in P$ such that $x \in B$; and
(b) for all $B, C \in P$, either $B = C$ or $B \cap C = \varnothing$.

This means that a partition P of A is a set of nonempty subsets of A such that every element of A is an element of exactly one element of P. Informally it is a way to split up A into non-overlapping pieces.

As usual it is helpful to understand partitions through some examples.

Examples A.14.

(a) Let $A = \{1, 2, 3, 4, 5\}$. Then:
  (i) $P = \{\{1, 3\}, \{2, 5\}, \{4\}\}$ is a partition of A.
  (ii) $P = \{\{1, 3, 4\}, \{2, 4, 5\}\}$ is not a partition of A, as 4 is in two elements of P.
  (iii) $P = \{\{1, 3\}, \{4\}, \{5\}\}$ is not a partition of A, because 2 does not lie in any element of P.

(b) Let A be any set. Then
  (i) $P = \{A\}$ is a partition of A.
  (ii) $P = \{\{a\} : a \in A\}$ is a partition of A.

A.6 Equivalence relations and partitions

We can check that each of the sets of equivalence classes in Examples A.12 is a partition. We show that this is a general property of equivalence relations in the following important theorem.

Theorem A.15. *Let A be a set, $\sim$ an equivalence relation on A, and $a, b \in A$. Then the following hold:*

(a) $a \in [a]_\sim$;
(b) $[a]_\sim = [b]_\sim$ *if and only if* $a \sim b$;
(c) $[a]_\sim = [b]_\sim$ *or* $[a]_\sim \cap [b]_\sim = \varnothing$.
(d) $A/\sim$ *is a partition of* A.

Proof. (a) By the reflexive property we have $a \sim a$. Therefore, $a \in [a]_\sim$.
(b) Suppose that $[a]_\sim = [b]_\sim$. By (a) we have $a \in [a]_\sim$, so $a \in [b]_\sim$ and $a \sim b$.
Next suppose that $a \sim b$.
Let $c \in [a]_\sim$. Then $c \sim a$, so by the transitive property we have $c \sim b$. Thus $c \in [b]_\sim$. Therefore, $[a]_\sim \subseteq [b]_\sim$.
Let $c \in [b]_\sim$. Then $c \sim b$. By the symmetry property we have $b \sim a$. Thus $c \sim a$ by the

transitive property, so $c \in [a]_\sim$. Therefore, $[b]_\sim \subseteq [a]_\sim$.
Hence, $[a]_\sim = [b]_\sim$.
(c) It suffices to show that if $[a]_\sim \cap [b]_\sim \neq \varnothing$, then $[a]_\sim = [b]_\sim$. So suppose that $[a]_\sim \cap [b]_\sim \neq \varnothing$, and let $c \in [a]_\sim \cap [b]_\sim$. Then $c \sim a$ and $c \sim b$. By the symmetry property we have $a \sim c$. Therefore, we have $a \sim c$ and $c \sim b$, so by the transitive property we have $a \sim b$. Hence, $[a]_\sim = [b]_\sim$ by (b).
(d) This follows from (a) and (c). □

Theorem A.15 shows that an equivalence relation gives rise to a partition. The following proposition, says that a partition determines an equivalence relation. The proof is exercise QA.7.

Proposition A.16. *Let A be a set and let $P \subseteq \mathcal{P}(A)$ be a partition. Define $\sim$ on A by*

$$a \sim b \text{ means there exists } B \in P \text{ such that } a, b \in B.$$

Then $\sim$ is an equivalence relation on A.

A consequence of Theorem A.15 and Proposition A.16 is that there is a correspondence between equivalence relations and partitions. Or in other words that equivalence relations and partitions are essentially the same thing – you'll have to think about what it really meant by this.

A.7 Exercises for Appendix A

QA.1. For the following relations $\sim$ determine whether they are:

(i) reflexive.
(ii) symmetric.
(iii) transitive.
(iv) an equivalence relation.

(a) Define $\sim$ on $\mathbb{N}$ by $x \sim y$ means xy is a perfect square.
(b) Define $\sim$ on $\mathbb{R}$ by $x \sim y$ means there exists $u \in \mathbb{Q}$ such that $ux = y$.

