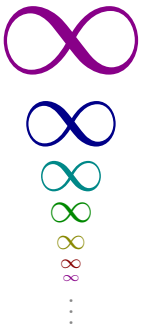
An Infinite Descent into Pure Mathematics



BY CLIVE NEWSTEAD

Version 0.7
Thursday 25th June 2026
infinitedescent.org

Version 0.7
ISBN 978-1-950215-99-7

Preview of First Edition (forthcoming)
ISBN 978-1-950215-00-3 (paperback)
ISBN 978-1-950215-01-0 (hardback)

A free PDF copy of *An Infinite Descent into Pure Mathematics* can be obtained from the book's website:
https://infinitedescent.org

0 2 4 6 8 10 9 7 5 3 1

To my parents,
Imogen and Matthew,
who taught me so much.

Contents

Preface

Hello, and thank you for taking the time to read this quick introduction to *An Infinite Descent into Pure Mathematics*! The most recent version of the book is freely available for download from the following website:

https://infinitedescent.org

The website also includes information about changes between different versions of the book, an archive of previous versions, and some resources for using LaTeX (see also Chapter D).

About the book

A student in a typical calculus class will learn the chain rule and then use it to solve some prescribed 'chain rule problems' such as computing the derivative of $\sin(1+x^2)$ with respect to x, or perhaps solving a word problem involving related rates of change. The expectation is that the student correctly apply the chain rule to derive the correct answer, and show enough work to be believed. In this sense, the student is a *consumer* of mathematics. They are given the chain rule as a tool to be accepted without question, and then use the tool to solve a narrow range of problems.

The goal of this book is to help the reader make the transition from being a *consumer* of mathematics to a *producer* of it. This is what is meant by 'pure' mathematics. While a consumer of mathematics might learn the chain rule and use it to compute a derivative, a producer of mathematics might derive the chain rule from the rigorous definition of a derivative, and then prove more abstract versions of the chain rule in more general contexts (such as multivariate analysis).

Consumers of mathematics are expected to say how they used their tools to find their answers. Producers of mathematics, on the other hand, have to do much more: they must be able to keep track of definitions and hypotheses, piece together facts in new and interesting ways, and make their own definitions of mathematical concepts. But even more importantly, once they have done this, they must communicate their findings in a way that others find intelligible, and they must convince others that what they have done is correct, appropriate and worthwhile.

It is this transition from consumption to production of mathematics that guided the principles I used to design and write this book. In particular:

- **Communication.** Above all, this book aims to help the reader to obtain mathematical literacy and express themselves mathematically. This occurs at many levels of magnification. For example,

consider the following expression:

$$\forall x \in \mathbb{R}, [\neg(x = 0) \Rightarrow (\exists y \in \mathbb{R}, y^2 < x^2)]$$

After working through this book, you will be able to say what the symbols $\forall$, $\in$, $\mathbb{R}$, $\neg$, $\Rightarrow$ and $\exists$ all mean intuitively and how they are defined precisely. But you will also be able to interpret what the expression means as a whole, explain what it means in plain terms to another person *without* using a jumble of symbols, prove that it is true, and communicate your proof to another person in a clear and concise way.

The kinds of tools needed to do this are developed in the main chapters of the book, and more focus is given to the writing side of things in Chapter A.

- **Inquiry.** The research is clear that people learn more when they find things out for themselves. If I took this to the extreme, the book would be blank; however, I do believe it is important to incorporate aspects of inquiry-based learning into the text.

  This principle manifests itself in that there are exercises scattered throughout the text, many of which simply require you to prove a result that has been stated. Many readers will find this frustrating, but this is for good reason: these exercises serve as checkpoints to make sure your understanding of the material is sufficient to proceed. That feeling of frustration is what learning feels like—embrace it!

- **Strategy.** Mathematical proof is much like a puzzle. At any given stage in a proof, you will have some definitions, assumptions and results that are available to be used, and you must piece them together using the logical rules at your disposal. Throughout the book, and particularly in the early chapters, I have made an effort to highlight useful proof strategies whenever they arise.

- **Content.** There isn't much point learning about mathematics if you don't have any concepts to prove results about. With this in mind, Part II includes several chapters dedicated to introducing some topic areas in pure mathematics, such as number theory, combinatorics, analysis and probability theory.

- **LaTeX.** The *de facto* standard for typesetting mathematics is LaTeX. I think it is important for mathematicians to learn this early in a guided way, so I wrote a brief tutorial in Chapter D and have included LaTeX code for all new notation as it is defined throughout the book.

Navigating the book

This book need not, and, emphatically *should not*, be read from front to back. The order of material is chosen so that material appearing later depends only on material appearing earlier, but following the material in the order it is presented may be a fairly dry experience.

The majority of introductory pure mathematics courses cover, at a minimum, symbolic logic, sets, functions and relations. This material is the content of Part I. Such courses usually cover additional topics from pure mathematics, with exactly *which* topics depending on what the course is preparing students for. With this in mind, Part II serves as an introduction to a range of areas of pure mathematics, including number theory, combinatorics, set theory, real analysis, probability theory and order theory.

It is not necessary to cover all of Part I before proceeding to topics in Part II. In fact, interspersing material from Part II can be a useful way of motivating many of the abstract concepts that arise in Part I.

The following table shows dependencies between sections. Previous sections within the same chapter as a section should be considered 'essential' prerequisites unless indicated otherwise.

Part	Section	Essential	Recommended	Useful
I	1.1	0		
	2.1	1.3		
	3.1	2.2		
	4.1	1.3	3.1	3.2
	5.1	2.1	3.1	3.2, 4.2
	6.1	3.2, 4.3	5.2	
II	7.1	1.3	2.1, 4.3	3.1
	7.3		5.2	
	8.1	6.1		
	9.1	4.1, 2.1		5.2
	9.2	3.1	9.1	
	9.3	3.1	9.1	7.3, 6.2
	10.1	6.2		
	10.2	8.1		
	11.1	8.1	6.2, 9.3	
	12.1	5.2		
	12.2	4.3, 3.2	6.2	12.1

Prerequisites are cumulative. For example, in order to cover Section 10.2, you should first cover Chapters 0 to 4 and Sections 6.1, 6.2, 8.1 and 10.1.

What the numbers, colours and symbols mean

Broadly speaking, the material in the book is broken down into enumerated items that fall into one of five categories: definitions, results, remarks, examples and exercises. In Chapter A, we also have proof extracts. To improve navigability, these categories are distinguished by name, colour and symbol, as indicated in the following table.

Category	Symbol	Colour	Category	Symbol	Colour
Definitions	✦	**Red**	Examples	✐	**Teal**
Results	✣	**Blue**	Exercises	✎	**Gold**
Remarks	❖	**Purple**	Proof extracts	❝	**Teal**

These items are enumerated according to their section—for example, Theorem 9.2.41 is in Section 9.2. Definitions and theorems (important results) appear in a box.

You will also encounter the symbols □ and ◁ whose meanings are as follows:

□ **End of proof.** It is standard in mathematical documents to identify when a proof has ended by drawing a small square or by writing '*Q.E.D.*' (The latter stands for *quod erat demonstrandum*, which is Latin for *which was to be shown*.)

◁ **End of item.** This is *not* a standard usage, and is included only to help you to identify when an item has finished and the main content of the book continues.

Some subsections are labelled with the symbol ⋆. This indicates that the material in that subsection can be skipped without dire consequences.

Licence

Comments and corrections

Any feedback, be it from students, teaching assistants, instructors or any other readers, would be very much appreciated. Particularly useful are corrections of typographical errors, suggestions for alternative ways to describe concepts or prove theorems, and requests for new content (e.g. if you know of a nice example that illustrates a concept, or if there is a relevant concept you wish were included in the book).

Such feedback can be sent to the author by email (cnewstead@infinitedescent.org).

Acknowledgements

When I reflect on the time I have spent writing this book, I am overwhelmed by the number of people who have had some kind of influence on its content.

This book would never have come to exist were it not for Chad Hershock's course 38-801 *Evidence-Based Teaching in the Sciences*, which I took in Fall 2014 as a graduate student at Carnegie Mellon University. His course heavily influenced my approach to teaching, and it motivated me to write this book in the first place. Many of the pedagogical decisions I made when writing this book were informed by research that I was exposed to as a student in Chad's class.

The legendary Carnegie Mellon professor, John Mackey, has been using this book (in various forms) as course notes for 21-128 *Mathematical Concepts and Proofs* and 15-151 *Mathematical Foundations of Computer Science* since Fall 2016. His influence can be felt throughout the book: thanks to discussions with John, many proofs have been reworded, sections restructured, and explanations improved. As a result, there is some overlap between the exercises in this book and the questions on his problem sheets. I am extremely grateful for his ongoing support.

Steve Awodey, who was my doctoral thesis advisor, has for a long time been a source of inspiration for me. Many of the choices I made when choosing how to present the material in this book are grounded in my desire to do mathematics *the right way*—it was this desire that led me to study category theory, and ultimately to become Steve's PhD student. I learnt a great deal from him and I greatly appreciated his patience and flexibility in helping direct my research despite my busy teaching schedule and extracurricular interests (such as writing this book).

Perhaps unbeknownst to them, many insightful conversations with the following people have helped shape the material in this book in one way or another: Jeremy Avigad, Deb Brandon, Santiago Cañez, Heather Dwyer, Thomas Forster, Will Gunther, Kate Hamilton, Jessica Harrell, Bob Harper, Brian Kell, Marsha Lovett, Ben Millwood, David Offner, Ruth Poproski, Emily Riehl, Hilary Schuldt, Gareth Taylor, Katie Walsh, Emily Weiss and Andy Zucker.

The *Stack Exchange* network has influenced the development of this book in two important ways. First, I have been an active member of *Mathematics Stack Exchange* (`https://math.stackexchange.com/`) since early 2012 and have learnt a great deal about how to effectively explain mathematical concepts; occasionally, a question on Mathematics Stack Exchange inspires me to add a new example or exercise to the book. Second, I have made frequent use of *LaTeX Stack Exchange* (`https://tex.stackexchange.com`) for implementing some of the more technical aspects of writing a book using LaTeX.

The Department of Mathematical Sciences at Carnegie Mellon University supported me academically, professionally and financially throughout my PhD and presented me with more opportunities than I could possibly have hoped for to develop as a teacher. This support is now continued by the Department of Mathematics at Northwestern University, where I am currently employed as a lecturer.

I would also like to thank everyone at Carnegie Mellon's and Northwestern's teaching centres, the Eberly Center and the Searle Center, respectively. Through various workshops, programs and fellowships at both teaching centres, I have learnt an incredible amount about how people learn, and I have transformed as a teacher. Their student-centred, evidence-based approach to the science of teaching and learning underlies everything I do as a teacher, including writing this book—their influence cannot be understated.

Finally, and importantly, I am grateful to the 1000+ students who have already used this book to learn mathematics. Every time a student contacts me to ask a question or point out an error, the book gets better; this is reflected in the dozens of typographical errors that have been fixed as a consequence.

Clive Newstead
January 2020
Evanston, Illinois

Chapter 0

Getting started

Consider the following statements:

(1) *This sentence is false.*

(2) *I have been far even as decided to use even go want to do more like.*

Are they true or false? If you think about (1) for a minute or two, then you'll get into a bit of a pickle, since both 'true' and 'false' seem to have nonsensical consequences; and (2) doesn't make any sense, so asking whether it is true or false is pointless.

Clearly we'll be wasting our time trying to write proofs of statements like the two listed above—we need to narrow our scope to statements that we might actually have a chance of proving (or perhaps refuting)! This motivates the following (informal) definition.

Definition 0.1

A **proposition** is a statement to which it makes sense to assign a **truth value** (which, for us, can be either **true** or **false**, and can't be both).

Thus the statements given at the beginning of this section are not propositions because there is no possible way of assigning them a truth value. Note that, in Definition 0.1, all that matters is that it *makes sense* to say that it is true or false, regardless of whether it actually *is* true or false—the truth value of many propositions is unknown, even very simple ones.

Exercise 0.2

Think of an example of (a) a true proposition; (b) a false proposition; and (c) a proposition whose truth value you don't know.

While propositions can be formed about any topic, the kinds of propositions that we will focus on in this textbook concern *mathematical objects*, such as numbers (representing quantities, differences, rates, and so on), shapes, sets and functions.

Crudely speaking, pure mathematics is done by defining mathematical objects and then finding the truth values of propositions about those objects. In reality there is far more to it than that—we don't

just make definitions and form propositions at random! There is usually a motivation, such as a real-world application, a problem we're trying to solve, a structure we hope to understand, or for aesthetic reasons. However, the crude 'define objects and find truth values of propositions' characterisation of pure mathematics is adequate as a starting point.

If we want to determine the truth value of a proposition, we had better agree upon a way of doing so. This is where *proofs* come in.

Definition 0.3

A **proof** of a true proposition is an argument that demonstrates its truth; the proof should start from what is already known or assumed to be true, should use logical steps that have been agreed upon as valid, and should be verifiable by a member of its intended target audience.

There is a lot going on in the definition of a proof:

- Starting from what is already known or assumed to be true: before we can prove anything, we need to agree upon some basic assumptions, such as the assumption that every object is equal to itself, and build up a body of knowledge from there. Technically these basic assumptions are known as **axioms**; we will not worry about these too much in the main body of the book, but an interested reader can find out more in Section B.1.
- Using logical steps that have been agreed upon as valid: in order to deduce the truth of new propositions, we must have a logical system of rules that tell us how to do so. In this chapter we will reason *informally*, but we will make our logical system precise in Chapter 1. Every proof we write can, in theory, be boiled down to a sequence of deductions made according to the rules outlined in that chapter.
- Verifiability: a reader from a proof's intended target audience should be able to understand why each deduction made in the proof is correct without having to ask for more information. Who the target audience is depends on the context—for example, much more detail would need to be provided when writing a proof in an introductory textbook on pure mathematics (such as this one) than in a mathematical research journal article.

Results in mathematical papers and textbooks may be referred to as *propositions*, but they may also be referred to as *theorems*, *lemmas* or *corollaries* depending on their intended usage.

- A **proposition** is an umbrella term which can be used for any result.
- A **theorem** is a key result which is particularly important.
- A **lemma** is a result that will (or can be) be used as a step in the proof of a theorem appearing later, or as a tool for the reader to use in their own proofs.
- A **corollary** is a result which follows as an immediate consequence of a theorem.

These are not precise definitions, and they are not meant to be; but using these words appropriately helps readers work out how to read a mathematical textbook or research article. For example, if you just want to skim an article and find its key results, you'd look for the theorems.

Right, then. Let's start proving some theorems!

…oh, wait a minute… we don't have anything to prove anything about yet.

That's what the rest of this chapter is for: we will preview some of the concepts that will be introduced more formally later in the book, and we will write a few starter proofs without worrying too much about the technicalities—for now. This will help prepare us for the more rigorous treatment of mathematical proof in the rest of the textbook.

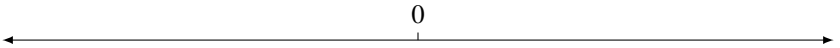
Numbers

Unsurprisingly, we will encounter numbers very often on our journey through pure mathematics. Different types of number are used for representing different kinds of quantity (amount, difference, rate, position, ...) of different kinds of object (discrete, continuous, finite, infinite, ...).

We will illustrate the different types of number of interest to us geometrically. To that end, here is an infinite line:

The arrows indicate that it is supposed to extend in both directions without end.

Now let's fix a point on this line, and label it '0':



This point can be thought of as representing the number zero. It is the point against which all other numbers will be measured. Numbers to the left of 0 on the number line are said to be *negative*, and those to the right are *positive*; 0 itself is neither positive nor negative.

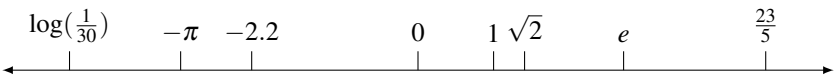
Finally, let's fix a unit of length:

This unit of length will be used, amongst other things, to compare the extent to which the other numbers differ from zero.

Definition 0.4

The above infinite line, together with its fixed zero point and fixed unit length, constitute the (**real**) **number line**. The numbers represented by points on the number line are called **real numbers**.

A few examples of real numbers and their positions on the number line are indicated in the diagram below.



For the time being, the geometric intuition provided by the real number line and our preexisting understanding of arithmetic and algebra is enough to get started with some basic proofs. However, so that we are all on the same page, some assumptions that we may make about real numbers are listed explicitly next:

Assumptions 0.5 · *Some properties of the real numbers*

(a) (**Closure**) The real numbers are closed under addition, subtraction, multiplication and division, which is to say that the sum, difference, product and ratio of any two real numbers is a real number (except where the denominator in the ratio is zero). Explicitly, whenever a and b are real numbers, so are $a+b$, $a-b$ and ab, and provided $b \neq 0$, so is $\frac{a}{b}$; we do not attempt to give meaning to the expression $\frac{a}{0}$.

(b) (**Commutativity**) Addition and multiplication are commutative operations, which is to say that the order in which real numbers are added or multiplied does not matter. Explicitly, whenever a and b are real numbers, we have $a+b=b+a$ and $ab=ba$.

(c) (**Associativity**) Addition and multiplication are associative operations, which is to say that in a chain of multiple additions of pairs of real numbers can be evaluated in any order, and likewise for products. For example, whenever a, b and c are real numbers, we may evaluate $a+b+c$ as $(a+b)+c$ or as $a+(b+c)$, and we may evaluate abc as either $(ab)c$ or $a(bc)$.

(d) (**Inverse**) Addition and subtraction by a real number are inverse operations; likewise for multiplication and division. Thus for example, whenever a and b are real numbers, we have $(a+b)-b=a$ and $\frac{a}{b}\times b=a$ (the latter assuming that $b\neq 0$).

(e) (**Distributivity**) Multiplication distributes over addition and subtraction, which is to say that whenever a, b and c are real numbers, we have $a(b+c)=ab+ac$ and $a(b-c)=ab-ac$.

Please note that Assumptions 0.5 is *not* an exhaustive list of assumptions, but it illustrates some of the basic assumptions (axioms) we might want to spell out explicitly when we reason more formally later on. We will additionally assume basic algebraic facts about equations and inequalities of real numbers involving basic operations such as addition, subtraction, negation, multiplication, division, powers, exponents and surds (also known as radicals). For example:

- If a, b and c are real numbers with $a \leqslant b$ and $c<0$, then $ca \geqslant cb$;
- If a, b, c and d are real numbers with $b\neq 0$ and $d\neq 0$, then $\dfrac{a}{b}+\dfrac{c}{d}=\dfrac{ad+bc}{bd}$.

We will not attempt to make anything resembling a complete list of properties of this kind. It is safe to assume any fact of this nature from school-level elementary algebra.

We will use the number line to provide geometric interpretations four particular types of real number: the *natural numbers* (Definition 0.6), the *integers* (Definition 0.8) and the *rational numbers* (Definition 0.15).[a] Each of these types of number has a different character and different properties, and is used for different purposes.

[a] The naming of the number sets can be problematic for people upon encountering them for the first time. The real numbers are no more or less real than any other notion of number (such as *imaginary numbers*, which in turn are no more or less imaginary than real numbers); the rational numbers are *ration*al only inasmuch as they represent *ratio*s; and the natural numbers are so unnatural that no-one can agree on what exactly they are.

Natural numbers

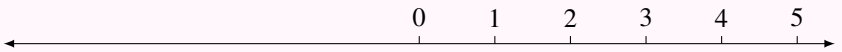
The *natural numbers* are the numbers used for counting—they are the answers to questions of the form 'how many'—for example, at the time of writing this paragraph, I have *three* guinea pigs and *zero* cats.

Counting is a skill humans have had for a very long time; we know this because there is evidence of people using tally marks tens of thousands of years ago. Tally marks provide one method of counting small numbers: starting with nothing, proceed through the objects you want to count one by one, and make a mark for every object. When you are finished, there will be as many marks as there are objects. We are taught from a young age to count with our fingers; this is another instance of making tally marks, where now instead of making a mark we raise a finger.

Making a tally mark represents an *increment* in quantity—that is, adding one. On our number line, we can represent an increment in quantity by moving to the right by the unit length. Then the distance from zero we have moved, which is equal to the number of times we moved right by the unit length, is therefore equal to the number of objects being counted.

Definition 0.6 · *Natural number*

The **natural numbers** are the nonnegative whole numbers; they are represented by the points on the number line that can be obtained by starting at 0 and moving right by a whole unit length any number of times:



The natural numbers have very important and interesting mathematical structure, and are central to the material in Chapter 8. A more precise characterisation of the natural numbers will be provided in Section 4.1, and a mathematical construction of the set of natural numbers can be found in Section B.1 (see Construction B.2.5). Central to these more precise characterisations will be the notions of 'zero' and of 'adding one'—just like making tally marks.

Aside

Some authors define the natural numbers to be the *positive* whole numbers, starting at 1 and excluding 0. The perspective taken in this book is that the natural numbers are *counting numbers*, and there will be plenty of times where we will count 0 of something—for example, the number of eighteen-letter words on this page is 0. That said, as with any mathematical definition, the choice about whether 0 is or is not a natural number a matter of taste or convenience, and is merely a convention—it is not something that can be proved or refuted.

For the time being, we will take for granted the arithmetic and algebraic facts about natural numbers that should already be familiar, some of which are outlined below:

Assumption 0.7 · *Closure properties of the natural numbers*
The natural numbers are *closed* under addition and multiplication, meaning that whenever a and b are natural numbers, so are $a+b$ and ab.

Note that adding or multiplying two natural numbers always results in another natural number; this is to say that the natural numbers are *closed* under the operations of addition and multiplication operations. More precisely, whenever a and b are natural numbers, so are $a+b$ and ab.

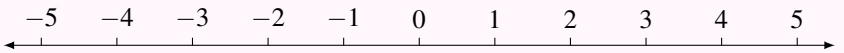
Importantly, the natural numbers are *not* closed under subtraction or division: for example, 1 and 2 are natural numbers, but $1-2$ and $\frac{1}{2}$ are not.

Integers

The *integers* can be used for measuring the difference between two natural numbers. For example, suppose I have five apples and five bananas. Another person, also holding apples and bananas, wishes to trade. After our exchange, I have seven apples and only one banana. Thus I have *two more* apples and *four fewer* bananas.

Since an increment in quantity can be represented by moving to the right on the number line by the unit length, a *decrement* in quantity can therefore be represented by moving to the *left* by the unit length. Doing so gives rise to the integers.

Definition 0.8
The **integers** are the whole numbers; they are represented by the points on the number line which can be obtained by starting at 0 and moving in either direction by the unit length any number of times:



Including the negative whole numbers ensures the closure of the integers under *subtraction*, in addition to the closure properties enjoyed by the natural numbers:

Assumption 0.9 · *Closure properties of the integers*
The integers are closed under addition, subtraction and multiplication, meaning that whenever a and b are integers, so are $a+b$, $a-b$ and ab.

However, the integers still fail to be closed under division, since for example $\frac{1}{2}$ is not an integer.

The integers have such a fascinating structure that a whole chapter of this book is devoted to them—see Chapter 7. This is to do with the fact that, although you can add, subtract and multiply two integers and obtain another integer, the same is not true of division. This 'bad behaviour' of division is what makes the integers interesting.

One of the properties of integers that is useful in proofs is that there is always a *next* one: given any integer n, the next integer is $n+1$, meaning that $n+1 > n$ and there are no integers k that satisfy the inequality $n < k < n+1$. Sometimes it is useful to be able to round a real number x up or down to the next integer; this is what the *floor* and *ceiling* operations do.

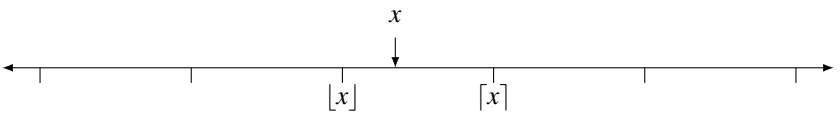
Definition 0.10 · *Floor and ceiling*

Let x be a real number.

- The **floor** of x is the integer n such that $n \leqslant x < n+1$; we write $\lfloor x \rfloor$ to denote the floor of x (LaTeX code: `\lfloor, \rfloor`).
- The **ceiling** of x is the integer n such that $n-1 < x \leqslant n$; we write $\lceil x \rceil$ to denote the ceiling of x (LaTeX code: `\lceil, \rceil`).

Intuitively, $\lfloor x \rfloor$ is the result of rounding x *down* to the next integer, and $\lceil x \rceil$ is the result of rounding x *up* to the next integer.

The floor and ceiling operations are illustrated on the following number line diagram, with the notches representing the integers.



Example 0.11

Some examples of floors and ceilings of real numbers include $\lfloor \pi \rfloor = 3$, $\lceil \pi \rceil = 4$, $\lfloor 1.3 \rfloor = 1$ and $\lceil 1.3 \rceil = 2$.

We now have enough to work with to write some basic proofs. At this stage, it is not necessary to worry too much about total mathematical rigour—we will save that for Chapter 1. For now, we will focus on reasoning clearly and working with definitions.

Proposition 0.12

Let n be an integer. Then $\lfloor n \rfloor = n = \lceil n \rceil$.

Proof

By Definition 0.10, $\lfloor n \rfloor$ is the integer m (we're calling it m here since the variable 'n' is already being used) such that $m \leqslant n < m+1$. Since $n \leqslant n < n+1$ and n is an integer, we must have $m = n$, and so $\lfloor n \rfloor = n$.

Likewise, since $n-1 < n \leqslant n$ and n is an integer, it follows from Definition 0.10 that $\lceil n \rceil = n$. □

Exercise 0.13

Suppose that x is a real number that is not an integer. Prove that $\lceil x \rceil = \lfloor x \rfloor + 1$.

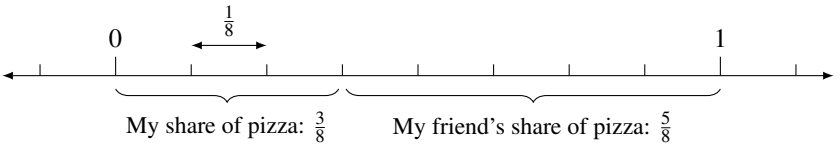
Exercise 0.14

Suppose that x is a real number. Prove that $\lfloor -x \rfloor = -\lceil x \rceil$ and $\lceil -x \rceil = -\lfloor x \rfloor$.

Rational numbers

Suppose I buy a pizza which is divided evenly into eight slices. A friend and I decide to share the pizza. I don't have much of an appetite, so I eat three slices and my friend eats five. Unfortunately, we cannot represent the proportion of the pizza each of us has eaten using natural numbers or integers: we each ate more than zero pizzas, but less than one pizza.

However, we're not far off: we can count the number of equal parts the pizza was split into (8), and of those parts, we can count how many we each had (3 and 5, respectively). On the number line, this could be represented by splitting the unit line segments into eight equal segments, whose length we call an *eighth* ($\frac{1}{8}$) of a unit. Counting eight segments is then equivalent to counting one unit, and each segment represents one slice of pizza.

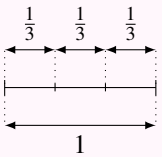


This kind of procedure gives rise to the *rational numbers*.

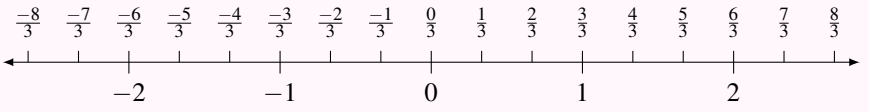
Definition 0.15

A **rational number** is a real number that can be expressed as a fraction whose numerator is an integer and whose denominator is a nonzero integer; that is, x is a rational number if and only if there are integers a and b, with $b \neq 0$, such that $x = \dfrac{a}{b}$.

The representation of rational numbers on the number line is a little more involved. Given a positive integer b, the real number $\frac{1}{b}$ represents the length obtained by breaking up the unit length into b equal parts:



The real numbers of the form $\frac{a}{b}$, where a is an integer, are those obtained by starting at 0 and moving to the left or right by $\frac{1}{b}$ units some number of times, with the direction and number of steps determined by the value of a:



When $b < 0$, we obtain the geometric interpretation of $\frac{a}{b}$ using the fact that $-b > 0$ and $\frac{a}{b} = \frac{-a}{-b}$.

In addition to the closure properties enjoyed by the integers, the rational numbers are also closed under division. Since we defined the rational numbers in terms of integers, instead of just in terms of points on a number line, we may actually *prove* these closure properties instead of simply assuming them.

Theorem 0.16 · *Closure properties of the rational numbers*
Suppose that x and y are rational numbers. Then $x+y$, $x-y$ and xy are also rational numbers, and that if $y \neq 0$ then $\dfrac{x}{y}$ is a rational number.

Proof that $x+y$ is rational
Since x and y are rational, it follows from Definition 0.15 that there are integers a, b, c, d with $b \neq 0$ and $d \neq 0$ such that $x = \dfrac{a}{b}$ and $y = \dfrac{c}{d}$.

Elementary algebra yields $x+y = \dfrac{a}{b} + \dfrac{c}{d} = \dfrac{ad+bc}{bd}$. Since $ad+bc$ and bd are integers, and since $bd \neq 0$ (because $b \neq 0$ and $d \neq 0$), it follows from Definition 0.15 that $x+y$ is rational. □

Exercise 0.17
Complete the proof of Theorem 0.16.

Rational numbers central to several areas of mathematics, including algebraic number theory and algebraic geometry.

Sets and elements

Fundamental to modern mathematics is the notion of a *set*. You will find sets in every chapter of the textbook, so we will take some time to think about them now. We will not treat sets formally at this stage—for now, the following definition will suffice.

Definition 0.18 · *to be revised in Section 2.1*
A **set** is a collection of objects. The objects in the set are called **elements** of the set. When X is a set and x is an object, we write $x \in X$ (LaTeX code: `x \in X`) to denote the assertion that x is an element of X.

Sets provide a convenient language for talking about objects of a certain kind without having to describe them every time, thereby making our mathematical arguments more concise and less verbose.

In fact, sets are so convenient that they are the basis of the (currently) most widely used logical foundation for mathematics—we would be getting too far ahead of ourselves to delve into this now, but details can be found in Section B.1. Sets are also interesting objects of study in their own right: we will study them in greater detail in Chapter 2.

Some ubiquitous examples of sets are given by sets of numbers of a given type, collectively called *number sets*.

Definition 0.19 · *Number sets*

The **number sets** are defined as follows:

- $\mathbb{N}$ is the set of all natural numbers;
- $\mathbb{Z}$ is the set of all integers;
- $\mathbb{Q}$ is the set of all rational numbers;
- $\mathbb{R}$ is the set of all real numbers.

Aside

Mathematics is both case-sensitive and font-sensitive, so for example it is possible in theory to use a, $\mathbf{a}$, A, $\mathbf{A}$, $\mathbb{A}$, and $\mathscr{A}$ to refer to six different things. It is important to keep this in mind when handwriting mathematics (see Chapter A) and typesetting mathematics (see Chapter D). The font used for the number sets is called *blackboard bold* (LaTeX code: `\mathbb{...}`), and is commonplace in modern mathematical texts; some texts use regular bold face (as in $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \dots$) instead.

Example 0.20

By Definition 0.18 and Definition 0.19, the expression $x \in \mathbb{Z}$ means that x is an integer, and the expression $x \notin \mathbb{Z}$ means that x is not an integer. Therefore, for example, the propositions $4 \in \mathbb{Z}$ and $\frac{5}{2} \notin \mathbb{Z}$ are both true.

Exercise 0.21

Determine the truth values of the following propositions.

(a) $2 \in \mathbb{N}$

(b) $-2 \in \mathbb{N}$

(c) $-2 \in \mathbb{Q}$

(d) $\dfrac{2}{5} \in \mathbb{Z}$

(e) $\dfrac{\sqrt{8}+\sqrt{2}}{\sqrt{8}-\sqrt{2}} \in \mathbb{Q}$

(f) $\pi^{\pi} \in \mathbb{Q}$

(g) $\sqrt{-7} \in \mathbb{R}$

(h) $\mathbb{N} \in \mathbb{Z}$

Specifying a set

We defined the sets in Definition 0.19 using words, which is fine if the set is easy to describe in that way, but in practice it can be very cumbersome to do so. Some other ways of specifying sets, which are usually more concise, are using *lists* and *set-builder notation*, described next.

- **Lists.** A set can be specified by simply listing its elements; the elements should be separated by commas and enclosed in {curly brackets} (LaTeX code: `\{...\}`). For example, the following is

a specification of a set A, whose elements are the natural numbers 0, 1, 2, 3, 4 and 5 (inclusive):

$$A = \{0,1,2,3,4,5\}$$

Sometimes a list might be too long to write out—maybe even infinite—or the length of the list might depend on a variable. In these cases it will be convenient to use an **implied list**, in which some elements of the list are written, and the rest are left implicit by writing an ellipsis '…' (LaTeX code: `\dots`). For example, the statement

$$B = \{0,1,4,9,16,25,\dots\}$$

means (most likely) that B is the set whose elements are all the squares of natural numbers. Note that there is some ambiguity here, since the reader is expected to be able to figure out what pattern is suggested by the '…'.

- **Set-builder notation.** List notation is only useful when discussing small, finite sets or sets whose elements can be listed with an obvious pattern. For some sets, it is not clear what order its elements can be listed in—in fact, as we will see in Section 6.2, there are some sets whose elements can't be listed at all (even in an infinite list)!

  To get around this, we can use *set-builder notation*, which is a means of specifying a set in terms of the properties its elements satisfy. The set of all objects x satisfying some property $p(x)$ is written in set-builder notation as:

  $$\{x \mid p(x)\} \quad \text{or} \quad \{x : p(x)\}$$

  An object a is an element of the set $\{x \mid p(x)\}$ if and only if the property $p(x)$ is true when a is substituted for x in the expression $p(x)$; that is, $a \in \{x \mid p(x)\}$ if and only if $p(a)$ is true.

  For example, if C is the set of all real numbers whose square is less than 9, we could write

  $$C = \{x \mid x \in \mathbb{R} \text{ and } x^2 < 9\}$$

  Then $2 \in C$ since the proposition '$2 \in \mathbb{R}$ and $2^2 < 9$' is true, but $4 \notin C$ since the proposition '$4 \in \mathbb{R}$ and $4^2 < 9$' is false.

  A couple of variations on set-builder notation are useful:

  (i) If X is a set that has already been defined, then $\{x \in X \mid p(x)\}$ denotes the set of all elements of X that satisfy the property $p(x)$. That is,

  $$\{x \in X \mid p(x)\} = \{x \mid x \in X \text{ and } p(x)\} \text{ is true}$$

  For example, the set C described above could be more succinctly expressed as

  $$C = \{x \in \mathbb{R} \mid x^2 < 9\}$$

  (ii) If $\text{expr}(x)$ is some expression involving the variable x, then $\{\text{expr}(x) \mid p(x)\}$ refers to the set of all objects that can be expressed in the form $\text{expr}(x)$ for some element x that satisfies the property $p(x)$. That is,

  $$\{\text{expr}(x) \mid p(x)\} = \{y \mid \text{there is an object } x \text{ such that } y = \text{expr}(x) \text{ and } p(x) \text{ is true}\}$$

  For example, the set $B = \{0,1,4,9,16,25,\dots\}$ described above could be more succinctly expressed as

  $$B = \{n^2 \mid n \in \mathbb{N}\}$$

  since its elements are those objects of the form n^2 where n is a natural number.

Example 0.22

The sets $\mathbb{N}$ and $\mathbb{Z}$ can be described in (implied) list notation as follows:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\} \quad \text{and} \quad \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, 3, \dots\}$$

It is not necessarily to list number sets in increasing order, and in fact we can avoid having two instances of '…' in the list notation for $\mathbb{Z}$ by alternating between non-negative and negative integers in the list as follows:

$$\mathbb{Z} = \{0, -1, 1, -2, 2, -3, 3, \dots\}$$

This seemingly trivial rearrangement of the list of integers has important consequences that will be explored in Chapter 6.

It is possible to express $\mathbb{Q}$ in list notation, but it is not at all obvious how to do so; in fact, from an expression of $\mathbb{Q}$ in list notation, it would not be clear that the set being listed was actually $\mathbb{Q}$ and not some other set. For this reason, list notation is not really suitable as a means of expressing $\mathbb{Q}$—the daring reader is invited to try! However, $\mathbb{Q}$ can readily be described in set-builder notation as

$$\mathbb{Q} = \left\{ x \in \mathbb{R} \;\middle|\; x = \frac{a}{b} \text{ for some } a \in \mathbb{Z} \text{ and } b \in \mathbb{Z} \text{ with } b \neq 0 \right\}$$

We could use variant (ii) described above to express $\mathbb{Q}$ more succinctly as

$$\mathbb{Q} = \left\{ \frac{a}{b} \;\middle|\; a \in \mathbb{Z} \text{ and } b \in \mathbb{Z} \text{ with } b \neq 0 \right\}$$

Perhaps surprisingly, it is *impossible* to express $\mathbb{R}$ in list notation at all—we will not be ready to prove this impossibility for quite some time yet, but it is an immediate consequence of of Theorem 9.3.21.

Aside

When we want to refer to more than one object being an element of a set, we sometimes abuse notation by simply separating the variables by commas. For example, in Example 0.22 we could have written

$$\mathbb{Q} = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z} \text{ with } b \neq 0 \right\}$$

Example 0.23

The set of all even integers can be written in set-builder notation as

$$\{n \in \mathbb{Z} \mid n \text{ is even}\}$$

For comparison, the set of all even natural numbers can be written as

$$\{n \in \mathbb{N} \mid n \text{ is even}\} = \{0, 2, 4, 6, \dots\}$$

Note that -6 is an element of the former set but not of the latter set, since -6 is an integer but is not a natural number.

Note moreover that the expression

$$\{n \in \mathbb{Q} \mid n \text{ is even}\}$$

is meaningless—at least, unless we define what 'even' means in the context of rational numbers that are not integers.

Exercise 0.24

A **dyadic rational** is a rational number that can be expressed as an integer divided by a power of 2. Express the set of all dyadic rationals using set-builder notation in at least two different ways.

Set-builder notation is useful for defining sets based on the properties they satisfy, as in Definitions 0.25 and 0.28 below.

Definition 0.25

Let $n \in \mathbb{N}$. The **canonical n-element set** is the set $[n]$ defined by

$$[n] = \{k \in \mathbb{N} \mid k < n\}$$

When speaking out loud it is common to read $[n]$ as 'bracket-n'.

Example 0.26

In list notation, we have $[1] = \{0\}$, $[2] = \{0,1\}$, $[3] = \{0,1,2\}$, $[4] = \{0,1,2,3\}$, and so on.

Note that $[0]$ has no elements (it is *empty*, a term we will define in Definition 2.1.19), since there are no natural numbers k satisfying the inequality $k < 0$.

Exercise 0.27

Let $A = \{x \in [10] \mid x \text{ is a power of } 2\}$, let $B = \{2^x \mid x \in [10]\}$ and let $C = \{-99, -97, \ldots, 95, 97, 99\}$.

(a) Find a list notation expression for A.

(b) Find a list notation expression for B.

(c) Find a set-builder notation expression for C that involves a set of the form $[n]$ for some $n \in \mathbb{N}$.

While not particularly interesting yet, sets of the form $[n]$ will be fundamental throughout Chapter 8, as they are used to define the notion of a *finite set*, as well as the *size* of a finite set.

Aside

The key property of $[n]$ (for $n \in \mathbb{N}$) that we will need later in the book is that it has n elements, but exactly *which* n elements those are doesn't matter too much—we could just as well have defined $[n] = \{k \in \mathbb{N} \mid 1 \leqslant k \leqslant n\}$, so that for instance we would have $[3] = \{1,2,3\}$ instead of $\{0,1,2\}$, and everything we do could be adapted to this alternative definition. In fact, it is more common to see the definition $[n] = \{k \in \mathbb{N} \mid 1 \leqslant k \leqslant n\}$ in other texts. However, the definition of $[n]$ as in Definition 0.25 simplifies some proofs and generalises more immediately to the infinite setting, so we will stick to that.

Intervals are particular subsets of $\mathbb{R}$ that are ubiquitous in mathematics, and especially in analysis and topology.

Definition 0.28 · *Intervals of the real line*

Let $a, b \in \mathbb{R}$ with $a < b$. The **open interval** (a,b), the **closed interval** $[a,b]$, and the **half-open intervals** $[a,b)$ and $(a,b]$ from a to b are defined by

$$(a,b) = \{x \in \mathbb{R} \mid a < x < b\} \qquad (a,b] = \{x \in \mathbb{R} \mid a < x \leqslant b\}$$
$$[a,b) = \{x \in \mathbb{R} \mid a \leqslant x < b\} \qquad [a,b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$$

In each of the above definitions, a is the **lower bound** and b is the **upper bound** of the interval, and these four kinds of intervals are called **bounded intervals**. We further define the **unbounded intervals** $(-\infty,b)$, $(-\infty,b]$, $[a,\infty)$ and (a,∞) (LaTeX code: `\infty`) by

$$(-\infty,b) = \{x \in \mathbb{R} \mid x < b\} \qquad (a,\infty) = \{x \in \mathbb{R} \mid x > a\}$$
$$(-\infty,b] = \{x \in \mathbb{R} \mid x \leqslant b\} \qquad [a,\infty) = \{x \in \mathbb{R} \mid x \geqslant a\}$$

and finally we define $(-\infty,\infty) = \mathbb{R}$.

Note that intervals are the set of all *real* numbers satisfying the defining inequality, so for example it would not be appropriate to express the set $\{x \in \mathbb{Z} \mid 2 \leqslant x \leqslant 5\}$ as $[2,5]$: the former only has integers as elements, while the latter includes all real numbers from 2 to 5.

Example 0.29

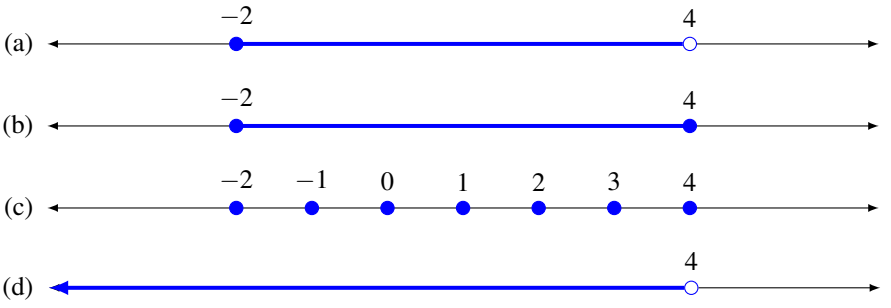
The following illustration depicts the open interval $(-2,5)$.