You should give a proof or counterexample for each assertion that you make. Hint: You may need Theorem 2.24 in (a)(iii).

QA.2. Define the relation $\sim$ on $\mathbb{R}^2$ by

$$(x_1, y_1) \sim (x_2, y_2) \quad \text{means} \quad x_1^2 + y_1^2 = x_2^2 + y_2^2.$$

(a) Prove that $\sim$ is an equivalence relation.
(b) Sketch the equivalence classes of $(1, 0)$ and $(1, 1)$.

QA.3. Define the relation $\sim$ on $\mathbb{Z} \times \mathbb{N}$ by

$$(a, b) \sim (c, d) \quad \text{means} \quad ad = bc$$

(a) Prove that $\sim$ is an equivalence relation.
(b) Let $E = (\mathbb{Z} \times \mathbb{N})/\sim$. Define the function $f : E \to \mathbb{Q}$ as follows. Let $x \in E$ and choose $(a, b) \in \mathbb{Z} \times \mathbb{N}$ such that $x = [(a, b)]_\sim$. Define

$$f(x) = \frac{a}{b}$$

Prove that f is well defined and is a bijection.

QA.4. Write down all the partitions of the set $\{0, 1, 2\}$.

QA.5. Let $k \in \mathbb{Z}$ and define

$$E_k = \{x \in \mathbb{R} \ : \ k \leq x < k + 1\}.$$

Prove that $\{E_k \ : \ k \in \mathbb{Z}\}$ is a partition of $\mathbb{R}$.

QA.6. Let A be a set and let $f : A \to \{1, 2, 3, 4\}$ be a function. For $i = 1, 2, 3, 4$, define

$$A_i = \{x \in A \ : \ f(x) = i\}.$$

Let $a \in A$ and $i, j \in \{1, 2, 3, 4\}$ and let $k = f(a) \in \{1, 2, 3, 4\}$. Prove that:

(a) $a \in A_k$.
(b) if $i \neq j$, then $A_i \cap A_j = \varnothing$.
(c) $\{A_1, A_2, A_3, A_4\}$ is a partition of A.

QA.7. (a) Prove Proposition A.16:

> **Proposition.** *Let A be a set and let $P \subseteq \mathcal{P}(A)$ be a partition. Define $\sim$ on A by*
>
> $$a \sim b \text{ means there exists } B \in P \text{ such that } a, b \in B.$$
>
> *Then $\sim$ is an equivalence relation on A.*

(b) Let A be a set and P a set of nonempty subsets of A. Define $\sim$ on A by

$$a \sim b \text{ means there exists } B \in P \text{ such that } a, b \in B.$$

(i) Give an example of A and P for which $\sim$ is not reflexive.
(ii) Give an example of A and P for which $\sim$ is not transitive.

You should justify your answers.

Appendix B

Functions

As we have used functions in Chapter 4 where we studied permutations, we include a recap on functions in this appendix. You learned about functions in 1ACa and in other courses, and should have covered all of the material here. As this appendix is quite brief in places you may benefit by looking in other places for more details. It is convenient to have this appendix here, as we can refer to it in Chapter 4.

B.1 Functions

We begin with the definition of a function.

Definition B.1. A *function* f consists of three things:

(a) a set $A = \mathrm{dom}(f)$ called the *domain of* f;
(b) a set $B = \mathrm{codom}(f)$ called the *codomain of* f; and
(c) a rule that assigns to each element $a \in A$ a unique element $f(a) \in B$.

We write $f : A \to B$ to mean that f is a function with domain A and codomain B, and say that f is a function from A to B.

Given $a \in A$, we say that $f(a)$ is the *image of* a *under* f.

The *image of* f is defined to be

$$\mathrm{im}(f) = \{b \in B \ : \text{ there exists } a \in A \text{ such that } b = f(a)\}.$$

We give some examples of functions.

Examples B.2. Let $A = \{2, 4, 6, 8\}$, $B = \{1, 2, 3, 4, 5\}$, $C = \{-2, -1, 0, 1, 2\}$ and $D = \{0, 1, 4\}$.