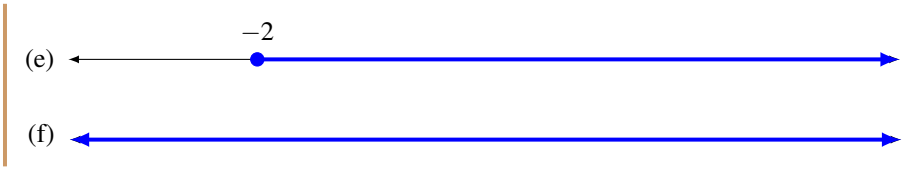
The hollow circles ∘ indicate that the endpoints are not included in the interval.

Be warned that the use of the symbol ∞ is misleading, since it suggests that the symbol ∞ on its own has a specific meaning (or, worse, that it refers to a real number). It doesn't—it is just a symbol that suggests unboundedness of the interval in question. A less misleading way of writing $[a,\infty)$, for instance, might be $[a,\rightarrow)$ or $\mathbb{R}^{\geqslant a}$; however, $[a,\infty)$ is standard, so it is what we will write.

Exercise 0.30

Express each of the sets illustrated below using the most appropriate notation. A filled circle • indicates that a point is included in the set, whereas a hollow circle ∘ indicates that an end-point is not included in the interval.





Getting started with proofs

The purpose of the remainder of this section is to introduce some topics involving the number sets defined above. We will follow the typical style of a mathematics text, by making definitions and proving facts about the definitions we made. Despite the fact that we have not introduced a formal system of logical reasoning, we will nonetheless reason precisely and attempt to keep our proofs as close as possible to the definitions we provide, without appealing to intuition.

Number bases

Writing numbers down is something that may seem easy to you now, but it likely took you several years as a child to truly understand what was going on. Historically, there have been many different systems for representing numbers symbolically, called *numeral systems*. First came the most primitive of all, tally marks, appearing in the Stone Age and still being used for some purposes today. Thousands of years and hundreds of numeral systems later, there is one dominant numeral system, understood throughout the world: the **Hindu–Arabic numeral system**. This numeral system consists of ten symbols, called *digits*. It is a *positional* numeral system, meaning that the position of a symbol in a string determines its numerical value.

In English, the *Arabic numerals* are used as the ten digits:

0 1 2 3 4 5 6 7 8 9

The right-most digit in a string is in the units place, and the value of each digit increases by a factor of ten moving to the left. For example, when we write ‘2812’, the left-most ‘2’ represents the number two thousand, whereas the last ‘2’ represents the number two.

The fact that there are ten digits, and that the numeral system is based on powers of ten, is a biological accident corresponding with the fact that most humans have ten fingers. For many purposes, this is inconvenient. For example, ten does not have many positive divisors (only four: 1, 2, 5 and 10)—this has implications for the ease of performing arithmetic; a system based on the number twelve, which has six positive divisors (1, 2, 3, 4, 6 and 12), might be more convenient. Another example is in computing and digital electronics, where it is more convenient to work in a *binary* system, with just two digits—0 and 1—which represent ‘off’ and ‘on’ (or ‘low voltage’ and ‘high voltage’), respectively; arithmetic can then be performed directly using sequences of *logic gates* in an electrical circuit.

It is therefore worthwhile to have some understanding of positional numeral systems based on numbers other than ten. The mathematical abstraction we make leads to the definition of *base-b expansion*.

Definition 0.31

Let b be a natural number greater than 1. The **base-b expansion** of a natural number n is the[b] string $d_r d_{r-1} \dots d_0$ such that:

(i) $n = d_r \cdot b^r + d_{r-1} \cdot b^{r-1} + \cdots + d_0 \cdot b^0$;

(ii) $0 \leqslant d_i < b$ for each i; and

(iii) If $n > 0$ then $d_r \neq 0$—the base-b expansion of zero is 0 in all bases b.

Certain number bases have names; for instance, the base-2, 3, 8, 10, 12 and 16 expansions are respectively called *binary*, *ternary*, *octal*, *decimal*, *duodecimal* (or *dozenal*) and *hexadecimal*.

Before we look at an example of Definition 0.31 in action, let's examine the definition, which is a little terse on first sight.

- Condition (i) tells us that the digits in the string tell us how many of each power of b are added up to obtain n. For example, when $b = 10$, the digits from right to left tell us the units, tens, hundreds, thousands, and so on.
- Condition (ii) tells us that the digits in a base-b expansion must be less than b—for example, the base-4 digits are 0, 1, 2 and 3. If we allowed more digits then silly things would happen—for example, if 'X' were a new base-10 digit representing the number ten, then 'X2' and '102' would be different strings both representing the number one hundred and two.
- Condition (iii) ensures that the string representing a positive number doesn't have any leading '0's—otherwise, for example, '01423' and '1423' would be different strings representing the same natural number.

Example 0.32

Consider the number 1023. Its decimal (base-10) expansion is 1023, since

$$1023 = 1 \cdot 10^3 + 0 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0$$

Its binary (base-2) expansion is 1111111111, since

$$1023 = 1 \cdot 2^9 + 1 \cdot 2^8 + 1 \cdot 2^7 + 1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0$$

We can express numbers in base-36 by using the ten usual digits 0 through 9 and the twenty-six letters A through Z; for instance, A represents 10, M represents 22 and Z represents 35. The base-36 expansion of 1023 is SF, since

$$1023 = 28 \cdot 36^1 + 15 \cdot 36^0 = \mathrm{S} \cdot 36^1 + \mathrm{F} \cdot 36^0$$

Exercise 0.33

Find the binary, ternary, octal, decimal, hexadecimal and base-36 expansions of the number 21127, using the letters A–F as additional digits for the hexadecimal expansion (representing the numbers 10–15, respectively), and the letters A–Z as additional digits for the base-36 expansion.

We sometimes wish to specify a natural number in terms of its base-b expansion; we have some notation for this.

Notation 0.34

Let $b > 1$. If the numbers $d_0, d_1, \ldots, d_r$ are base-b digits (in the sense of Definition 0.31), then we write

$$d_r d_{r-1} \ldots d_{0(b)} = d_r \cdot b^r + d_{r-1} \cdot b^{r-1} + \cdots + d_0 \cdot b^0$$

for the natural number whose base-b expansion is $d_r d_{r-1} \ldots d_0$. If there is no subscript (b) and a base is not specified explicitly, the expansion will be assumed to be in base-10.

Example 0.35

Using our new notation, we have

$$1023 = 1111111111_{(2)} = 1101220_{(3)} = 1777_{(8)} = 1023_{(10)} = 3\mathrm{FF}_{(16)} = \mathrm{SF}_{(36)}$$

Division of integers

Dividing one integer by another integer is not necessarily another integer. However, the result is *sometimes* another integer; for example, I can divide six apples between three people, and each person will receive an integral number of apples; that is not the case if I try to divide seven apples between three people. This makes division interesting: how can we measure the failure of one integer's divisibility by another? How can we deduce when one integer is divisible by another? What is the structure of the set of integers when viewed through the lens of division? This motivates Definition 0.36.

Definition 0.36 · *to be repeated in Definition 7.1.5*

Let $a, b \in \mathbb{Z}$. We say b **divides** a if $a = qb$ for some integer q. There are many other ways of saying that b divides a, such as: a is *divisible by* b, b is a *divisor* of a, b is a *factor* of a, or a is a *multiple* of b.

Note that, perhaps counterintuitively, the definition of divisibility does not involve the arithmetic operation of division: it is defined in terms of multiplication.

Example 0.37

The integer 12 is divisible by 1, 2, 3, 4, 6 and 12, since

$$12 = 12 \cdot 1 = 6 \cdot 2 = 4 \cdot 3 = 3 \cdot 4 = 2 \cdot 6 = 1 \cdot 12$$

It is also divisible by the negatives of all of those numbers; for example, 12 is divisible by -3 since $12 = (-4) \cdot (-3)$.

Exercise 0.38

Prove that 1 divides every integer, and that every integer divides 0.

A consequence of Exercise 0.38 is that 0 is divisible by 0. This is surprising: we've been told our whole lives that we can't divide by zero, but now we discover that we can divide zero by zero... how can that be? This highlights why it was so important for the definition of divisibility (Definition 0.36) to be given in terms of multiplication, without using the division operation: saying that 0 divides 0 simply means that 0 can be multiplied by an integer to obtain 0 (which is true)—but this does not imply that the expression '$\frac{0}{0}$' can (or should) be meaningfully defined.

Using Definition 0.36, we can prove some general basic facts about divisibility.

Proposition 0.39
Let $a, b, c \in \mathbb{Z}$. If c divides b and b divides a, then c divides a.

Proof
Suppose that c divides b and b divides a. By Definition 0.36, it follows that

$$b = qc \quad \text{and} \quad a = rb$$

for some integers q and r. Using the first equation, we may substitute qc for b in the second equation, to obtain

$$a = r(qc)$$

But $r(qc) = (rq)c$, and rq is an integer, so it follows from Definition 0.36 that c divides a. □

Exercise 0.40
Let $a, b, d \in \mathbb{Z}$. Suppose that d divides a and d divides b. Given integers u and v, prove that d divides $au + bv$.

Some familiar concepts, such as evenness and oddness, can be characterised in terms of divisibility.

Definition 0.41
An integer n is **even** if it is divisible by 2; otherwise, n is **odd**.

It is not just interesting to know when one integer *does* divide another; however, proving that one integer *doesn't* divide another is much harder. Indeed, to prove that an integer b does not divide an integer a, we must prove that $a \neq qb$ for *any* integer q at all. We will look at methods for doing this in Chapter 1; these methods use the following extremely important result, which will underlie all of Chapter 7.

Theorem 0.42 · *Division theorem, to be repeated in Theorem 7.1.1*
Let $a, b \in \mathbb{Z}$ with $b \neq 0$. There is exactly one way to write

$$a = qb + r$$

such that q and r are integers, and $0 \leqslant r < |b|$.

The number q in Theorem 0.42 is called the **quotient** of a when divided by b, and the number r is called the **remainder**.

Example 0.43
The number 12 leaves a remainder of 2 when divided by 5, since $12 = 2 \cdot 5 + 2$.

Here's a slightly more involved example.

Proposition 0.44
Suppose an integer a leaves a remainder of r when divided by an integer b, and that $r > 0$. Then $-a$ leaves a remainder of $b - r$ when divided by b.

Proof
Suppose a leaves a remainder of r when divided by b. Then

$$a = qb + r$$

for some integer q. A bit of algebra yields

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

Since $0 < r < b$, we have $0 < b - r < b$. Hence $-(q+1)$ is the quotient of $-a$ when divided by b, and $b - r$ is the remainder. □

Exercise 0.45
Prove that if an integer a leaves a remainder of r when divided by an integer b, then a leaves a remainder of r when divided by $-b$.

We will finish this part on division of integers by connecting it with the material on number bases—we can use the division theorem (Theorem 0.42) to find the base-b expansion of a given natural number. It is based on the following observation: the natural number n whose base-b expansion is $d_r d_{r-1} \cdots d_1 d_0$ is equal to

$$d_0 + b(d_1 + b(d_2 + \cdots + b(d_{r-1} + bd_r) \cdots))$$

Moreover, $0 \leqslant d_i < b$ for all i. In particular n leaves a remainder of d_0 when divided by b. Hence

$$\frac{n - d_0}{b} = d_1 + d_2 b + \cdots + d_r b^{r-1}$$

The base-b expansion of $\frac{n-d_0}{b}$ is therefore

$$d_r d_{r-1} \cdots d_1$$

In other words, the remainder of n when divided by b is the last base-b digit of n, and then subtracting this number from n and dividing the result by b truncates the final digit. Repeating this process gives us d_1, and then d_2, and so on, until we end up with 0.

This suggests the following algorithm for computing the base-b expansion of a number n:

- **Step 1.** Let d_0 be the remainder when n is divided by b, and let $n_0 = \frac{n-d_0}{b}$ be the quotient. Fix $i = 0$.

- **Step 2.** Suppose n_i and d_i have been defined. If $n_i = 0$, then proceed to Step 3. Otherwise, define d_{i+1} to be the remainder when n_i is divided by b, and define $n_{i+1} = \frac{n_i - d_{i+1}}{b}$. Increment i, and repeat Step 2.
- **Step 3.** The base-b expansion of n is
$$d_i d_{i-1} \cdots d_0$$

Example 0.46

We compute the base-17 expansion of 15213, using the letters A–G to represent the numbers 10 through 16.

- $15213 = 894 \cdot 17 + 15$, so $d_0 = 15 = \mathrm{F}$ and $n_0 = 894$.
- $894 = 52 \cdot 17 + 10$, so $d_1 = 10 = \mathrm{A}$ and $n_1 = 52$.
- $52 = 3 \cdot 17 + 1$, so $d_2 = 1$ and $n_2 = 3$.
- $3 = 0 \cdot 17 + 3$, so $d_3 = 3$ and $n_3 = 0$.
- The base-17 expansion of 15213 is therefore 31AF.

A quick verification gives

$$31\mathrm{AF}_{(17)} = 3 \cdot 17^3 + 1 \cdot 17^2 + 10 \cdot 17 + 15 = 15213$$

as desired.

Exercise 0.47

Find the base-17 expansion of 408 735 787 and the base-36 expansion of 1 442 151 747.

Irrational numbers

Before we can talk about irrational numbers, we should say what they are.

Definition 0.48

An **irrational number** is a real number that is not rational.

Unlike $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, there is no standard single letter expressing the irrational numbers. However, by the end of Section 2.2, we will be able to write the set of irrational numbers as '$\mathbb{R} \setminus \mathbb{Q}$'.

Proving that a real number is *irrational* is not particularly easy, in general. We will get our foot in the door by allowing ourselves to assume the following result, which is restated and proved in Proposition 4.3.12.

Assumption 0.49

The real number $\sqrt{2}$ is irrational.

We can use the fact that $\sqrt{2}$ is irrational to prove some facts about the relationship between rational numbers and irrational numbers.

Proposition 0.50
There exist irrational numbers a and b such that ab is rational.

Proof
Let $a = b = \sqrt{2}$. Then a and b are irrational, and $ab = 2 = \frac{2}{1}$, which is rational. □

Exercise 0.51
Let r be a rational number and let a be an irrational number. Prove that it is possible that ra be rational, and it is possible that ra be irrational.

Polynomials and complex numbers

The closure properties of the number sets make them suitable for defining *polynomials*, which are expressions involving powers of variables representing unknown quantities. Polynomials are fundamental objects of study in algebra and have far-reaching applications, both within mathematics (e.g. Taylor polynomials in calculus) and in mathematical models of physical phenomena.

Definition 0.52
Let $\mathbb{S}$ be a number set ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$). A (**univariate**) **polynomial over** $\mathbb{S}$ in the **indeterminate** x is an expression of the form

$$a_0 + a_1 x + \cdots + a_n x^n$$

where $n \in \mathbb{N}$ and each $a_k \in \mathbb{S}$. The numbers a_k are called the **coefficients** of the polynomial. If not all coefficients are zero, the largest value of k for which $a_k \neq 0$ is called the **degree** of the polynomial. By convention, the degree of the polynomial 0 is $-\infty$.

Polynomials of degree $\leqslant 0$, 1, 2, 3, 4 and 5 are respectively called *constant*, *linear*, *quadratic*, *cubic*, *quartic* and *quintic* polynomials.

Notation 0.53
Polynomials are typically denoted using expressions like $p(x)$ or $q(u)$, where the variable in parentheses denotes the indeterminate of the polynomial. The same notation is used for predicates (Definition 1.2.2), but is disambiguated by context.

If $p(x)$ is a polynomial over a number set $\mathbb{S}$, and $\alpha \in \mathbb{S}$, then $p(\alpha)$ denotes the element of $\mathbb{S}$ obtained by substituting α for x in $p(x)$ and evaluating the arithmetic operations.

Example 0.54
Some examples of polynomials are the following:

- Define $p(x) = 2$. Then $p(x)$ is a polynomial of degree 0 (a *constant* polynomial) over $\mathbb{N}$, over $\mathbb{Z}$, over $\mathbb{Q}$ and over $\mathbb{R}$.

- Define $q(u) = 2u - 1$. Then $q(u)$ is a polynomial of degree 1 (a *linear* polynomial) over $\mathbb{Z}$, over $\mathbb{Q}$ and over $\mathbb{R}$, but not over $\mathbb{N}$ since $-1 \notin \mathbb{N}$.
- Define $r(t) = \frac{2}{3}t^2 - t$. Then $r(t)$ is a polynomial of degree 2 (a *quadratic* polynomial) over $\mathbb{R}$ and over $\mathbb{Q}$, but not over $\mathbb{Z}$ or $\mathbb{N}$ since $\frac{2}{3}$ is not an integer.

With $p(x)$, $q(u)$ and $r(t)$ defined as above, we have

$$p(3) = 2, \quad q(3) = 2 \cdot 3 - 1 = 5 \quad \text{and} \quad r(3) = \frac{2}{3} \cdot 3^2 - 3 = 3$$

Exercise 0.55
Find a quartic polynomial $s(w)$ over $\mathbb{R}$ that is not a polynomial over $\mathbb{Q}$ but for which $s(3) \in \mathbb{Q}$.

Note that, if $\mathbb{S}$ is any of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$, and $p(x)$ is a polynomial over $\mathbb{S}$, then $p(\alpha) \in \mathbb{S}$ for all $\alpha \in \mathbb{S}$. This is because all of the number sets $\mathbb{S}$ are closed under addition and multiplication, and polynomials are defined using only the operations of addition and multiplication (and exponentiation by a natural number, which is equivalent to repeated multiplication).

Definition 0.56
Let $p(x)$ be a polynomial. A **root** of $p(x)$ is a number α such that $p(\alpha) = 0$.

Verifying that numbers are or are not roots of a given polynomial is generally very easy: simply substitute and evaluate!

Example 0.57
Let $p(x) = x^2 - 2$. Then $\sqrt{2}$ is a root of $p(x)$ since

$$p(\sqrt{2}) = \sqrt{2}^2 - 2 = 2 - 2 = 0$$

but $\sqrt{3}$ is not a root of $p(x)$ since

$$p(\sqrt{3}) = \sqrt{3}^2 - 2 = 3 - 2 = 1 \neq 0$$

Finding the roots of a polynomial is a much harder problem in general. We will start with the simplest non-constant case:

Exercise 0.58
Let $p(x) = ax + b$ be a linear polynomial over $\mathbb{Z}$.

(a) Find a root of $p(x)$, expressing it in terms of a and b.

(b) Prove that the root you found in (a) is, in fact, the *only* root of $p(x)$.

(c) Under what conditions on a and b is the root of $p(x)$ an integer?

Finding the roots of quadratic polynomials is harder than for linear polynomials, but not much more so: it can be done by rearranging equations using techniques from elementary algebra.

Definition 0.59
Let $p(x) = ax^2 + bx + c$ be a quadratic polynomial over $\mathbb{R}$. The **discriminant** of $p(x)$ is the real number Δ defined by $\Delta = b^2 - 4ac$.

Theorem 0.60 · *Quadratic formula*
Let $p(x) = ax^2 + bx + c$ be a quadratic polynomial over $\mathbb{R}$ (so that $a, b, c \in \mathbb{R}$ and $a \neq 0$), let Δ be its discriminant, and suppose that $\Delta \geqslant 0$. Then the roots of $p(x)$ are precisely α and β, where

$$\alpha = \frac{-b + \sqrt{\Delta}}{2a} \quad \text{and} \quad \beta = \frac{-b - \sqrt{\Delta}}{2a}$$

Proof
Completing the square yields

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

Now:

- α is a root of $p(x)$. To see this, note that

  $$p(\alpha) = a\left(\frac{\sqrt{\Delta}}{2a}\right)^2 - \frac{\Delta}{4a} = \frac{\Delta}{4a} - \frac{\Delta}{4a} = 0$$

  Likewise, β is a root of $p(x)$ since

  $$p(\beta) = a\left(-\frac{\sqrt{\Delta}}{2a}\right)^2 - \frac{\Delta}{4a} = \frac{\Delta}{4a} - \frac{\Delta}{4a} = 0$$

- To see that there are no other roots, suppose that γ is another root of $p(x)$; we will show that γ must be equal to either α or β.

  Since $p(\gamma) = 0$, we have $a\left(\gamma + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a} = 0$. Adding $\frac{\Delta}{4a}$ to both sides and dividing through by a gives $\left(\gamma + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a^2}$, and then taking square roots gives

  $$\gamma + \frac{b}{2a} = \frac{\sqrt{\Delta}}{2a} \quad \text{or} \quad \gamma + \frac{b}{2a} = -\frac{\sqrt{\Delta}}{2a}$$

  Rearranging the first equation gives $\gamma = \alpha$, and rearranging the second equation gives $\gamma = \beta$, as required. □

The symbol '$\pm$' is commonly used in the context of the quadratic formula, since it allows us to say concisely that the roots of $p(x) = ax^2 + bx + c$ are $\frac{-b \pm \sqrt{\Delta}}{2a}$. We have taken care to avoid this: the

symbol '$\pm$' is dangerous because it may suppress the word 'and' or the word 'or', depending on context—this kind of ambiguity is not something that we will want to deal with when discussing the logical structure of a proposition in Chapter 1!

Exercise 0.61
Let $p(x)$ be a polynomial over $\mathbb{R}$ with discriminant Δ. Prove that $p(x)$ has exactly two real roots if $\Delta > 0$, exactly one real root if $\Delta = 0$, and no real roots if $\Delta < 0$.

Although quadratic polynomials with negative discriminant have no *real* roots, there is another number set in which all quadratic polynomials (in fact all *non-constant* polynomials!) have roots, namely the set of all *complex numbers*, which are the result of extending $\mathbb{R}$ by adjoining a square root of -1.

Definition 0.62
The **imaginary unit** i represents a fixed solution to the equation $i^2 = -1$. Given a negative real number a, we define $\sqrt{a} = \sqrt{|a|}i$, so in particular $\sqrt{-1} = i$.

A **complex number** is an expression of the form $a+bi$, where $a, b \in \mathbb{R}$. The set of all complex numbers is denoted by $\mathbb{C}$.

We can perform arithmetic operations with complex numbers by treating i as a variable, simplifying powers of i using the fact that $i^2 = -1$.

Example 0.63
Let $z = 1+3i$ and $w = 2-i$. Then:

- $z+w = (1+3i)+(2-i) = (1+2)+(3i-i) = 3+2i$;
- $z-w = (1+3i)-(2-i) = (1-2)+(3i-(-i)) = -1+4i$; and
- $zw = (1+3i)(2-i) = 2+5i-3i^2 = 2+5i-(-3) = 5+5i$.

Division by a complex number is slightly more complicated: suppose $a, b \in \mathbb{R}$ not both zero. Then

$$(a+bi)(a-bi) = a^2 - (bi)^2 = a^2 + b^2 \neq 0$$

and so

$$\frac{1}{a+bi} = \frac{1}{a+bi}\frac{a-bi}{a-bi} = \frac{a-bi}{a^2+b^2}$$

Thus we have

$$\frac{z}{w} = (1+3i)\cdot\frac{1}{2-i} = \frac{(1+3i)(2+i)}{5} = \frac{-1+7i}{5}$$

Exercise 0.64
Express i^n in the form $a+bi$ for each $n \in \mathbb{Z}$.

Exercise 0.65

Verify that the conclusion of the quadratic formula (Theorem 0.60) is true for *all* quadratic polynomials over $\mathbb{R}$, regardless of their discriminant.

Exercise 0.66

Let $\alpha = a + bi$ be a complex number, where $a, b \in \mathbb{R}$. Prove that α is the root of a quadratic polynomial over $\mathbb{R}$, and find the other root of this polynomial.

Exercise 0.67

Prove a general version of the quadratic formula for quadratic polynomials over $\mathbb{C}$.

⋆ Geometric interpretation of order and arithmetic

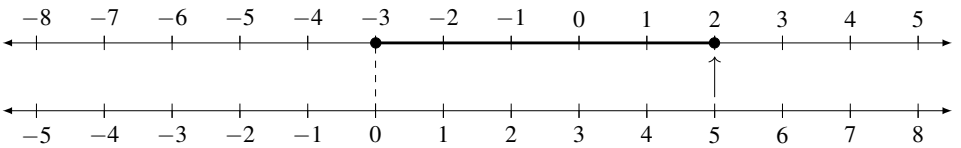
Earlier in this section, we motivated the definitions of the number sets geometrically using an infinite number line, but most of our proofs have been algebraic in nature. It is interesting to see how the symbolic operations and relations that appear in our algebraic proofs translate to this geometric interpretation.

- **Ordering.** A real number a is less than a real number b, written $a < b$, if a lies to the left of b on the number line. The usual conventions for the symbols $\leqslant$ (LaTeX code: `\le`), $>$ and $\geqslant$ (LaTeX code: `\ge`) apply, for instance '$a \leqslant b$' means that either $a < b$ or $a = b$.

  For example, the following diagram demonstrates that $-3 < 2$ and that $5 \geqslant 2$:

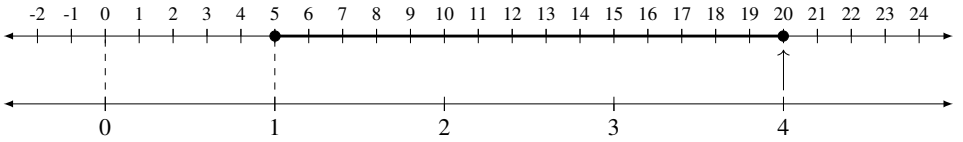
- **Addition.** Suppose we want to add a real number a to a real number b. To do this, we *translate* a by b units to the right—if $b < 0$ then this amounts to translating a by an equivalent number of units to the left. Concretely, take two copies of the number line, one above the other, with the same choice of unit length; move the 0 of the lower number line beneath the point a of the upper number line. Then $a + b$ is the point on the upper number line lying above the point b of the lower number line.

  Here is an illustration of the fact that $(-3) + 5 = 2$:

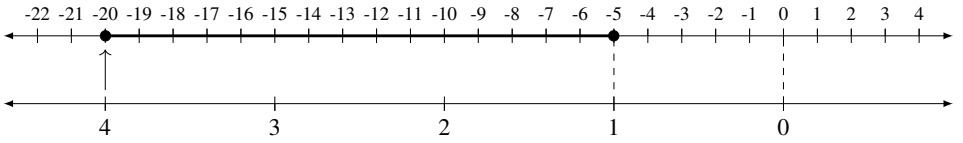


- **Multiplication.** This one is fun. Suppose we want to multiply a real number a by a real number b. To do this, we *scale* the number line, and perhaps *reflect* it. Concretely, take two copies of the number line, one above the other; align the 0 points on both number lines, and stretch the lower number line evenly until the point 1 on the lower number line is below the point a on the upper number line—note that if $a < 0$ then the number line must be reflected in order for this to happen. Then $a \cdot b$ is the point on the upper number line lying above b on the lower number line.

  Here is an illustration of the fact that $5 \cdot 4 = 20$.



and here is an illustration of the fact that $(-5)\cdot 4 = -20$:



Exercise 0.68

Interpret the operations of subtraction and division as geometric transformations of the real number line.

Section 0.E

Chapter 0 exercises

0.E.1. The video-sharing website *YouTube* assigns to each video a unique identifier, which is a string of 11 characters from the set

$$\{\texttt{A},\texttt{B},\ldots,\texttt{Z},\texttt{a},\texttt{b},\ldots,\texttt{z},0,1,2,3,4,5,6,7,8,9,\texttt{-},\texttt{\_}\}$$

This string is actually a natural number expressed in base-64, where the characters in the above set represent the numbers 0 through 63, in the same order—thus `C` represents 2, `c` represents 28, `3` represents 55, and `_` represents 63. According to this schema, find the natural number whose base-64 expansion is `dQw4w9WgXcQ`, and find the base-64 expansion of the natural number 7 159 047 702 620 056 984.

0.E.2. Let $a,b,c,d \in \mathbb{Z}$. Under what conditions is $(a+b\sqrt{2})(c+d\sqrt{2})$ an integer?

0.E.3. Suppose an integer m leaves a remainder of i when divided by 3, and an integer n leaves a remainder of j when divided by 3, where $0 \leqslant i, j < 3$. Prove that $m+n$ and $i+j$ leave the same remainder when divided by 3.

0.E.4. What are the possible remainders of n^2 when divided by 3, where $n \in \mathbb{Z}$?

Definition 0.E.5
A set X is **closed** under an operation $\odot$ if, whenever a and b are elements of X, $a \odot b$ is also an element of X.

In Questions 0.E.6 to 0.E.11, determine which of the number sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are closed under the operation $\odot$ defined in the question.

0.E.6. $a \odot b = a+b$

0.E.7. $a \odot b = a-b$

0.E.8. $a \odot b = (a-b)(a+b)$

0.E.9. $a \odot b = (a-1)(b-1)+2(a+b)$

0.E.10. $a \odot b = \dfrac{a}{b^2+1}$

0.E.11. $a \odot b = \dfrac{a}{\sqrt{b^2+1}}$

Definition 0.E.12
A complex number α is **algebraic** if $p(\alpha) = 0$ for some nonzero polynomial $p(x)$ over $\mathbb{Q}$.

0.E.13. Let x be a rational number. Prove that x is an algebraic number.

0.E.14. Prove that $\sqrt{2}$ is an algebraic number.

0.E.15. Prove that $\sqrt{2}+\sqrt{3}$ is an algebraic number.

0.E.16. Prove that $x+yi$ is an algebraic number, where x and y are any two rational numbers.

True–False questions

In Questions 0.E.17 to 0.E.23, determine (with proof) whether the statement is true or false.

0.E.17. Every integer is a natural number.

0.E.18. Every integer is a rational number.

0.E.19. Every integer divides zero.

0.E.20. Every integer divides its square.

0.E.21. The square of every rational number is a rational number.

0.E.22. The square root of every positive rational number is a rational number.

0.E.23. When an integer a is divided by a positive integer b, the remainder is always less than a.

0.E.24. Every quadratic polynomial has two distinct complex roots.

Always–Sometimes–Never questions

In Questions 0.E.25 to 0.E.32, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

0.E.25. Let $n, b_1, b_2 \in \mathbb{N}$ with $1 < n < b_1 < b_2$. Then the base-b_1 expansion of n is equal to the base-b_2 expansion of n.

0.E.26. Let $n, b_1, b_2 \in \mathbb{N}$ with $1 < b_1 < b_2 < n$. Then the base-b_1 expansion of n is equal to the base-b_2 expansion of n.

0.E.27. Let $a, b, c \in \mathbb{Z}$ and suppose that a divides c and b divides c. Then ab divides c.

0.E.28. Let $a, b, c \in \mathbb{Z}$ and suppose that a divides c and b divides c. Then ab divides c^2.

0.E.29. Let $x, y \in \mathbb{Q}$ and let $a, b, c, d \in \mathbb{Z}$ with $cy + d \neq 0$. Then $\dfrac{ax+b}{cy+d} \in \mathbb{Q}$.

0.E.30. Let $\dfrac{a}{b}$ be a rational number. Then $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$.

0.E.31. Let $x \in \mathbb{R}$ and assume that $x^2 \in \mathbb{Q}$. Then $x \in \mathbb{Q}$.

0.E.32. Let $x \in \mathbb{R}$ and assume that $x^2 + 1 \in \mathbb{Q}$ and $x^5 + 1 \in \mathbb{Q}$. Then $x \in \mathbb{Q}$.

0.E.33. Let $p(x) = ax^2 + bx + c$ be a polynomial with $a, b, c \in \mathbb{R}$ and $a \neq 0$, and suppose that $u + vi$ be a complex root of $p(x)$ with $v \neq 0$. Then $u - vi$ is a root of $p(x)$.

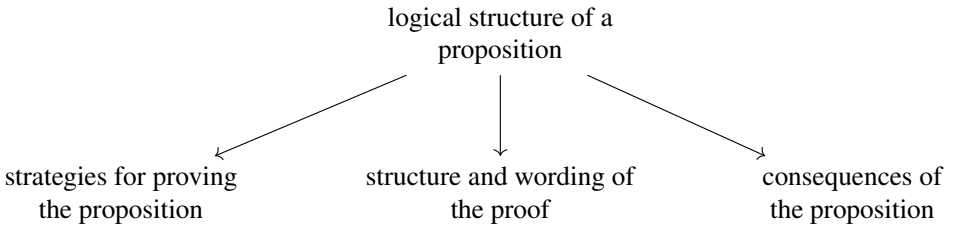
Part I

Core concepts

Chapter 1

Logical structure

The goal of this chapter is to develop a methodical way of breaking up a proposition into smaller components and seeing how these components fit together—this is called the *logical structure* of a proposition. The logical structure of a proposition is very informative: it tells us what we need to do in order to prove it, what we need to write in order to communicate our proof, and how to explore the consequences of the proposition after it has been proved.



Sections 1.1 and 1.2 are dedicated to developing a system of *symbolic logic* for reasoning about propositions. We will be able to represent a proposition using a string of variables and symbols, and this expression will guide how we can prove the proposition and explore its consequences. In Section 1.3 we will develop techniques for manipulating these logical expressions algebraically—this, in turn, will yield new proof techniques (some have fancy names like 'proof by contraposition', but some do not).

Exploring how the logical structure of a proposition informs the structure and wording of its proof is the content of Section A.2.

Section 1.1

Propositional logic

Every mathematical proof is written in the context of certain *assumptions* being made, and certain *goals* to be achieved.

- **Assumptions** are the propositions which are known to be true, or which we are assuming to be true for the purposes of proving something. They include theorems that have already been proved, prior knowledge which is assumed of the reader, and assumptions which are explicitly made using words like 'suppose' or 'assume'.
- **Goals** are the propositions we are trying to prove in order to complete the proof of a result, or perhaps just a step in the proof.

With every phrase we write, our assumptions and goals change. This is perhaps best illustrated by example. In Example 1.1.1 below, we will examine the proof of Proposition 0.39 in detail, so that we can see how the words we wrote affected the assumptions and goals at each stage in the proof. We will indicate our assumptions and goals at any given stage using tables—the assumptions listed will only be those assumptions which are made explicitly; prior knowledge and previously proved theorems will be left implicit.

Example 1.1.1

The statement of Proposition 0.39 was as follows:

> Let $a, b, c \in \mathbb{Z}$. If c divides b and b divides a, then c divides a.

The set-up of the proposition instantly gives us our initial assumptions and goals:

Assumptions	Goals
$a, b, c \in \mathbb{Z}$	If c divides b and b divides a, then c divides a

We will now proceed through the proof, line by line, to see what effect the words we wrote had on the assumptions and goals.

Since our goal was an expression of the form 'if…then…', it made sense to start by assuming the 'if' statement, and using that assumption to prove the 'then' statement. As such, the first thing we wrote in our proof was:

> Suppose that c divides b and b divides a.

Our updated assumptions and goals are reflected in the following table.

Assumptions	Goals
$a, b, c \in \mathbb{Z}$ c divides b b divides a	c divides a

Our next step in the proof was to unpack the definitions of 'c divides b' and 'b divides a', giving us more to work with.

> Suppose that c divides b and b divides a. By Definition 0.36, it follows that
>
> $$b = qc \quad \text{and} \quad a = rb$$
>
> for some integers q and r.

This introduces two new variables q, r and allows us to replace the assumptions 'c divides b' and 'b divides a' with their definitions.

Assumptions	**Goals**
$a, b, c, q, r \in \mathbb{Z}$ $b = qc$ $a = rb$	c divides a

At this point we have pretty much exhausted all of the assumptions we can make, and so our attention turns towards the goal—that is, we must prove that c divides a. At this point, it helps to 'work backwards' by unpacking the goal: what does it mean for c to divide a? Well, by Definition 0.36, we need to prove that a is equal to some integer multiplied by c—this will be reflected in the following table of assumptions and goals.

Since we are now trying to express a in terms of c, it makes sense to use the equations we have relating a with b, and b with c, to relate a with c.

> Suppose that c divides b and b divides a. By Definition 0.36, it follows that
>
> $$b = qc \quad \text{and} \quad a = rb$$
>
> for some integers q and r. Using the first equation, we may substitute qc for b in the second equation, to obtain
>
> $$a = r(qc)$$

We are now very close, as indicated in the following table.

Assumptions	**Goals**
$a, b, c, q, r \in \mathbb{Z}$ $b = qc$ $a = rb$ $a = r(qc)$	$a = [\text{some integer}] \cdot c$

Our final step was to observe that the goal has at last been achieved:

Suppose that c divides b and b divides a. By Definition 0.36, it follows that

$$b = qc \quad \text{and} \quad a = rb$$

for some integers q and r. Using the first equation, we may substitute qc for b in the second equation, to obtain

$$a = r(qc)$$

But $r(qc) = (rq)c$, and rq is an integer,

Assumptions	Goals
$a, b, c, q, r \in \mathbb{Z}$	
$b = qc$	
$a = rb$	
$a = r(qc)$	
$a = (rq)c$	
$rq \in \mathbb{Z}$	

Now that there is nothing left to prove, it is helpful to reiterate that point so that the reader has some closure on the matter.

Suppose that c divides b and b divides a. By Definition 0.36, it follows that

$$b = qc \quad \text{and} \quad a = rb$$

for some integers q and r. Using the first equation, we may substitute qc for b in the second equation, to obtain

$$a = r(qc)$$

But $r(qc) = (rq)c$, and rq is an integer, so it follows from Definition 0.36 that c divides a.

Symbolic logic

Consider again the proposition that we proved in Proposition 0.39 (for given integers a, b, c):

If c divides b and b divides a, then c divides a.

The three statements 'c divides b', 'b divides a' and 'c divides a' are all propositions in their own right, despite the fact that they all appear inside a more complex proposition. We can examine the logical structure of the proposition by replacing these simpler propositions with symbols, called *propositional variables*. Writing P to represent 'c divides b', Q to represent 'b divides a' and R to represent 'c divides a', we obtain:

If P and Q, then R.

Breaking down the proposition in this way makes it clear that a feasible way to prove it is to *assume* P and Q, and then *derive* R from these assumptions—this is exactly what we did in the proof, which we examined in great detail in Example 1.1.1. But importantly, it suggests that the same proof strategy might work for other propositions which are also of the form 'if P and Q, then R', such as the following proposition (for a given integer n):

If $n > 2$ and n is prime, then n is odd.

Observe that the simpler propositions are joined together to form a more complex proposition using language, namely the word 'and' and the construction 'if… then…'—we will represent these constructions symbolically using *logical operators*, which will be introduced in Definition 1.1.3.

Zooming in even more closely, we can use Definition 0.36 to observe that 'c divides b' really means '$b = qc$ for some $q \in \mathbb{Z}$'. The expression 'for some $q \in \mathbb{Z}$' introduces a new variable q, which we must deal with appropriately in our proof. Words which we attach to variables in our proofs—such as 'any', 'exists', 'all', 'some', 'unique' and 'only'—will be represented symbolically using *quantifiers*, which we will study in Section 1.2.

By breaking down a complex proposition into simpler statements which are connected together using logical operators and quantifiers, we can more precisely identify what assumptions we can make at any given stage in a proof of the proposition, and what steps are needed in order to finish the proof.

Propositional formulae

We begin our development of symbolic logic with some definitions to fix our terminology.

Definition 1.1.2

A **propositional variable** is a symbol that represents a proposition. Propositional variables may be assigned **truth values** ('true' or 'false').

We will typically use the lower-case letters p, q, r and s as our propositional variables.

We will be able to form more complex expressions representing propositions by connecting together simpler ones using *logical operators* such as $\wedge$ (which represents 'and'), $\vee$ (which represents 'or'), $\Rightarrow$ (which represents 'if…then…') and $\neg$ (which represents 'not').

The definition of the notions of *logical operator* and *propositional formula* given below is a little bit difficult to digest, so it is best understood by considering examples of propositional formulae and instances of logical operators. Fortunately we will see plenty of these, since they are the central objects of study for the rest of this section.

Definition 1.1.3

A **propositional formula** is an expression that is either a propositional variable, or is built up from simpler propositional formulae ('subformulae') using a **logical operator**. In the latter case, the truth value of the propositional formula is determined by the truth values of the subformulae according to the rules of the logical operator.

On first sight, Definition 1.1.3 seems circular—it defines the term 'propositional formula' in terms of propositional formulae! But in fact it is not circular; it is an example of a *recursive* definition (we avoid circularity with the word 'simpler'). To illustrate, consider the following example of a propositional formula:

$$(p \wedge q) \Rightarrow r$$

This expression represents a proposition of the form 'if p and q, then r', where p, q, r are themselves propositions. It is built from the subformulae $p \wedge q$ and r using the logical operator $\Rightarrow$, and $p \wedge q$ is itself built up from the subformulae p and q using the logical operator $\wedge$.

The truth value of $(p \wedge q) \Rightarrow r$ is then determined by the truth values of the constituent propositional variables (p, q and r) according to the rules for the logical operators $\wedge$ and $\Rightarrow$.

If this all seems a bit abstract, that is because it *is* abstract, and you are forgiven if it makes no sense to you yet. From this point onwards, we will only study particular instances of logical operators, which will make it all much easier to understand.

Conjunction ('and', $\wedge$)

Conjunction is the logical operator which makes precise what we mean when we say 'and'.

Definition 1.1.4

The **conjunction** operator is the logical operator $\wedge$ (LaTeX code: `\wedge`), defined according to the following rules:

- ($\wedge$I) If p is true and q is true, then $p \wedge q$ is true;
- ($\wedge$E$_1$) If $p \wedge q$ is true, then p is true;
- ($\wedge$E$_2$) If $p \wedge q$ is true, then q is true.

The expression $p \wedge q$ represents 'p and q'.

It is not always obvious when conjunction is being used; sometimes it sneaks in without the word 'and' ever being mentioned! Be on the look-out for occasions like this, such as in the following exercise.

Example 1.1.5

We can express the proposition '7 is a prime factor of 28' in the form $p \wedge q$, by letting p represent the proposition '7 is prime' and letting q represent the proposition '7 divides 28'.

Exercise 1.1.6

Express the proposition 'John is a mathematician who lives in Pittsburgh' in the form $p \wedge q$, for propositions p and q.

The rules in Definition 1.1.4 are examples of *rules of inference*—they tell us how to deduce (or 'infer') the truth of one propositional formula from the truth of other propositional formulae. In particular, rules of inference never directly tell us when a proposition is *false*—in order to prove something is false, we will prove its *negation* is true (see Definition 1.1.35).

Rules of inference tell us how to use the logical structure of propositions in proofs:

- The rule ($\wedge$I) is an *introduction rule*, meaning that it tells us how to *prove a goal* of the form $p \wedge q$—indeed, if we want to prove that $p \wedge q$ is true, ($\wedge$I) tells us that it suffices to prove that p is true and prove that q is true.
- The rules ($\wedge$E$_1$) and ($\wedge$E$_2$) are *elimination rules*, meaning that they tell us how to *use an assumption* of the form $p \wedge q$—indeed, if we are assuming that $p \wedge q$ is true, we are then free to use the fact that p is true and the fact that q is true.