(a) Define $f : A \to B$ by
$$f(x) = \frac{x}{2} + 1.$$

(b) Define $g : B \to C$ by
$$g(x) = x - 3.$$

(c) Define $h : C \to D$ by
$$h(x) = x^2.$$

(d) Define $k : C \to B$ by

$$k(x) = x + 3.$$

We now define what it means for two functions to be equal.

Definition B.3. Let $f : A \to B$ and $g : C \to D$ be functions. We say that f *is equal to* g and write $f = g$ if the following three conditions hold:

(a) $A = C$;
(b) $B = D$; and
(c) $f(a) = g(a)$ for all $a \in A$.

We stress that the definition says that for functions to be equal they have to have the same domain and codomain. It is not enough for them to just have the same rule defining them.

B.2 Composition of functions

Below we define the composition of two functions, which just means doing one function after the other.

Definition B.4. Let $f : A \to B$ and $g : B \to C$ be functions. The *composition of* f *and* g is the function $g \circ f : A \to C$ defined by

$$(g \circ f)(x) = g(f(x))).$$

We demonstrate composition of functions in the next example.

Example B.5. For f, g, h as in Examples B.2, we have

$$(g \circ f)(x) = \frac{x}{2} - 2 \quad \text{for } x \in A. \tag{B.1}$$

and

$$(h \circ g)(x) = (x - 3)^2 \quad \text{for } x \in B. \tag{B.2}$$

The following lemma says that composition of functions is associative.

Lemma B.6. *Let $f : A \to B$, $g : B \to C$ and $h : C \to D$ be functions. Then*

$$(f \circ g) \circ h = f \circ (g \circ h).$$

Proof. Let $a \in A$. Then we have

$$((f \circ g) \circ h)(a) = (f \circ g)(h(a)) = f(g(h(a))).$$

Similarly,

$$(f \circ (g \circ h))(a) = f(g(h(a))).$$

Hence,

$$((f \circ g) \circ h)(a) = (f \circ (g \circ h))(a).$$

This holds for all $a \in A$, so

$$(f \circ g) \circ h = f \circ (g \circ h). \quad \square$$

B.3 Injections, surjections and bijections

We give the definition of an injections, surjections and bijections.

Definition B.7. Let $f : A \to B$ be a function. We say that:

(a) f is an *injection* if
for all $a, a' \in A$, if $f(a) = f(a')$, then $a = a'$.
An injection is sometimes called an injective function or a one-to-one function.

(b) f is a *surjection* if
for all $b \in B$, there exists $a \in A$ such that $f(a) = b$;
equivalently, $\mathrm{im}(f) = B$.
A surjection is sometimes called a surjective function or an onto function.

(c) f is a *bijection* if it is both an injection and a surjection.
A bijection is sometimes called a bijective function.

We demonstrate these concepts with some examples.

Example B.8. In Examples B.2, the functions f, g and k are injective, but h is not injective, because $h(1) = h(-1)$. The composition $g \circ f$ given in (B.1) is injective.

The functions g, h and k are surjective, but f is not surjective, because there is no $a \in A$ such that $f(a) = 1$. The composition $h \circ g$ given in (B.2) is surjective.

Therefore, the functions g and k are bijective, but f and h are not bijective.

The next lemma tells us that compositions of injective functions are injective, and similarly for surjective and bijective functions.

Lemma B.9. *Let $f : A \to B$ and $g : B \to C$ be functions.*

(a) *Suppose that f and g are injections. Then $g \circ f$ is an injection.*
(b) *Suppose that f and g are surjections. Then $g \circ f$ is a surjection.*
(c) *Suppose that f and g are bijections. Then $g \circ f$ is a bijection.*

Proof. (a) Let $a, a' \in A$ such that $(g \circ f)(a) = (g \circ f)(a')$.
Then $g(f(a)) = g(f(a'))$, so $f(a) = f(a')$, because g is injective.
Thus $a = a'$, because f is injective.
Hence, $g \circ f$ is injective.
(b) Let $c \in C$.
Since g is surjective, there exists $b \in B$ such that $g(b) = c$.
Since f is surjective, there exists $a \in A$ such that $f(a) = b$.
Therefore,

$$\begin{aligned}(g \circ f)(a) &= g(f(a)) \\ &= g(b) \\ &= c.\end{aligned}$$

Hence, $g \circ f$ is surjective.
(c) This follows immediately from (a) and (b). □

B.4 Identity functions and inverse functions

In this section we define identity functions and inverse functions.