Each logical operator will come equipped with some introduction and/or elimination rules, which tell us how to prove goals or use assumptions which include the logical operator in question. It is in this way that the logical structure of a proposition informs *proof strategies*, like the following:

❖ **Strategy 1.1.7** · *Proving conjunctions*
A proof of the proposition $p \wedge q$ can be obtained by tying together two proofs, one being a proof that p is true and one being a proof that q is true.

✎ **Example 1.1.8**
Suppose we are required to prove that 7 is a prime factor of 28. In Example 1.1.5 we expressed '7 is a prime factor of 28' as the conjunction of the propositions '7 is prime' and '7 divides 28', and so Strategy 1.1.7 breaks down the proof into two steps: first prove that 7 is prime, and then prove that 7 divides 28.

Much like Strategy 1.1.7 was informed by the introduction rule for $\wedge$, the elimination rules inform how we may make use of an assumption involving a conjunction.

❖ **Strategy 1.1.9** · *Assuming conjunctions*
If an assumption in a proof has the form $p \wedge q$, then we may assume p and assume q in the proof.

✎ **Example 1.1.10**
Suppose that, somewhere in the process of proving a proposition, we arrive at the fact that 7 is a prime factor of 28. Strategy 1.1.9 then allows us to use the separate facts that 7 is prime and that 7 divides 28.

Strategies 1.1.7 and 1.1.9 seem almost *obvious*. To an extent they are obvious, and that is why we are stating them first. But the real reason we are going through the process of precisely defining logical operators, their introduction and elimination rules, and the corresponding proof strategies, is that when you are in the middle of the proof of a complicated result, it is all too easy to lose track of what you have already proved and what remains to be proved. Keeping track of the assumptions and goals in a proof, and understanding what must be done in order to complete the proof, is a difficult task.

To avoid drawing this process out too long, we need a compact way of expressing rules of inference that allows us to simply read off corresponding proof strategies. We *could* use tables of assumptions and goals like in Example 1.1.1, but this quickly becomes clunky—indeed, even the very simple conjunction introduction rule ($\wedge$I) doesn't look very nice in this format:

Assumptions	Goals		Assumptions	Goals
⋮	$p \wedge q$	$\rightsquigarrow$	⋮	p
⋮			⋮	q

Instead, we will represent rules of inference in the style of *natural deduction*. In this style, we write the *premises* $p_1, p_2, \ldots, p_k$ of a rule above a line, with a single *conclusion* q below the line, representing the assertion that the truth of a proposition q follows from the truth of (all of) the premises $p_1, p_2, \ldots, p_k$.

$$\frac{p_1 \qquad p_2 \qquad \cdots \qquad p_k}{q}$$

For instance, the introduction and elimination rules for conjunction can be expressed concisely follows:

$$\frac{p \qquad q}{p \wedge q}\ (\wedge \mathrm{I}) \qquad \frac{p \wedge q}{p}\ (\wedge \mathrm{E}_1) \qquad \frac{p \wedge q}{q}\ (\wedge \mathrm{E}_2)$$

In addition to its clean and compact nature, this way of writing rules of inference is useful because we can combine them into *proof trees* in order to see how to prove more complicated propositions. For example, consider the following proof tree, which combines two instances of the conjunction introduction rule.

$$\dfrac{\dfrac{p \qquad q}{p \wedge q} \qquad r}{(p \wedge q) \wedge r}$$

From this proof tree, we obtain a strategy for proving a proposition of the form $(p \wedge q) \wedge r$. Namely, first prove p and prove q, to conclude $p \wedge q$; and then prove r, to conclude $(p \wedge q) \wedge r$. This illustrates that the logical structure of a proposition informs how we may structure a proof of the proposition.

✎ **Exercise 1.1.11**

Write a proof tree whose conclusion is the propositional formula $(p \wedge q) \wedge (r \wedge s)$, where p, q, r, s are propositional variables. Express '2 is an even prime number and 3 is an odd prime number' in the form $(p \wedge q) \wedge (r \wedge s)$, for appropriate propositions p, q, r and s, and describe how your proof tree suggests what a proof might look like.

Disjunction ('or', $\vee$)

✦ **Definition 1.1.12**

The **disjunction** operator is the logical operator $\vee$ (LaTeX code: `\vee`), defined according to the following rules:

- ($\vee I_1$) If p is true, then $p \vee q$ is true;
- ($\vee I_2$) If q is true, then $p \vee q$ is true;
- ($\vee E$) If $p \vee q$ is true, and if r can be derived from p and from q, then r is true.

The expression $p \vee q$ represents 'p or q'.

The introduction and elimination rules for disjunction are represented diagrammatically as follows.

$$\frac{p}{p \vee q}\,(\vee I_1) \qquad \frac{q}{p \vee q}\,(\vee I_2) \qquad \frac{p \vee q \qquad \begin{matrix}[p] \\ \vdots \\ r\end{matrix} \qquad \begin{matrix}[q] \\ \vdots \\ r\end{matrix}}{r}\,(\vee E)$$

We will discuss what the notation $[p] \rightsquigarrow r$ and $[q] \rightsquigarrow r$ means momentarily. First, we zoom in on how the disjunction introduction rules inform proofs of propositions of the form 'p or q'.

❖ **Strategy 1.1.13** · *Proving disjunctions*
In order to prove a proposition of the form $p \vee q$, it suffices to prove just one of p or q.

✎ **Example 1.1.14**
Suppose we want prove that 8192 is not divisible by 3. We know by the division theorem (Theorem 0.42) that an integer is not divisible by 3 if and only if it leaves a remainder of 1 or 2 when divided by 3, and so it suffices to prove the following:

$$\begin{array}{c}\text{8192 leaves a remainder of 1} \\ \text{when divided by 3}\end{array} \quad \vee \quad \begin{array}{c}\text{8192 leaves a remainder of 2} \\ \text{when divided by 3}\end{array}$$

A quick computation reveals that $8192 = 2730 \times 3 + 2$, so that 8192 leaves a remainder of 2 when divided by 3. By Strategy 1.1.13, the proof is now complete, since the full disjunction follows by ($\vee I_2$).

✎ **Example 1.1.15**
Let p, q, r, s be propositional variables. The propositional formula $(p \vee q) \wedge (r \vee s)$ represents 'p or q, and r or s'. What follows are two examples of truth trees for this propositional formula.

$$\frac{\dfrac{p}{p \vee q}\,(\vee I_1) \qquad \dfrac{r}{r \vee s}\,(\vee I_1)}{(p \vee q) \wedge (r \vee s)}\,(\wedge I) \qquad \frac{\dfrac{q}{p \vee q}\,(\vee I_2) \qquad \dfrac{s}{r \vee s}\,(\vee I_2)}{(p \vee q) \wedge (r \vee s)}\,(\wedge I)$$

The proof tree on the left suggests the following proof strategy for $(p \vee q) \wedge (r \vee s)$. First prove p, and deduce $p \vee q$; then prove r, and deduce $r \vee s$; and finally deduce $(p \vee q) \wedge (r \vee s)$. The proof tree on the right suggests a different strategy, where $p \vee q$ is deduced by proving q instead of p, and $r \vee s$ is deduced by proving s instead of r.

Selecting which (if any) of these to use in a proof might depend on what we are trying to prove. For example, for a fixed natural number n, let p represent 'n is even', let q represent 'n is odd', let r represent '$n \geqslant 2$' and let s represent 'n is a perfect square'. Proving $(p \vee q) \wedge (r \vee s)$ when $n = 2$ would be most easily done using the left-hand proof tree above, since p and r are evidently true when $n = 2$. However, the second proof tree would be more appropriate for proving $(p \vee q) \wedge (r \vee s)$ when $n = 1$.

❖ **Aside**

If you haven't already mixed up $\wedge$ and $\vee$, you probably will soon, so here's a way of remembering which is which:

fish n chips

If you forget whether it's $\wedge$ or $\vee$ that means 'and', just write it in place of the 'n' in 'fish n chips':

fish $\wedge$ chips fish $\vee$ chips

Clearly the first looks more correct, so $\wedge$ means 'and'. This mnemonic can be modified to accommodate a wide variety of dietary restrictions; for instance 'quinoa n kale'.

Recall the diagrammatic statement of the disjunction elimination rule:

$$\dfrac{p \vee q \qquad \begin{matrix}[p] \\ \vdots \\ r\end{matrix} \qquad \begin{matrix}[q] \\ \vdots \\ r\end{matrix}}{r}\ (\vee\text{E})$$

The curious notation $[p] \rightsquigarrow r$ indicates that p is a *temporary assumption*. In the part of the proof corresponding to $[p] \rightsquigarrow r$, we would assume that p is true and derive r from that assumption, and remove the assumption that p is true from that point onwards. Likewise for $[q] \rightsquigarrow r$.

The proof strategy obtained from the disjunction elimination rule is called *proof by cases*.

❖ **Strategy 1.1.16** · *Assuming disjunctions—proof by cases*

If an assumption in a proof has the form $p \vee q$, then we may derive a proposition r by splitting into two cases: first, derive r from the temporary assumption that p is true, and then derive r from the assumption that q is true.

The following example illustrates how Strategies 1.1.13 and 1.1.16 can be used together in a proof.

✎ **Example 1.1.17**

Let n be a positive proper factor of 4, and suppose we want to prove that n is either even or a perfect square.

- Our assumption that n is a positive proper factor of 4 can be expressed as the disjunction $n = 1 \vee n = 2$.

- Our goal is to prove the disjunction 'n is even $\vee$ n is a perfect square'.

According to Strategy 1.1.16, we split into two cases, one in which $n = 1$ and one in which $n = 2$. In each case, we must derive 'n is even $\vee$ n is a perfect square', for which it suffices by Strategy 1.1.13 to derive either that n is even or that n is a perfect square. Thus a proof might look something like this:

> Since n is a positive proper factor of 4, either $n = 1$ or $n = 2$.
>
> - **Case 1.** Suppose $n = 1$. Then since $1^2 = 1$ we have $n = 1^2$, so that n is a perfect square.
> - **Case 2.** Suppose $n = 2$. Then since $2 = 2 \times 1$, we have that n is even.
>
> Hence n is either even or a perfect square.

Notice that in both Case 1 and Case 2, we did not explicitly mention that we had proved that 'n is even $\vee$ n is a perfect square', leaving that deduction to the reader—we only mentioned it after the proofs in each case were complete.

The proof of Proposition 1.1.18 below splits into *three* cases, rather than just two.

Proposition 1.1.18
Let $n \in \mathbb{Z}$. Then n^2 leaves a remainder of 0 or 1 when divided by 3.

Proof
Let $n \in \mathbb{Z}$. By the division theorem (Theorem 0.42), one of the following must be true for some $k \in \mathbb{Z}$:

$$n = 3k \quad \text{or} \quad n = 3k+1 \quad \text{or} \quad n = 3k+2$$

- Suppose $n = 3k$. Then
$$n^2 = (3k)^2 = 9k^2 = 3 \cdot (3k^2)$$
So n^2 leaves a remainder of 0 when divided by 3.
- Suppose $n = 3k+1$. Then
$$n^2 = (3k+1)^2 = 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1$$
So n^2 leaves a remainder of 1 when divided by 3.
- Suppose $n = 3k+2$. Then
$$n^2 = (3k+2)^2 = 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1$$
So n^2 leaves a remainder of 1 when divided by 3.

In all cases, n^2 leaves a remainder of 0 or 1 when divided by 3. □

Note that in the proof of Proposition 1.1.18, unlike in Example 1.1.17, we did not explicitly use

the word 'case', even though we were using proof by cases. Whether or not to make your proof strategies explicit is up to you—discussion of this kind of matter can be found in Section A.2.

When completing the following exercises, try to keep track of exactly where you use the introduction and elimination rules that we have seen so far.

✎ **Exercise 1.1.19**
Let n be an integer. Prove that n^2 leaves a remainder of 0, 1 or 4 when divided by 5.

✎ **Exercise 1.1.20**
Let $a, b \in \mathbb{R}$ and suppose $a^2 - 4b \neq 0$. Let α and β be the (distinct) roots of the polynomial $x^2 + ax + b$. Prove that there is a real number c such that either $\alpha - \beta = c$ or $\alpha - \beta = ci$.

Implication ('if…then…', $\Rightarrow$)

✦ **Definition 1.1.21**
The **implication** operator is the logical operator $\Rightarrow$ (LaTeX code: `\Rightarrow`), defined according to the following rules:

- ($\Rightarrow$I) If q can be derived from the assumption that p is true, then $p \Rightarrow q$ is true;
- ($\Rightarrow$E) If $p \Rightarrow q$ is true and p is true, then q is true.

The expression $p \Rightarrow q$ represents 'if p, then q'.

$$\dfrac{\begin{array}{c}[p] \\ \vdots \\ q\end{array}}{p \Rightarrow q}\ (\Rightarrow\text{I}) \qquad \dfrac{p \Rightarrow q \qquad p}{q}\ (\Rightarrow\text{E})$$

❖ **Strategy 1.1.22** · *Proving implications*
In order to prove a proposition of the form $p \Rightarrow q$, it suffices to assume that p is true, and then derive q from that assumption.

The following proposition illustrates how Strategy 1.1.22 can be used in a proof.

✣ **Proposition 1.1.23**
Let x and y be real numbers. If x and $x + y$ are rational, then y is rational.

Proof
Suppose x and $x + y$ are rational. Then there exist integers a, b, c, d with $b, d \neq 0$ such that

$$x = \frac{a}{b} \quad \text{and} \quad x + y = \frac{c}{d}$$

It then follows that

$$y = (x+y) - x = \frac{c}{d} - \frac{a}{b} = \frac{bc-ad}{bd}$$

Since $bc-ad$ and bd are integers, and $bd \neq 0$, it follows that y is rational. □

The key phrase in the above proof was 'Suppose x and $x+y$ are rational.' This introduced the assumptions $x \in \mathbb{Q}$ and $x+y \in \mathbb{Q}$, and reduced our goal to that of deriving a proof that y is rational—this was the content of the rest of the proof.

✎ **Exercise 1.1.24**

Let $p(x)$ be a polynomial over $\mathbb{C}$ (see Definition 0.62). Prove that if α is a root of $p(x)$, and $a \in \mathbb{C}$, then α is a root of $(x-a)p(x)$.

The elimination rule for implication (⇒E) is more commonly known by the Latin name *modus ponens*.

❖ **Strategy 1.1.25** · *Assuming implications—modus ponens*

If an assumption in a proof has the form $p \Rightarrow q$, and p is also assumed to be true, then we may deduce that q is true.

Strategy 1.1.16 is frequently used to reduce a more complicated goal to a simpler one. Indeed, if we know that $p \Rightarrow q$ is true, and if p is easy to verify, then it allows us to prove q by proving p instead.

✐ **Example 1.1.26**

Let $f(x) = x^2 + ax + b$ be a polynomial with $a, b \in \mathbb{R}$, and let $\Delta = a^2 - 4b$ be its discriminant. In Exercise 0.61 we proved that:

(i) If $\Delta > 0$, then f has two real roots;

(ii) If $\Delta = 0$, then f has one real root;

(iii) If $\Delta < 0$, then f has no real roots.

Given the polynomial $f(x) = x^2 - 68x + 1156$, it would be a pain to go through the process of solving the equation $f(x) = 0$ in order to determine how many real roots f has. However, each of the propositions (i), (ii) and (iii) take the form $p \Rightarrow q$, so Strategy 1.1.25 reduces the problem of finding how many real roots f has to that of evaluating Δ and comparing it with 0. And indeed, $(-68)^2 - 4 \times 1156 = 0$, so the implication (ii) together with (⇒E) tell us that f has one real root.

A common task faced by mathematicians is to prove that two conditions are equivalent. For example, given a polynomial $f(x) = x^2 + ax + b$ with $a, b \in \mathbb{R}$, we know that *if* $a^2 - 4b > 0$ *then* f has two real roots, but is it also true that if f has two real roots then $a^2 - 4b > 0$? (The answer is 'yes'.) The relationship between these two implications is that each is the *converse* of the other.

✦ **Definition 1.1.27**

The **converse** of a proposition of the form $p \Rightarrow q$ is the proposition $q \Rightarrow p$.

A quick remark on terminology is pertinent. The following table summarises some common ways of referring to the propositions '$p \Rightarrow q$' and '$q \Rightarrow p$'.

$p \Rightarrow q$	$q \Rightarrow p$
if p, then q	if q, then p
p only if q	p if q
p is sufficient for q	p is necessary for q

We so often encounter the problem of proving both an implication and its converse that we introduce a new logical operator that represents the conjunction of both.

Definition 1.1.28

The **biconditional** operator is the logical operator $\Leftrightarrow$ (LaTeX code: `\Leftrightarrow`), defined by declaring $p \Leftrightarrow q$ to mean $(p \Rightarrow q) \wedge (q \Rightarrow p)$. The expression $p \Leftrightarrow q$ represents 'p if and only if q'.

Many examples of biconditional statements come from solving equations; indeed, to say that the values $\alpha_1, \ldots, \alpha_n$ are the solutions to a particular equation is precisely to say that

$$x \text{ is a solution} \quad \Leftrightarrow \quad x = \alpha_1 \text{ or } x = \alpha_2 \text{ or } \cdots \text{ or } x = \alpha_n$$

Example 1.1.29

We find all real solutions x to the equation

$$\sqrt{x-3} + \sqrt{x+4} = 7$$

Let's rearrange the equation to find out what the possible solutions may be.

$$\begin{aligned}
&\sqrt{x-3} + \sqrt{x+4} = 7 \\
\Rightarrow\ & (x-3) + 2\sqrt{(x-3)(x+4)} + (x+4) = 49 && \text{squaring} \\
\Rightarrow\ & 2\sqrt{(x-3)(x+4)} = 48 - 2x && \text{rearranging} \\
\Rightarrow\ & 4(x-3)(x+4) = (48-2x)^2 && \text{squaring} \\
\Rightarrow\ & 4x^2 + 4x - 48 = 2304 - 192x + 4x^2 && \text{expanding} \\
\Rightarrow\ & 196x = 2352 && \text{rearranging} \\
\Rightarrow\ & x = 12 && \text{dividing by 196}
\end{aligned}$$

You might be inclined to stop here. Unfortunately, all we have proved is that, given a real number x, *if* x solves the equation $\sqrt{x-3} + \sqrt{x+4} = 7$, *then* $x = 12$. This narrows down the set of possible solutions to just one candidate—but we still need to check the converse, namely that *if* $x = 12$, *then* x is a solution to the equation.

As such, to finish off the proof, note that

$$\sqrt{12-3} + \sqrt{12+4} = \sqrt{9} + \sqrt{16} = 3 + 4 = 7$$

and so the value $x = 12$ is indeed a solution to the equation.

The last step in Example 1.1.29 may have seemed a little bit silly; but Example 1.1.30 demonstrates that proving the converse when solving equations truly is necessary.

Example 1.1.30

We find all real solutions x to the equation

$$x + \sqrt{x} = 0$$

We proceed as before, rearranging the equation to find all possible solutions.

$$\begin{aligned} & x + \sqrt{x} = 0 \\ \Rightarrow\ & x = -\sqrt{x} && \text{rearranging} \\ \Rightarrow\ & x^2 = x && \text{squaring} \\ \Rightarrow\ & x(x-1) = 0 && \text{rearranging} \\ \Rightarrow\ & x = 0 \text{ or } x = 1 \end{aligned}$$

Now certainly 0 is a solution to the equation, since

$$0 + \sqrt{0} = 0 + 0 = 0$$

However, 1 is *not* a solution, since

$$1 + \sqrt{1} = 1 + 1 = 2$$

Hence it is actually the case that, given a real number x, we have

$$x + \sqrt{x} = 0 \quad \Leftrightarrow \quad x = 0$$

Checking the converse here was vital to our success in solving the equation!

Another class of examples of biconditional propositions arise in finding necessary and sufficient criteria for an integer n to be divisible by some number—for example, that an integer is divisible by 10 if and only if its base-10 expansion ends with the digit 0.

Example 1.1.31

Let $n \in \mathbb{N}$. We will prove that n is divisible by 8 if and only if the number formed of the last three digits of the base-10 expansion of n is divisible by 8.

First, we will do some 'scratch work'. Let $d_r d_{r-1} \ldots d_1 d_0$ be the base-10 expansion of n. Then

$$n = d_r \cdot 10^r + d_{r-1} \cdot 10^{r-1} + \cdots + d_1 \cdot 10 + d_0$$

Define

$$n' = d_2 d_1 d_0 \quad \text{and} \quad n'' = n - n' = d_r d_{r-1} \ldots d_4 d_3 000$$

Now $n - n' = 1000 \cdot d_r d_{r-1} \ldots d_4 d_3$ and $1000 = 8 \cdot 125$, so it follows that 8 divides n''.

Our goal is now to prove that 8 divides n if and only if 8 divides n'.

- ($\Rightarrow$) Suppose 8 divides n. Since 8 divides n'', it follows from Exercise 0.40 that 8 divides

$an + bn''$ for all $a, b \in \mathbb{Z}$. But

$$n' = n - (n - n') = n - n'' = 1 \cdot n + (-1) \cdot n''$$

so indeed 8 divides n', as required.

- ($\Leftarrow$) Suppose 8 divides n'. Since 8 divides n'', it follows from Exercise 0.40 that 8 divides $an' + bn''$ for all $a, b \in \mathbb{Z}$. But

$$n = n' + (n - n') = n' + n'' = 1 \cdot n' + 1 \cdot n''$$

so indeed 8 divides n, as required.

Exercise 1.1.32

Prove that a natural number n is divisible by 3 if and only if the sum of its base-10 digits is divisible by 3.

Negation ('not', $\neg$)

So far we only officially know how to prove that true propositions are *true*. The negation operator makes precise what we mean by 'not', which allows us to prove that false propositions are *false*.

Definition 1.1.33

A **contradiction** is a proposition that is known or assumed to be false. We will use the symbol $\bot$ (LaTeX code: `\bot`) to represent an arbitrary contradiction.

Example 1.1.34

Some examples of contradictions include the assertion that $0 = 1$, or that $\sqrt{2}$ is rational, or that the equation $x^2 = -1$ has a solution $x \in \mathbb{R}$.

Definition 1.1.35

The **negation** operator is the logical operator $\neg$ (LaTeX code: `\neg`), defined according to the following rules:

- ($\neg$I) If a contradiction can be derived from the assumption that p is true, then $\neg p$ is true;
- ($\neg$E) If $\neg p$ and p are both true, then a contradiction may be derived.

The expression $\neg p$ represents 'not p' (or 'p is false').

$$\dfrac{\begin{array}{c}[p]\\ \wr \\ \bot\end{array}}{\neg p}\ (\neg\text{I}) \qquad\qquad \dfrac{\neg p \qquad p}{\bot}\ (\neg\text{E})$$

Aside

The rules ($\neg$I) and ($\neg$E) closely resemble ($\Rightarrow$I) and ($\Rightarrow$E)—indeed, we could simply define $\neg p$ to mean '$p \Rightarrow \bot$', where $\bot$ represents an arbitrary contradiction, but it will be easier later on to have a primitive notion of negation.

The introduction rule for negation ($\neg$I) gives rise to a proof strategy called *proof by contradiction*, which turns out to be extremely useful.

Strategy 1.1.36 · *Proving negations—proof by contradiction*

In order to prove a proposition p is false (that is, that $\neg p$ is true), it suffices to assume that p is true and derive a contradiction.

The following proposition has a classic proof by contradiction.

Proposition 1.1.37

Let r be a rational number and let a be an irrational number. Then $r+a$ is irrational.

Proof

By Definition 0.48, we need to prove that $r+a$ is real and not rational. It is certainly real, since r and a are real, so it remains to prove that $r+a$ is not rational.

Suppose $r+a$ is rational. Since r is rational, it follows from Proposition 1.1.23 that a is rational, since

$$a = (r+a) - r$$

This contradicts the assumption that a is irrational. It follows that $r+a$ is not rational, and is therefore irrational. □

Now you can try proving some elementary facts by contradiction.

Exercise 1.1.38

Let $x \in \mathbb{R}$. Prove by contradiction that if x is irrational then $-x$ and $\frac{1}{x}$ are irrational.

Exercise 1.1.39

Prove by contradiction that there is no least positive real number. That is, prove that there is not a positive real number a such that $a \leqslant b$ for all positive real numbers b.

A proof need not be a 'proof by contradiction' in its entirety—indeed, it may be that only a small portion of the proof uses contradiction. This is exhibited in the proof of the following proposition.

Proposition 1.1.40

Let a be an integer. Then a is odd if and only if $a = 2b+1$ for some integer b.

Proof
Suppose a is odd. By the division theorem (Theorem 0.42), either $a = 2b$ or $a = 2b+1$, for some $b \in \mathbb{Z}$. If $a = 2b$, then 2 divides a, contradicting the assumption that a is odd; so it must be the case that $a = 2b+1$.

Conversely, suppose $a = 2b+1$. Then a leaves a remainder of 1 when divided by 2. However, by the division theorem, the even numbers are precisely those that leave a remainder of 0 when divided by 2. It follows that a is not even, so is odd. □

The elimination rule for the negation operator ($\neg$E) simply says that a proposition can't be true and false at the same time.

Strategy 1.1.41 · *Assuming negations*
If an assumption in a proof has the form $\neg p$, then any derivation of p leads to a contradiction.

The main use of Strategy 1.1.41 is for obtaining the contradiction in a proof by contradiction—in fact, we have already used it in our examples of proof by contradiction! As such, we will not dwell on it further.

Logical axioms

We wrap up this section by introducing a couple of additional logical rules (*axioms*) that we will use in our proofs.

The first is the so-called *law of excluded middle*, which appears so obvious that it is not even worth stating (let alone naming)—what it says is that every proposition is either true or false. But beware, as looks can be deceiving; the law of excluded middle is a non-constructive axiom, meaning that it should not be accepted in settings it is important to keep track of how a proposition is proved—simply knowing that a proposition is either true or false tells us nothing about how it might be proved or refuted. In most mathematical contexts, though, it is accepted without a second's thought.

Axiom 1.1.42 · *Law of excluded middle*
Let p be a propositional formula. Then $p \vee (\neg p)$ is true.

The law of excluded middle can be represented diagrammatically as follows; there are no premises above the line, since we are simply asserting that it is true.

$$\frac{}{p \vee (\neg p)}\ \text{LEM}$$

Strategy 1.1.43 · *Using the law of excluded middle*
In order to prove a proposition q is true, it suffices to split into cases based on whether some other proposition p is true or false, and prove that q is true in each case.

The proof of Proposition 1.1.44 below makes use of the law of excluded middle—note that we defined 'odd' to mean 'not even' (Definition 0.41).

Proposition 1.1.44
Let $a, b \in \mathbb{Z}$. If ab is even, then either a is even or b is even (or both).

Proof
Suppose $a, b \in \mathbb{Z}$ with ab even.

- Suppose a is even—then we're done.
- Suppose a is odd. If b is also odd, then by Proposition 1.1.40 can write

  $$a = 2k+1 \quad \text{and} \quad b = 2\ell+1$$

  for some integers k, ℓ. This implies that

  $$ab = (2k+1)(2\ell+1) = 4k\ell+2k+2\ell+1 = 2(\underbrace{2k\ell+k+\ell}_{\in\mathbb{Z}})+1$$

  so that ab is odd. This contradicts the assumption that ab is even, and so b must in fact be even.

In both cases, either a or b is even. □

Exercise 1.1.45
Reflect on the proof of Proposition 1.1.44. Where in the proof did we use the law of excluded middle? Where in the proof did we use proof by contradiction? What was the contradiction in this case? Prove Proposition 1.1.44 twice more, once using contradiction and not using the law of excluded middle, and once using the law of excluded middle and not using contradiction.

Exercise 1.1.46
Let a and b be irrational numbers. By considering the number $\sqrt{2}^{\sqrt{2}}$, prove that it is possible that a^b be rational.

Another logical rule that we will use is the *principle of explosion*, which is also known by its Latin name, *ex falso sequitur quodlibet*, which approximately translates to '*from falsity follows whatever you like*'.

Axiom 1.1.47 · *Principle of explosion*
If a contradiction is assumed, any consequence may be derived.

$$\frac{\bot}{p}\ \text{Expl}$$

The principle of explosion is a bit confusing on first sight. To shed a tiny bit of intuition on it, think of it as saying that both true and false propositions are consequences of a contradictory assumption.

For instance, suppose that $-1 = 1$. From this we can obtain consequences that are false, such as $0 = 2$ by adding 1 to both sides of the equation, and consequences that are true, such as $1 = 1$ by squaring both sides of the equation.

We will rarely use the principle of explosion directly in our mathematical proofs, but we will use it in Section 1.3 for proving logical formulae are equivalent.

Section 1.2

Variables and quantifiers

Free and bound variables

Everything we did in Section 1.1 concerned *propositions* and the logical rules concerning their proofs. Unfortunately if all we have to work with is propositions then our ability to do mathematical reasoning will be halted pretty quickly. For example, consider the following statement:

$$x \text{ is divisible by } 7$$

This statement seems like the kind of thing we should probably be able to work with if we're doing mathematics. It makes sense if x is a integer, such as 28 or 41; but it doesn't make sense at all if x is a parrot called Alex.[a] In any case, even when it does make sense, its truth value depends on x; indeed, '28 is divisible by 7' is a true proposition, but '41 is divisible by 7' is a false proposition.

This means that the statement 'x is divisible by 7' isn't a proposition—*quel horreur*! But it *almost* is a proposition: if we know that x refers somehow to an integer, then it becomes a proposition as soon as a particular numerical value of x is specified. The symbol x is called a *free variable*.

Definition 1.2.1

Let x be a variable that is understood to refer to an element of a set X. In a statement involving x, we say x is **free** if it makes sense to substitute particular elements of X in the statement; otherwise, we say x is **bound**.

To represent statements that have free variables in them abstractly, we generalise the notion of a propositional variable (Definition 1.1.2) to that of a *predicate*.

Definition 1.2.2

A **predicate** is a symbol p together with a specified list of free variables $x_1, x_2, \ldots, x_n$ (where $n \in \mathbb{N}$) and, for each free variable x_i, a specification of a set X_i called the **domain of discourse** (or **range**) of x_i. We will typically write $p(x_1, x_2, \ldots, x_n)$ in order to make the variables explicit.

The statements represented by predicates are those that become propositions when specific values are substituted for their free variables from their respective domains of discourse. For example, 'x is divisible by 7' is not a proposition, but it becomes a proposition when specific integers (such as 28 or 41) are substituted for x.

This is a lot to take in, so let's look at some examples.

[a] Alex the parrot is the only non-human animal to have ever been observed to ask an existential question; he died in September 2007 so we may never know if he was divisible by 7, but it is unlikely. According to *Time*, his last words were 'you be good, see you tomorrow, I love you'. The reader is advised to finish crying before they continue reading about variables and quantifiers.

Example 1.2.3

(i) We can represent the statement 'x is divisible by 7' discussed above by a predicate $p(x)$ whose only free variable x has $\mathbb{Z}$ as its domain of discourse. Then $p(28)$ is the true proposition '28 is divisible by 7' and $p(41)$ is the false proposition '41 is divisible by 7'.

(ii) A predicate with no free variables is precisely a propositional variable. This means that the notion of a predicate generalises that of a propositional variable.

(iii) The expression '$2^n - 1$ is prime' can be represented by a predicate $p(n)$ with one free variable n, whose domain of discourse is the set $\mathbb{N}$ of natural numbers. Then $p(3)$ is the true proposition '$2^3 - 1$ is prime' and $p(4)$ is the false proposition '$2^4 - 1$ is prime'.

(iv) The expression '$x - y$ is rational' can be represented by a predicate $q(x, y)$ with free variables x and y, whose domain of discourse is the set $\mathbb{R}$ of real numbers.

(v) The expression 'there exist integers a and b such that $x = a^2 + b^2$' has free variable x and bound variables a, b. It can be represented by a predicate $r(x)$ with one free variable x, whose domain of discourse is $\mathbb{Z}$.

(vi) The expression 'every even natural number $n \geqslant 2$ is divisible by k' has free variable k and bound variable n. It can be represented by a predicate $s(k)$ with one free variable k, whose domain of discourse is $\mathbb{N}$.

Quantifiers

Look again at the statements in parts (v) and (vi) of Example 1.2.3. Both contained bound variables, which were so because we used words like 'there exists' and 'every'—had we not used these words, those variables would be free, as in '$x = a^2 + b^2$' and 'n is divisible by k'.

Expressions that refer to *how many* elements of a set make a statement true, such as 'there exists' and 'every', turn free variables into bound variables. We represent such expressions using symbols called *quantifiers*, which are the central objects of study of this section.

The two main quantifiers used throughout mathematics are the *universal* quantifier $\forall$ and the *existential* quantifier $\exists$. We will define these quantifiers formally later in this section, but for now, the following informal definitions suffice:

- The expression '$\forall x \in X, \ldots$' denotes 'for all $x \in X, \ldots$' and will be defined formally in Definition 1.2.9;
- The expression '$\exists x \in X, \ldots$' denotes 'there exists $x \in X$ such that $\ldots$' and will be defined formally in Definition 1.2.17.

Note that we always place the quantifier *before* the statement, so even though we might write or say things like '$n = 2k$ for some integer k' or '$x^2 \geqslant 0$ for all $x \in \mathbb{R}$', we would express these statements symbolically as '$\exists k \in \mathbb{Z}, n = 2k$' and '$\forall x \in \mathbb{R}, x^2 \geqslant 0$', respectively.

We will define a third quantifier $\exists!$ in terms of $\forall$ and $\exists$ to say that there is *exactly one* element of a set making a statement true. There are plenty of other quantifiers out there, but they tend to be specific

to particular fields—examples include 'almost everywhere' in measure theory, 'almost surely' in probability theory, 'for all but finitely many' in set theory and related disciplines, and 'for fresh' in the theory of nominal sets.

Using predicates, logical operators and quantifiers, we are able to build up more complicated expressions, called *logical formulae*. Logical formulae generalise propositional formulae (Definition 1.1.3) in by allowing (free and bound) variables and quantification to occur.

Definition 1.2.4

A **logical formula** is an expression that is built from predicates using logical operators and quantifiers; it may have both free and bound variables. The truth value of a logical formula depends on its free variables according to the rules for logical operators and quantifiers.

Translating between plain English statements and purely symbolic logical formulae is an important skill to obtain:

- The plain English statements are easier to understand and are the kinds of things you would speak aloud or write down when discussing the mathematical ideas involved.
- The symbolic logical formulae are what provide the precision needed to guide a proof of the statement being discussed—we will see strategies for proving statements involving quantifiers soon.

The following examples and exercise concern translating between plain English statements and purely symbolic logical formulae.

Example 1.2.5

Recall that an integer n is even if and only if it is divisible by 2. According to Definition 0.36, that is to say that 'n is even' means '$n = 2k$ for some integer k'. Using quantifiers, we can express 'n is even' as '$\exists k \in \mathbb{Z}, n = 2k$'.

The (false) proposition 'every integer is even' can then be written symbolically as follows. First introduce a variable n to refer to an integer; to say 'every integer is even' is to say '$\forall n \in \mathbb{Z}$, n is even', and so using the symbolic representation of 'n is even', we can express 'every integer is even' as '$\forall n \in \mathbb{Z}, \exists k \in \mathbb{Z}, n = 2k$'.

Exercise 1.2.6

Find logical formulae that represent each of the following English statements.

(a) There is an integer that is divisible by every integer.

(b) There is no greatest odd integer.

(c) Between any two distinct rational numbers is a third distinct rational number.

(d) If an integer has a rational square root, then that root is an integer.

Example 1.2.7

Consider the following logical formula.

$$\forall a \in \mathbb{R},\, (a \geqslant 0 \Rightarrow \exists b \in \mathbb{R},\, a = b^2)$$

If we translate this expression symbol-for-symbol, what it says is:

For every real number a, if a is non-negative,
then there exists a real number b such that $a = b^2$.

Read in this way, it is not a particularly enlightening statement. However, we can distill the robotic nature of the symbol-for-symbol reading by thinking more carefully about what the statement *really* means.

Indeed, to say '$a = b^2$ for some real number b' is exactly to say that a has a real square root—after all, what is a square root of a if not a real number whose square is equal to a? This translation eliminates explicit reference to the bound variable b, so that the statement now reads:

For every real number a, if a is non-negative, then a has a real square root.

We're getting closer. Next note that instead of the clunky expression 'for every real number a, if a is non-negative, then …', we could just say 'for every non-negative real number a, …'.

For every non-negative real number a, a has a real square root.

Finally, we can eliminate the bound variable a by simply saying:

Every non-negative real number has a real square root.

This is now a meaningful expression that is much easier to understand than the logical formula we started with.

Exercise 1.2.8

Find statements in plain English, involving as few variables as possible, that are represented by each of the following logical formulae. (The domains of discourse of the free variables are indicated in each case.)

(a) $\exists q \in \mathbb{Z},\, a = qb$ — free variables $a, b \in \mathbb{Z}$

(b) $\exists a \in \mathbb{Z},\, \exists b \in \mathbb{Z},\, (b \neq 0 \wedge bx = a)$ — free variable $x \in \mathbb{R}$

(c) $\forall d \in \mathbb{N},\, [(\exists q \in \mathbb{Z},\, n = qd) \Rightarrow (d = 1 \vee d = n)]$ — free variable $n \in \mathbb{N}$

(d) $\forall a \in \mathbb{R},\, [a > 0 \Rightarrow \exists b \in \mathbb{R},\, (b > 0 \wedge a < b)]$ — no free variables

Now that we have a better understanding of how to translate between plain English statements and

logical formulae, we are ready to give a precise mathematical treatment of quantifiers. Just like with logical operators in Section 1.1, quantifiers will be defined according to *introduction rules*, which tell us how to prove a quantified formula, and *elimination rules*, which tell us how to use an assumption that involves a quantifier.

Universal quantification ('for all', $\forall$)

The universal quantifier makes precise what we mean when we say 'for all', or '$p(x)$ is always true no matter what value x takes'.

Definition 1.2.9

The **universal quantifier** is the quantifier $\forall$ (LaTeX code: `\forall`); if $p(x)$ is a logical formula with free variable x with range X, then $\forall x \in X,\, p(x)$ is the logical formula defined according to the following rules:

- ($\forall$I) If $p(x)$ can be derived from the assumption that x is an arbitrary element of X, then $\forall x \in X,\, p(x)$;
- ($\forall$E) If $a \in X$ and $\forall x \in X,\, p(x)$ is true, then $p(a)$ is true.

The expression $\forall x \in X,\, p(x)$ represents 'for all $x \in X,\, p(x)$'.

$$\frac{\begin{array}{c}[x \in X] \\ \vdots \\ p(x)\end{array}}{\forall x \in X,\, p(x)} \qquad \frac{\forall x \in X,\, p(x) \qquad a \in X}{p(a)}$$

Strategy 1.2.10 · *Proving universally quantified statements*

To prove a proposition of the form $\forall x \in X,\, p(x)$, it suffices to prove $p(x)$ for an arbitrary element $x \in X$—in other words, prove $p(x)$ whilst assuming nothing about the variable x other than that it is an element of X.

Useful phrases for introducing an arbitrary variable of a set X in a proof include 'fix $x \in X$' or 'let $x \in X$' or 'take $x \in X$'—more on this is discussed in Section A.2.

The proofs of the following propositions illustrate how a proof of a universally quantified statement might look.

Proposition 1.2.11

The square of every odd integer is odd.

Proof

Let n be an odd integer. Then $n = 2k+1$ for some $k \in \mathbb{Z}$ by the division theorem (Theorem 0.42), and so

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

Since $2k^2 + 2k \in \mathbb{Z}$, we have that n^2 is odd, as required. □

Note that in the proof of Proposition 1.2.11, we did not assume anything about n other than that it is an odd integer.

Proposition 1.2.12
The base-10 expansion of the square of every natural number ends in one of the digits 0, 1, 4, 5, 6 or 9.

Proof
Fix $n \in \mathbb{N}$, and let

$$n = d_r d_{r-1} \dots d_0$$

be its base-10 expansion. Write

$$n = 10m + d_0$$

where $m \in \mathbb{N}$—that is, m is the natural number obtained by removing the final digit from n. Then

$$n^2 = 100m^2 + 20md_0 + d_0^2 = 10m(10m + 2d_0) + d_0^2$$

Hence the final digit of n^2 is equal to the final digit of d_0^2. But the possible values of d_0^2 are

$$0 \quad 1 \quad 4 \quad 9 \quad 16 \quad 25 \quad 36 \quad 49 \quad 64 \quad 81$$

all of which end in one of the digits 0, 1, 4, 5, 6 or 9. □

Exercise 1.2.13
Prove that every integer is rational.

Exercise 1.2.14
Prove that every linear polynomial over $\mathbb{Q}$ has a rational root.

Exercise 1.2.15
Prove that, for all real numbers x and y, if x is irrational, then $x + y$ and $x - y$ are not both rational.

Before advancing too much further, beware of the following common error that arises when dealing with universal quantifiers.

Common error
Consider the following (non-)proof of the proposition $\forall n \in \mathbb{Z}, n^2 \geqslant 0$.

> Let n be an arbitrary integer, say $n = 17$. Then $17^2 = 289 \geqslant 0$, so the statement is true.

The error made here is that the *writer* has picked an arbitrary value of n, not the *reader*. (In fact, the above argument actually proves $\exists n \in \mathbb{Z}, n^2 \geqslant 0$.)

The proof should make no assumptions about the value of n other than that it is an integer. Here is a correct proof:

> Let n be an arbitrary integer. Either $n \geqslant 0$ or $n < 0$. If $n \geqslant 0$ then $n^2 \geqslant 0$, since the product of two nonnegative numbers is nonnegative; if $n < 0$ then $n^2 \geqslant 0$, since the product of two negative numbers is positive.

The strategy suggested by the elimination rule for the universal quantifier is one that we use almost without thinking about it.

❖ **Strategy 1.2.16** · *Assuming universally quantified statements*
If an assumption in a proof has the form $\forall x \in X, p(x)$, then we may assume that $p(a)$ is true whenever a is an element of X.

Existential quantification ('there exists', $\exists$)

✦ **Definition 1.2.17**
The **existential quantifier** is the quantifier $\exists$ (LaTeX code: `\exists`) if $p(x)$ is a logical formula with free variable x with range X, then $\exists x \in X, p(x)$ is the logical formula defined according to the following rules:

- ($\exists$I) If $a \in X$ and $p(a)$ is true, then $\exists x \in X, p(x)$;
- ($\exists$E) If $\exists x \in X, p(x)$ is true, and q can be derived from the assumption that $p(a)$ is true for some fixed $a \in X$, then q is true.