Definition B.10. Let A be a set. The *identity function on* A is the function $\mathrm{id}_A : A \to A$ defined by $\mathrm{id}_A(x) = x$.

Next we give an elementary lemma about identity functions.

Lemma B.11. *Let $f : A \to B$ be a function. Then*

(a) $f \circ \mathrm{id}_A = f$; *and*
(b) $\mathrm{id}_B \circ f = f$.

Proof. (a) Let $a \in A$. Then

$$\begin{aligned}(f \circ \mathrm{id}_A)(a) &= f(\mathrm{id}_A(a)) \\ &= f(a)\end{aligned}$$

This holds for all $a \in A$, so $f = f \circ \mathrm{id}_A$.
(b) A similar argument proves that $\mathrm{id}_B \circ f = f$. □

Definition B.12. Let $f : A \to B$ be a bijection. The *inverse of* f is the function $f^{-1} : B \to A$ defined by

$$f^{-1}(x) \text{ is the unique element } y \in A \text{ such that } f(y) = x.$$

To justify this definition, we need the following two facts:

- there exists $y \in A$ such that $f(y) = x$, because f is surjective; and
- y is unique because f is injective.

In the next example with very quickly demonstrate an inverse function.

Example B.13. In Examples B.2, k is the inverse of g.

We have the following lemma giving some properties of inverses. The proof is exercise QB.2.

Lemma B.14. *Let $f : A \to B$ be a bijection. Then*

(a) *for all $a \in A$, we have $f^{-1}(f(a)) = a$, so $f^{-1} \circ f = \mathrm{id}_A$;*
(b) *for all $b \in B$, we have $f(f^{-1}(b)) = b$, so $f \circ f^{-1} = \mathrm{id}_B$;*
(c) *f^{-1} is a bijection; and*
(d) *$(f^{-1})^{-1} = f$.*

Our next lemma is about inverses of compositions.

Lemma B.15. *Let $f : A \to B$ and $g : B \to C$ be bijections. Then*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Proof. Note that by Lemma B.9, $g \circ f$ is a bijection so $(g \circ f)^{-1}$ is defined. For $x \in C$, we have

$$\begin{aligned}(g \circ f)((f^{-1} \circ g^{-1})(x)) &= (g \circ f)(f^{-1}(g^{-1}(x))) \\ &= g(f(f^{-1}(g^{-1}(x)))) \\ &= g(g^{-1}(x)) \\ &= x.\end{aligned}$$

Therefore, the definition of $(g \circ f)^{-1}$ says that

$$(g \circ f)^{-1}(x) = (f^{-1} \circ g^{-1})(x),$$

because $y = (f^{-1} \circ g^{-1})(x)$ satisfies $(g \circ f)(y) = x$. This holds for all $x \in C$, so

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \quad \square$$

Note that when we take the inverse of the composition $g \circ f$ we have to swap the order. This is similar to what happens when we take inverses of matrices, i.e. if A and B are invertible $n \times n$ matrices, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. This is no coincidence, as matrices correspond to certain functions between vector spaces.

B.5 Exercises for Appendix B

QB.1. Let $f : A \to B$ and $g : B \to C$ be functions.

(a) Suppose that $g \circ f$ is an injection. Prove that f is an injection.
(b) Suppose that $g \circ f$ is a surjection. Prove that g is a surjection.
(c) Give an example of functions f and g such that $g \circ f$ is a bijection, but neither f nor g is a bijection.

Hint: We proved that functions are injections and surjections in Lemma B.9, so this should give you an idea of how to set out your proofs.

QB.2. Prove Lemma B.14:

Lemma. *Let $f : A \to B$ be a bijection. Then*

(a) *for all $a \in A$, we have $f^{-1}(f(a)) = a$, so $f^{-1} \circ f = \mathrm{id}_A$;*
(b) *for all $b \in B$, we have $f(f^{-1}(b)) = b$, so $f \circ f^{-1} = \mathrm{id}_B$;*
(c) *f^{-1} is a bijection; and*
(d) *$(f^{-1})^{-1} = f$.*