The expression $\exists x \in X, p(x)$ represents 'there exists $x \in X$ such that $p(x)$'.

$$\frac{a \in X \qquad p(a)}{\exists x \in X, p(x)}\ (\exists\text{I}) \qquad\qquad \frac{\exists x \in X, p(x) \qquad \begin{array}{c}[a \in X], [p(a)] \\ \vdots \\ q\end{array}}{q}\ (\exists\text{E})$$

❖ **Strategy 1.2.18** · *Proving existentially quantified statements*
To prove a proposition of the form $\exists x \in X, p(x)$, it suffices to prove $p(a)$ for some specific element $a \in X$, which should be explicitly defined.

✎ **Example 1.2.19**
We prove that there is a natural number that is a perfect square and is one more than a perfect cube. That is, we prove

$$\exists n \in \mathbb{N}, ([\exists k \in \mathbb{Z}, n = k^2] \wedge [\exists \ell \in \mathbb{Z}, n = \ell^3 + 1])$$

So define $n = 9$. Then $n = 3^2$ and $n = 2^3 + 1$, so that n is a perfect square and is one more than a perfect cube, as required.

The following proposition involves an existentially quantified statement—indeed, to say that a polynomial $f(x)$ has a real root is to say $\exists x \in \mathbb{R}$, $f(x) = 0$.

Proposition 1.2.20
Fix $a \in \mathbb{R}$. The cubic polynomial $x^3 + (1-a^2)x - a$ has a real root.

Proof
Let $f(x) = x^3 + (1-a^2)x - a$. Define $x = a$; then

$$f(x) = f(a) = a^3 + (1-a^2)a - a = a^3 + a - a^3 - a = 0$$

Hence a is a root of $f(x)$. Since a is real, $f(x)$ has a real root. □

The following exercises require you to prove existentially quantified statements.

Exercise 1.2.21
Prove that there is a real number which is irrational but whose square is rational.

Exercise 1.2.22
Prove that there is an integer which is divisible by zero.

Example 1.2.23
Prove that, for all $x, y \in \mathbb{Q}$, if $x < y$ then there is some $z \in \mathbb{Q}$ with $x < z < y$.

The elimination rule for the existential quantifier gives rise to the following proof strategy.

Strategy 1.2.24 · *Assuming existentially quantified statements*
If an assumption in the proof has the form $\exists x \in X$, $p(x)$, then we may introduce a new variable $a \in X$ and assume that $p(a)$ is true.

It ought to be said that when using existential elimination in a proof, the variable a used to denote a particular element of X for which $p(a)$ is true should not already be in use earlier in the proof.

Strategy 1.2.24 is very useful in proofs of divisibility, since the expression 'a divides b' is an existentially quantified statement—this was Exercise 1.2.8(a).

Proposition 1.2.25
Let $n \in \mathbb{Z}$. If n^3 is divisible by 3, then $(n+1)^3 - 1$ is divisible by 3.

Proof
Suppose n^3 is divisible by 3. Take $q \in \mathbb{Z}$ such that $n^3 = 3q$. Then

$$\begin{aligned}
&(n+1)^3 - 1 \\
&= (n^3 + 3n^2 + 3n + 1) - 1 && \text{expanding} \\
&= n^3 + 3n^2 + 3n && \text{simplifying} \\
&= 3q + 3n^2 + 3n && \text{since } n^3 = 3q \\
&= 3(q + n^2 + n) && \text{factorising}
\end{aligned}$$

Since $q + n^2 + n \in \mathbb{Z}$, we have proved that $(n+1)^3 - 1$ is divisible by 3, as required. □

Uniqueness

The concept of uniqueness arises whenever we want to use the word 'the'. For example, in Definition 0.31 we defined the base-b expansion of a natural number n to be *the* string $d_r d_{r-1} \dots d_1 d_0$ satisfying some properties. The issue with the word 'the' here is that we don't know ahead of time whether a natural number n may have base-b expansions other than $d_r d_{r-1} \dots d_1 d_0$—this fact actually requires proof. To prove this fact, we would need to assume that $e_s e_{s-1} \dots e_1 e_0$ were another base-b expansion of n, and prove that the strings $d_r d_{r-1} \dots d_1 d_0$ and $e_s e_{s-1} \dots e_1 e_0$ are the same—this is done in Theorem 7.3.51.

Uniqueness is typically coupled with *existence*, since we usually want to know if there is *exactly one* object satisfying a property. This motivates the definition of the *unique existential* quantifier, which encodes what we mean when we say 'there is exactly one $x \in X$ such that $p(x)$ is true'. The 'existence' part ensures that at least one $x \in X$ makes $p(x)$ true; the 'uniqueness' part ensures that x is the only element of X making $p(x)$ true.

Definition 1.2.26
The **unique existential quantifier** is the quantifier $\exists!$ ((LaTeX code: `\exists!`)) defined such that $\exists! x \in X,\, p(x)$ is shorthand for

$$\underbrace{(\exists x \in X,\, p(x))}_{\text{existence}} \wedge \underbrace{(\forall a \in X, \forall b \in X,\, [(p(a) \wedge p(b)) \Rightarrow a = b])}_{\text{uniqueness}}$$

Example 1.2.27
Every positive real number has a unique positive square root. We can write this symbolically as

$$\forall a \in \mathbb{R},\, (a > 0 \Rightarrow \exists! b \in \mathbb{R},\, (b > 0 \wedge b^2 = a))$$

Reading this from left to right, this says: for every real number a, if a is positive, then there exists a unique real number b, which is positive and whose square is a.

Discussion 1.2.28

Explain why Definition 1.2.26 captures the notion of there being 'exactly one' element $x \in X$ making $p(x)$ true. Can you think of any other ways that $\exists! x \in X,\, p(x)$ could be defined?

Strategy 1.2.29 · *Proving unique-existentially quantified statements*

A proof of a statement of the form $\exists! x \in X,\, p(x)$, consists of two parts:

- **Existence** — prove that $\exists x \in X,\, p(x)$ is true (e.g. using Strategy 1.2.18);
- **Uniqueness** — let $a, b \in X$, assume that $p(a)$ and $p(b)$ are true, and derive $a = b$.

Alternatively, prove existence to obtain a fixed $a \in X$ such that $p(a)$ is true, and then prove $\forall x \in X,\, [p(x) \Rightarrow x = a]$.

Example 1.2.30

We prove Example 1.2.27, namely that for each real $a > 0$ there is a unique $b > 0$ such that $b^2 = a$. So first fix $a > 0$.

- (**Existence**) The real number $\sqrt{a}$ is positive and satisfies $(\sqrt{a})^2 = a$ by definition. Its existence will be deferred to a later time, but an informal argument for its existence could be provided using 'number line' arguments as in Chapter 0.
- (**Uniqueness**) Let $y, z > 0$ be real numbers such that $y^2 = a$ and $z^2 = a$. Then $y^2 = z^2$. Rearranging and factorising yields
$$(y - z)(y + z) = 0$$
so either $y - z = 0$ or $y + z = 0$. If $y + z = 0$ then $z = -y$, and since $y > 0$, this means that $z < 0$. But this contradicts the assumption that $z > 0$. As such, it must be the case that $y - z = 0$, and hence $y = z$, as required.

Exercise 1.2.31

For each of the propositions, write it out as a logical formula involving the $\exists!$ quantifier and then prove it, using the structure of the logical formula as a guide.

(a) For each real number a, the equation $x^2 + 2ax + a^2 = 0$ has exactly one real solution x.

(b) There is a unique real number a for which the equation $x^2 + a^2 = 0$ has a real solution x.

(c) There is a unique natural number with exactly one positive divisor.

The unique existential quantifier will play a large role when we study functions in Section 3.1.

Quantifier alternation

Compare the following two statements:

(i) For every door, there is a key that can unlock it.

(ii) There is a key that can unlock every door.

Letting the variables x and y refer to doors and keys, respectively, and letting $p(x,y)$ be the statement 'door x can be unlocked by key y', we can formulate these statements as:

(i) $\forall x, \exists y,\ p(x,y)$

(ii) $\exists y, \forall x,\ p(x,y)$

This is a typical 'real-world' example of what is known as *quantifier alternation*—the two statements differ only by the order of the front-loaded quantifiers, and yet they say very different things. Statement (i) requires every door to be unlockable, but the keys might be different for different doors; statement (ii), however, implies the existence of some kind of 'master key' that can unlock all the doors.

Here's another example with a more mathematical nature:

Exercise 1.2.32

Let $p(x,y)$ be the statement '$x+y$ is even'.

- Prove that $\forall x \in \mathbb{Z}, \exists y \in \mathbb{Z},\ p(x,y)$ is true.
- Prove that $\exists y \in \mathbb{Z}, \forall x \in \mathbb{Z},\ p(x,y)$ is false.

In both of the foregoing examples, you might have noticed that the '$\forall\exists$' statement says something *weaker* than the '$\exists\forall$' statement—in some sense, it is easier to make a $\forall\exists$ statement true than it is to make an $\exists\forall$ statement true.

This idea is formalised in Theorem 1.2.33 below, which despite its abstract nature, has an extremely simple proof.

Theorem 1.2.33

Let $p(x,y)$ be a logical formula with free variables $x \in X$ and $y \in Y$. Then

$$\exists y \in Y, \forall x \in X,\ p(x,y) \Rightarrow \forall x \in X, \exists y \in Y,\ p(x,y)$$

Proof

Suppose $\exists y \in Y, \forall x \in X,\ p(x,y)$ is true. We need to prove $\forall x \in X, \exists y \in Y,\ p(x,y)$, so fix $a \in X$—our goal is now to prove $\exists y \in Y,\ p(a,y)$.

Using our assumption $\exists y \in Y, \forall x \in X,\ p(x,y)$, we may choose $b \in Y$ such that $\forall x,\ p(x,b)$ is true. But then $p(a,b)$ is true, so we have proved $\exists y \in Y,\ p(a,y)$, as required. □

Statements of the form $\exists y \in Y, \forall x \in X,\ p(x,y)$ imply some kind of *uniformity*: a value of y making $\forall x \in X,\ p(x,y)$ true can be thought of as a 'one size fits all' solution to the problem of proving $p(x,y)$ for a given $x \in X$. Later in your studies, it is likely that you will encounter the word 'uniform' many times—it is precisely this notion of quantifier alternation that the word 'uniform' refers to.

Section 1.3

Logical equivalence

We motivate the content of this section with an example.

Example 1.3.1

Consider the following two logical formulae, where P denotes the set of all prime numbers.

(1) $\forall n \in P, (n > 2 \Rightarrow [\exists k \in \mathbb{Z}, n = 2k+1])$;

(2) $\neg\exists n \in P, (n > 2 \wedge [\exists k \in \mathbb{Z}, n = 2k])$.

The logical formula (1) translates to 'every prime number greater than two is odd', and the logical formula (2) translates to 'there does not exist an even prime number greater than two'. These statements are evidently *equivalent*—they mean the same thing—but they suggest different proof strategies:

(1) Fix a prime number n, assume that $n > 2$, and then prove that $n = 2k+1$ for some $k \in \mathbb{Z}$.

(2) Assume that there is some prime number n such that $n > 2$ and $n = 2k$ for some $k \in \mathbb{Z}$, and derive a contradiction.

While statement (1) more directly translates the plain English statement 'every prime number greater than two is odd', it is the proof strategy suggested by (2) that is easier to use. Indeed, if n is a prime number such that $n > 2$ and $n = 2k$ for some $k \in \mathbb{Z}$, then 2 is a divisor of n other than 1 and n (since $1 < 2 < n$), contradicting the assumption that n is prime.

The notion of *logical equivalence*, captures precisely the sense in which the logical formulae in (1) and (2) in Example 1.3.1 'mean the same thing'. Being able to transform a logical formula into a different (but equivalent) form allows us to identify a wider range of feasible proof strategies.

Definition 1.3.2

Let p and q be logical formulae. We say that p and q are **logically equivalent**, and write $p \equiv q$ (LaTeX code: `\equiv`), if q can be derived from p and p can be derived from q.

Logical equivalence of propositional formulae

While Definition 1.3.2 defines logical equivalence between arbitrary logical formulae, we will start by focusing our attention on logical equivalence between *propositional* formulae, like those we saw in Section 1.1.

First, let's look at a couple of examples of what proofs of logical equivalence might look like. Be warned—they're not very nice to read! But there is light at the end of the tunnel. After struggling through Examples 1.3.3 and 1.3.4 and Exercise 1.3.5, we will introduce a very quick and easy tool for proving propositional formulae are logically equivalent.

Example 1.3.3

We demonstrate that $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$, where p, q and r are propositional variables.

- First assume that $p \wedge (q \vee r)$ is true. Then p is true and $q \vee r$ is true by definition of conjunction. By definition of disjunction, either q is true or r is true.
  - ◇ If q is true, then $p \wedge q$ is true by definition of conjunction.
  - ◇ If r is true, then $p \wedge r$ is true by definition of conjunction.

  In both cases we have that $(p \wedge q) \vee (p \wedge r)$ is true by definition of disjunction.
- Now assume that $(p \wedge q) \vee (p \wedge r)$ is true. Then either $p \wedge q$ is true or $p \wedge r$ is true, by definition of disjunction.
  - ◇ If $p \wedge q$ is true, then p is true and q is true by definition of conjunction.
  - ◇ If $p \wedge r$ is true, then p is true and r is true by definition of conjunction.

  In both cases we have that p is true, and that $q \vee r$ is true by definition of disjunction. Hence $p \wedge (q \vee r)$ is true by definition of conjunction.

Since we can derive $(p \wedge q) \vee (p \wedge r)$ from $p \wedge (q \vee r)$ and vice versa, it follows that

$$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$$

as required.

Example 1.3.4

We prove that $p \Rightarrow q \equiv (\neg p) \vee q$, where p, q and r are propositional variables.

- First assume that $p \Rightarrow q$ is true. By the law of excluded middle (Axiom 1.1.42), either p is true or $\neg p$ is true—we derive $(\neg p) \vee q$ in each case.
  - ◇ If p is true, then since $p \Rightarrow q$ is true, it follows from ($\Rightarrow$E) that q is true, and so $(\neg p) \vee q$ is true by ($\vee \mathrm{I}_2$);
  - ◇ If $\neg p$ is true, then $(\neg p) \vee q$ is true by ($\vee \mathrm{I}_1$).

  In both cases, we see that $(\neg p) \vee q$ is true.
- Now assume that $(\neg p) \vee q$ is true. To prove that $p \Rightarrow q$ is true, it suffices by ($\Rightarrow$I) to assume that p is true and derive q. So assume p is true. Since $(\neg p) \vee q$ is true, we have that either $\neg p$ is true or q is true.
  - ◇ If $\neg p$ is true, then we obtain a contradiction from the assumption that p is true, and so q is true by the principle of explosion (Axiom 1.1.47).
  - ◇ If q is true… well, then q is true—there is nothing more to prove!

  In both cases we have that q is true. Hence $p \Rightarrow q$ is true.

We have derived $(\neg p) \vee q$ from $p \Rightarrow q$ and vice versa, and so the two formulae are logically equivalent.

Exercise 1.3.5

Let p, q and r be propositional variables. Prove that the propositional formula $(p \vee q) \Rightarrow r$ is logically equivalent to $(p \Rightarrow r) \wedge (q \Rightarrow r)$.

Working through the derivations each time we want to prove logical equivalence can become cumbersome even for small examples like Examples 1.3.3 and 1.3.4 and Exercise 1.3.5.

The following theorem reduces the problem of proving logical equivalence between *propositional* formulae to the purely algorithmic task of checking when the formulae are true and when they are false in a (relatively) small list of cases. We will streamline this process even further using *truth tables* (Definition 1.3.7).

Theorem 1.3.6

Two propositional formulae are logically equivalent if and only if their truth values are the same under any assignment of truth values to their constituent propositional variables.

Idea of proof

A formal proof of this fact is slightly beyond our reach at this point, although we will be able to prove it formally by *structural induction*, introduced in Section 12.2.

The idea of the proof is that, since propositional formulae are built up from simpler propositional formulae using logical operators, the truth value of a more complex propositional formula is determined by the truth values of its simpler subformulae. If we keep 'chasing' these subformulae, we end up with just propositional variables.

For example, the truth value of $(p \Rightarrow r) \wedge (q \Rightarrow r)$ is determined by the truth values of $p \Rightarrow r$ and $q \Rightarrow r$ according to the rules for the conjunction operator $\wedge$. In turn, the truth value of $p \Rightarrow r$ is determined by the truth values of p and r according to the implication operator $\Rightarrow$, and the truth value of $q \Rightarrow r$ is determined by the truth values of q and r according to the implication operator again. It follows that the truth value of the whole propositional formula $(p \Rightarrow r) \wedge (q \Rightarrow r)$ is determined by the truth values of p, q, r according to the rules for $\wedge$ and $\Rightarrow$.

If some assignment of truth values to propositional variables makes one propositional formula true but another false, then it must be impossible to derive one from the other—otherwise we'd obtain a contradiction. Hence both propositional formulae must have the same truth values no matter what assignment of truth values is given to their constituent propositional variables.

We now develop a systematic way of checking the truth values of a propositional formula under each assignment of truth values to its constituent propositional variables.

Definition 1.3.7

The **truth table** of a propositional formula is the table with one row for each possible assignment of truth values to its constituent propositional variables, and one column for each subformula (including the propositional variables and the propositional formula itself). The entries of the truth table are the truth values of the subformulae.

Example 1.3.8

The following are the truth tables for $\neg p$, $p \wedge q$, $p \vee q$ and $p \Rightarrow q$.

p	$\neg p$
✓	×
×	✓

p	q	$p \wedge q$
✓	✓	✓
✓	×	×
×	✓	×
×	×	×

p	q	$p \vee q$
✓	✓	✓
✓	×	✓
×	✓	✓
×	×	×

p	q	$p \Rightarrow q$
✓	✓	✓
✓	×	×
×	✓	✓
×	×	✓

In Example 1.3.8 we have used the symbol ✓ (LaTeX code: `\checkmark`) to mean 'true' and × (LaTeX code: `\times`) to mean 'false'. Some authors adopt other conventions, such as T, F or $\top, \bot$ (LaTeX code: `\top,\bot`) or $1, 0$ or $0, 1$—the possibilites are endless!

Exercise 1.3.9

Use the definitions of $\wedge$, $\vee$ and $\Rightarrow$ to justify the truth tables in Example 1.3.8.

The next example shows how the truth tables for the individual logical operators (as in Example 1.3.8) may be combined to form a truth table for a more complicated propositional formula that involves three propositional variables.

Example 1.3.10

The following is the truth table for $(p \wedge q) \vee (p \wedge r)$.

p	q	r	$p \wedge q$	$p \wedge r$	$(p \wedge q) \vee (p \wedge r)$
✓	✓	✓	✓	✓	✓
✓	✓	×	✓	×	✓
✓	×	✓	×	✓	✓
✓	×	×	×	×	×
×	✓	✓	×	×	×
×	✓	×	×	×	×
×	×	✓	×	×	×
×	×	×	×	×	×
propositional variables			intermediate subformulae		main formula

Some comments about the construction of this truth table are pertinent:

- The propositional variables appear first. Since there are three of them, there are $2^3 = 8$ rows. The column for p contains four ✓s followed by four ×s; the column for q contains two ✓s, two ×s, and then repeats; and the column for r contains one ✓, one ×, and then repeats.
- The next group of columns are the next-most complicated subformulae. Each is constructed by looking at the relevant columns further to the left and comparing with the truth table for conjunction.
- The final column is the main formula itself, which again is constructed by looking at the relevant columns further to the left and comparing with the truth table for disjunction.

Our choices of where to put the vertical bars and what order to put the rows in were not the only choices that could have been made, but when constructing truth tables for more complex logical formulae, it is useful to develop a system and stick to it.

Returning to Theorem 1.3.6, we obtain the following strategy for proving that two propositional formulae are logically equivalent.

❖ **Strategy 1.3.11** · *Logical equivalence using truth tables*
In order to prove that propositional formulae are logically equivalent, it suffices to show that they have identical columns in a truth table.

✎ **Example 1.3.12**
In Example 1.3.3 we proved that $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$. We prove this again using truth tables. First we construct the truth table for $p \wedge (q \vee r)$:

p	q	r	$q \vee r$	$p \wedge (q \vee r)$
✓	✓	✓	✓	✓
✓	✓	×	✓	✓
✓	×	✓	✓	✓
✓	×	×	×	×
×	✓	✓	✓	×
×	✓	×	✓	×
×	×	✓	✓	×
×	×	×	×	×

Note that the column for $p \wedge (q \vee r)$ is identical to that of $(p \wedge q) \vee (p \wedge r)$ in Example 1.3.10. Hence the two formulae are logically equivalent.

To avoid having to write out two truth tables, it can be helpful to combine them into one. For example, the following truth table exhibits that $p \wedge (q \vee r)$ is logically equivalent to $(p \wedge q) \vee (p \wedge r)$:

p	q	r	$q \vee r$	$p \wedge (q \vee r)$	$p \wedge q$	$p \wedge r$	$(p \wedge q) \vee (p \wedge r)$
✓	✓	✓	✓	✓	✓	✓	✓
✓	✓	×	✓	✓	✓	×	✓
✓	×	✓	✓	✓	×	✓	✓
✓	×	×	×	×	×	×	×
×	✓	✓	✓	×	×	×	×
×	✓	×	✓	×	×	×	×
×	×	✓	✓	×	×	×	×
×	×	×	×	×	×	×	×

In the following exercises, we use truth tables to repeat the proofs of logical equivalence from Example 1.3.4 and Exercise 1.3.5.

Exercise 1.3.13
Use a truth table to prove that $p \Rightarrow q \equiv (\neg p) \vee q$.

Exercise 1.3.14
Let p, q and r be propositional variables. Use a truth table to prove that the propositional formula $(p \vee q) \Rightarrow r$ is logically equivalent to $(p \Rightarrow r) \wedge (q \Rightarrow r)$.

Some proof strategies

We are now in good shape to use logical equivalence to derive some more sophisticated proof strategies.

Theorem 1.3.15 · *Law of double negation*
Let p be a propositional variable. Then $p \equiv \neg\neg p$.

Proof
The proof is almost trivialised using truth tables. Indeed, consider the following truth table.

p	$\neg p$	$\neg\neg p$
✓	×	✓
×	✓	×

The columns for p and $\neg\neg p$ are identical, and so $p \equiv \neg\neg p$. □

The law of double negation is important because it suggests a second way that we can prove statements by contradiction. Indeed, it says that proving a proposition p is equivalent to proving $\neg\neg p$, which amounts to assuming $\neg p$ and deriving a contradiction.

Strategy 1.3.16 · *Proof by contradiction (indirect version)*
In order to prove a proposition p is true, it suffices to assume that p is false and derive a contradiction.

At first sight, Strategy 1.3.16 looks very similar to Strategy 1.1.36, which we also termed *proof by contradiction*. But there is an important difference between the two:

- Strategy 1.1.36 says that to prove that a proposition is *false*, it suffices to assume that it is *true* and derive a contradiction;
- Strategy 1.3.16 says that to prove that a proposition is *true*, it suffices to assume that it is *false* and derive a contradiction.

The former is a *direct* proof technique, since it arises directly from the definition of the negation operator; the latter is an *indirect* proof technique, since it arises from a logical equivalence, namely the law of double negation.

Example 1.3.17

We prove that if a, b and c are non-negative real numbers satisfying $a^2+b^2=c^2$, then $a+b \geqslant c$.

Indeed, let $a,b,c \in \mathbb{R}$ with $a,b,c \geqslant 0$, and assume that $a^2+b^2=c^2$. Towards a contradiction, assume that it is not the case that $a+b \geqslant c$. Then we must have $a+b<c$. But then

$$(a+b)^2=(a+b)(a+b)<(a+b)c<c\cdot c=c^2$$

and so

$$c^2>(a+b)^2=a^2+2ab+b^2=c^2+2ab \geqslant c^2$$

This implies that $c^2>c^2$, which is a contradiction. So it must be the case that $a+b \geqslant c$, as required.

The next proof strategy we derive concerns proving implications.

Definition 1.3.18

The **contrapositive** of a proposition of the form $p \Rightarrow q$ is the proposition $\neg q \Rightarrow \neg p$.

Theorem 1.3.19 · *Law of contraposition*

Let p and q be propositional variables. Then $p \Rightarrow q \equiv (\neg q) \Rightarrow (\neg p)$.

Proof

We build the truth tables for $p \Rightarrow q$ and $(\neg q) \Rightarrow (\neg p)$.

p	q	$p \Rightarrow q$	$\neg q$	$\neg p$	$(\neg q) \Rightarrow (\neg p)$
✓	✓	✓	×	×	✓
✓	×	×	✓	×	×
×	✓	✓	×	✓	✓
×	×	✓	✓	✓	✓

The columns for $p \Rightarrow q$ and $(\neg q) \Rightarrow (\neg p)$ are identical, so they are logically equivalent. □

Theorem 1.3.19 suggests the following proof strategy.

Strategy 1.3.20 · *Proof by contraposition*

In order to prove a proposition of the form $p \Rightarrow q$, it suffices to assume that q is false and derive that p is false.

Example 1.3.21

Fix two natural numbers m and n. We will prove that if $mn>64$, then either $m>8$ or $n>8$.

By contraposition, it suffices to assume that it is *not* the case that $m>8$ or $n>8$, and derive that it is not the case that $mn>64$.

So assume that neither $m>8$ nor $n>8$. Then $m \leqslant 8$ and $n \leqslant 8$, so that $mn \leqslant 64$, as required.

Exercise 1.3.22

Use the law of contraposition to prove that $p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge ((\neg p) \Rightarrow (\neg q))$, and use the proof technique that this equivalence suggests to prove that an integer is even if and only if its square is even.

It feels good to invoke impressive-sounding results like *proof by contraposition*, but in practice, the logical equivalence between *any* two different propositional formulae suggests a new proof technique, and not all of these techniques have names. And indeed, the proof strategy in the following exercise, while useful, has no slick-sounding name—at least, not one that would be widely understood.

Exercise 1.3.23

Prove that $p \vee q \equiv (\neg p) \Rightarrow q$. Use this logical equivalence to suggest a new strategy for proving propositions of the form $p \vee q$, and use this strategy to prove that if two integers sum to an even number, then either both integers are even or both are odd.

Negation

In pure mathematics it is common to ask whether or not a certain property holds of a mathematical object. For example, in Section 9.2, we will look at convergence of sequences of real numbers: to say that a sequence $x_0, x_1, x_2, \ldots$ of real numbers *converges* (Definition 9.2.15) is to say

$$\exists a \in \mathbb{R}, \forall \varepsilon \in \mathbb{R}, (\varepsilon > 0 \Rightarrow \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

This is already a relatively complicated logical formula. But what if we wanted to prove that a sequence *does not* converge? Simply assuming the logical formula above and deriving a contradiction might work sometimes, but it is not particularly enlightening.

Our next goal is to develop a systematic method for negating complicated logical formulae. With this done, we will be able to negate the logical formula expressing 'the sequence $x_0, x_1, x_2, \ldots$ converges' as follows

$$\forall a \in \mathbb{R}, \exists \varepsilon \in \mathbb{R}, (\varepsilon > 0 \wedge \forall N \in \mathbb{N}, \exists n \in \mathbb{N}, [n \geqslant N \wedge |x_n - a| \geqslant \varepsilon])$$

Granted, this is still a complicated expression, but when broken down element by element, it provides useful information about how it may be proved.

The rules for negating conjunctions and disjunctions are instances of *de Morgan's laws*, which exhibit a kind of duality between $\wedge$ and $\vee$.

Theorem 1.3.24 · *de Morgan's laws for logical operators*

Let p and q be logical formulae. Then:

(a) $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$; and

(b) $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.

Proof of (a)
Consider the following truth table.

p	q	$p \wedge q$	$\neg(p \wedge q)$	$\neg p$	$\neg q$	$(\neg p) \vee (\neg q)$
✓	✓	✓	×	×	×	×
✓	×	×	✓	×	✓	✓
×	✓	×	✓	✓	×	✓
×	×	×	✓	✓	✓	✓

The columns for $\neg(p \wedge q)$ and $(\neg p) \vee (\neg q)$ are identical, so they are logically equivalent. □

Exercise 1.3.25
Prove Theorem 1.3.24(b) thrice: once using the definition of logical equivalence directly (like we did in Examples 1.3.3 and 1.3.4 and Exercise 1.3.5), once using a truth table, and once using part (a) together with the law of double negation.

Example 1.3.26
We often use de Morgan's laws for logical operators without thinking about it. For example, to say that 'neither 3 nor 7 is even' is equivalent to saying '3 is odd and 7 is odd'. The former statement translates to

$$\neg[(3 \text{ is even}) \vee (7 \text{ is even})]$$

while the second statement translates to

$$[\neg(3 \text{ is even})] \wedge [\neg(7 \text{ is even})]$$

Exercise 1.3.27
Prove that $\neg(p \Rightarrow q) \equiv p \wedge (\neg q)$ twice, once using a truth table, and once using Exercise 1.3.13 together with de Morgan's laws and the law of double negation.

Although conjunctions can be negated using de Morgan's laws, the following exercise offers a different way of negating conjunctions that is often more practical when being used in a proof.

Exercise 1.3.28
Prove that $\neg(p \wedge q) \equiv p \Rightarrow (\neg q)$. What strategy does this suggest for proving that a conjunction is false? How does this compare with the strategy suggested by de Morgan's laws for logical operators?

De Morgan's laws for logical operators generalise to statements about quantifiers, expressing a similar duality between $\forall$ and $\exists$ as we have between $\wedge$ and $\vee$.

Theorem 1.3.29 · *de Morgan's laws for quantifiers*
let $p(x)$ be a logical formula with free variable x ranging over a set X. Then:

(a) $\neg\forall x \in X, p(x) \equiv \exists x \in X, \neg p(x)$; and

(b) $\neg\exists x \in X,\, p(x) \equiv \forall x \in X,\, \neg p(x)$.

Proof
Unfortunately, since these logical formulae involve quantifiers, we do not have truth tables at our disposal, so we must assume each formula and derive the other.

We start by proving the equivalence in part (b), and then we derive (a) as a consequence.

- Assume $\neg\exists x \in X,\, p(x)$. To prove $\forall x \in X,\, \neg p(x)$, fix some $x \in X$. If $p(x)$ were true, then we'd have $\exists x \in X,\, p(x)$, which contradicts our main assumption; so we have $\neg p(x)$. But then $\forall x \in X,\, \neg p(x)$ is true.
- Assume $\forall x \in X,\, \neg p(x)$. For the sake of contradiction, assume $\exists x \in X,\, p(x)$ were true. Then we obtain some $a \in X$ for which $p(a)$ is true. But $\neg p(a)$ is true by the assumption that $\forall x \in X,\, \neg p(a)$, so we obtain a contradiction. Hence $\neg\exists x \in X,\, p(x)$ is true.

This proves that $\neg\exists x \in X,\, p(x) \equiv \forall x \in X,\, \neg p(x)$.

Now (a) follows from (b) using the law of double negation (Theorem 1.3.15):

$$\exists x \in X,\, \neg p(x) \equiv \neg\neg\exists x \in X,\, \neg p(x) \overset{(b)}{\equiv} \neg\forall x \in X,\, \neg\neg p(x) \equiv \neg\forall x \in X,\, p(x)$$

as required. □

The proof strategy suggested by the logical equivalence in Theorem 1.3.29(b) is so important that it has its own name.

Strategy 1.3.30 · *Proof by counterexample*
To prove that a proposition of the form $\forall x \in X,\, p(x)$ is false, it suffices to find a single element $a \in X$ such that $p(a)$ is false. The element a is called a **counterexample** to the proposition $\forall x \in X,\, p(x)$.

Example 1.3.31
We prove by counterexample that not every integer is divisible by a prime number. Indeed, let $x = 1$. The only integral factors of 1 are 1 and -1, neither of which are prime, so that 1 is not divisible by any primes.

Exercise 1.3.32
Prove by counterexample that not every rational number can be expressed as $\frac{a}{b}$ where $a \in \mathbb{Z}$ is even and $b \in \mathbb{Z}$ is odd.

We have now seen how to negate the logical operators $\neg$, $\wedge$, $\vee$ and $\Rightarrow$, as well as the quantifiers $\forall$ and $\exists$.

Definition 1.3.33

A logical formula is **maximally negated** if the only instances of the negation operator $\neg$ appear immediately before a predicate (or other proposition not involving logical operators or quantifiers).

Example 1.3.34

The following propositional formula is maximally negated:

$$[p \wedge (q \Rightarrow (\neg r))] \Leftrightarrow (s \wedge (\neg t))$$

Indeed, all instances of $\neg$ appear immediately before propositional variables.

However the following propositional formula is *not* maximally negated:

$$(\neg\neg q) \Rightarrow q$$

Here the subformula $\neg\neg q$ contains a negation operator immediately before another negation operator $(\neg\neg q)$. However by the law of double negation, this is equivalent to $q \Rightarrow q$, which is maximally negated trivially since there are no negation operators to speak of.

Exercise 1.3.35

Determine which of the following logical formulae are maximally negated.

(a) $\forall x \in X, (\neg p(x)) \Rightarrow \forall y \in X, \neg(r(x,y) \wedge s(x,y))$;

(b) $\forall x \in X, (\neg p(x)) \Rightarrow \forall y \in X, (\neg r(x,y)) \vee (\neg s(x,y))$;

(c) $\forall x \in \mathbb{R}, [x > 1 \Rightarrow (\exists y \in \mathbb{R}, [x < y \wedge \neg(x^2 \leqslant y)])]$;

(d) $\neg\exists x \in \mathbb{R}, [x > 1 \wedge (\forall y \in \mathbb{R}, [x < y \Rightarrow x^2 \leqslant y])]$.

The following theorem allows us to replace logical formulae by maximally negated ones, which in turn suggests proof strategies that we can use for proving that complicated-looking propositions are *false*.

Theorem 1.3.36

Every logical formula (built using only the logical operators and quantifiers we have seen so far) is logically equivalent to a maximally negated logical formula.

Idea of proof

Much like Theorem 1.3.6, a precise proof of Theorem 1.3.36 requires some form of induction argument, so instead we will give an idea of the proof.

Every logical formula we have seen so far is built from predicates using the logical operators $\wedge, \vee, \Rightarrow$ and $\neg$ and the quantifiers $\forall$ and $\exists$—indeed, the logical operator $\Leftrightarrow$ was defined in terms of $\wedge$ and $\Rightarrow$, and the quantifier $\exists$ was defined in terms of the quantifiers $\forall$ and $\exists$ and the logical operators $\wedge$ and $\Rightarrow$.

But the results in this section allow us to push negations 'inside' each of these logical operators and quantifiers, as summarised in the following table.

Negation outside		Negation inside	Proof
$\neg(p \wedge q)$	$\equiv$	$(\neg p) \vee (\neg q)$	Theorem 1.3.24(a)
$\neg(p \vee q)$	$\equiv$	$(\neg p) \wedge (\neg q)$	Theorem 1.3.24(b)
$\neg(p \Rightarrow q)$	$\equiv$	$p \wedge (\neg q)$	Exercise 1.3.27
$\neg(\neg p)$	$\equiv$	p	Theorem 1.3.15
$\neg\forall x \in X,\, p(x)$	$\equiv$	$\exists x \in X,\, \neg p(x)$	Theorem 1.3.29(a)
$\neg\exists x \in X,\, p(x)$	$\equiv$	$\forall x \in X,\, \neg p(x)$	Theorem 1.3.29(b)

Repeatedly applying these rules to a logical formula eventually yields a logically equivalent, maximally negated logical formula.

Example 1.3.37

Recall the logical formula from page 69 expressing the assertion that a sequence $x_0, x_1, x_2, \ldots$ of real numbers converges:

$$\exists a \in \mathbb{R},\, \forall \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \Rightarrow \exists N \in \mathbb{N},\, \forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

We will maximally negate this to obtain a logical formula expressing the assertion that the sequence does not converge.

Let's start at the beginning. The negation of the formula we started with is:

$$\neg\exists a \in \mathbb{R},\, \forall \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \Rightarrow \exists N \in \mathbb{N},\, \forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

The key to maximally negating a logical formula is to ignore information that is not immediately relevant. Here, the expression that we are negating takes the form $\neg\exists a \in \mathbb{R}$, (stuff). It doesn't matter what the 'stuff' is just yet; all that matters is that we are negating an existentially quantified statement, and so de Morgan's laws for quantifiers tells us that this is logically equivalent to $\forall a \in \mathbb{R}, \neg$(stuff). We apply this rule and just re-write the 'stuff', to obtain:

$$\forall a \in \mathbb{R},\, \neg\forall \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \Rightarrow \exists N \in \mathbb{N},\, \forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

Now we are negating a universally quantified statement, $\neg\forall \varepsilon \in \mathbb{R}$, (stuff) which, by de Morgan's laws for quantifiers, is equivalent to $\exists \varepsilon \in \mathbb{R}, \neg$(stuff):

$$\forall a \in \mathbb{R},\, \exists \varepsilon \in \mathbb{R},\, \neg(\varepsilon > 0 \Rightarrow \exists N \in \mathbb{N},\, \forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

At this point, the statement being negated is of the form (stuff) $\Rightarrow$ (junk), which by Exercise 1.3.27 negates to (stuff) $\wedge \neg$(junk). Here, 'stuff' is $\varepsilon > 0$ and 'junk' is $\exists N \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon]$. So performing this negation yields:

$$\forall a \in \mathbb{R},\, \exists \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \wedge \neg\exists N \in \mathbb{N},\, \forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

Now we are negating an existentially quantified formula again, so using de Morgan's laws for quantifiers gives:

$$\forall a \in \mathbb{R},\, \exists \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \wedge \forall N \in \mathbb{N},\, \neg\forall n \in \mathbb{N},\, [n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

The formula being negated here is universally quantified, so using de Morgan's laws for quantifiers *again* gives:

$$\forall a \in \mathbb{R},\, \exists \varepsilon \in \mathbb{R},\, (\varepsilon > 0 \wedge \forall N \in \mathbb{N},\, \exists n \in \mathbb{N},\, \neg[n \geqslant N \Rightarrow |x_n - a| < \varepsilon])$$

We're almost there! The statement being negated here is an implication, so applying the rule $\neg(p \Rightarrow q) \equiv p \wedge (\neg q)$ again yields:

$$\forall a \in \mathbb{R}, \exists \varepsilon \in \mathbb{R}, (\varepsilon > 0 \wedge \forall N \in \mathbb{N}, \exists n \in \mathbb{N}, [n \geqslant N \wedge \neg(|x_n - a| < \varepsilon)])$$

At this point, strictly speaking, the formula is maximally negated, since the statement being negated does not involve any other logical operators or quantifiers. However, since $\neg(|x_n - a| < \varepsilon)$ is equivalent to $|x_n - a| \geqslant \varepsilon$, we can go one step further to obtain:

$$\forall a \in \mathbb{R}, \exists \varepsilon \in \mathbb{R}, (\varepsilon > 0 \wedge \forall N \in \mathbb{N}, \exists n \in \mathbb{N}, [n \geqslant N \wedge |x_n - a| \geqslant \varepsilon])$$

This is as negated as we could ever dream of, and so we stop here.

✎ **Exercise 1.3.38**

Find a maximally negated propositional formula that is logically equivalent to $\neg(p \Leftrightarrow q)$.

✎ **Exercise 1.3.39**

Maximally negate the following logical formula, then prove that it is true or prove that it is false.

$$\exists x \in \mathbb{R}, [x > 1 \wedge (\forall y \in \mathbb{R}, [x < y \Rightarrow x^2 \leqslant y])]$$

Tautologies

The final concept that we introduce in this chapter is that of a *tautology*, which can be thought of as the opposite of a contradiction. The word 'tautology' has other implications when used colloquially, but in the context of symbolic logic it has a precise definition.

✦ **Definition 1.3.40**

A **tautology** is a proposition or logical formula that is true, no matter how truth values are assigned to its component propositional variables and predicates.

The reason we are interested in tautologies is that tautologies can be used as assumptions at any point in a proof, for any reason.

❖ **Strategy 1.3.41** · *Assuming tautologies*

Let p be a proposition. Any tautology may be assumed in any proof of p.

✐ **Example 1.3.42**

The law of excluded middle (Axiom 1.1.42) says precisely that $p \vee (\neg p)$ is a tautology. This means that when proving any result, we may split into cases based on whether a proposition is true or false, just as we did in Proposition 1.1.44.

Example 1.3.43

The formula $p \Rightarrow (q \Rightarrow p)$ is a tautology.

A direct proof of this fact is as follows. In order to prove $p \Rightarrow (q \Rightarrow p)$ is true, it suffices to assume p and derive $q \Rightarrow p$. So assume p. Now in order to prove $q \Rightarrow p$, it suffices to assume q and derive p. So assume q. But we're already assuming that p is true! So $q \Rightarrow p$ is true, and hence $p \Rightarrow (q \Rightarrow p)$ is true.

A proof using truth tables is as follows:

p	q	$q \Rightarrow p$	$p \Rightarrow (q \Rightarrow p)$
✓	✓	✓	✓
✓	×	✓	✓
×	✓	×	✓
×	×	✓	✓

We see that $p \Rightarrow (q \Rightarrow p)$ is true regardless of the truth values of p and q.

Exercise 1.3.44

Prove that each of the following is a tautology:

(a) $[(p \Rightarrow q) \wedge (q \Rightarrow r)] \Rightarrow (p \Rightarrow r)$;

(b) $[p \Rightarrow (q \Rightarrow r)] \Rightarrow [(p \Rightarrow q) \Rightarrow (p \Rightarrow r)]$;

(c) $\exists y \in Y, \forall x \in X, p(x,y) \Rightarrow \forall x \in X, \exists y \in Y, p(x,y)$;

(d) $[\neg(p \wedge q)] \Leftrightarrow [(\neg p) \vee (\neg q)]$;

(e) $(\neg \forall x \in X, p(x)) \Leftrightarrow (\exists x \in X, \neg p(x))$.

For each, try to interpret what it means, and how it might be useful in a proof.

You may have noticed parallels between de Morgan's laws for logical operators and quantifiers, and parts (d) and (e) of Exercise 1.3.44, respectively. They almost seem to say the same thing, except that in Exercise 1.3.44 we used '$\Leftrightarrow$' and in Theorems 1.3.24 and 1.3.29 we used '$\equiv$'. There is an important difference, though: if p and q are logical formulae, then $p \Rightarrow q$ is itself a logical formula, which we may study as a mathematical object in its own right. However, $p \equiv q$ is not a logical formula: it is an assertion *about* logical formulae, namely that the logical formulae p and q are equivalent.

There is, nonetheless, a close relationship between $\Leftrightarrow$ and $\equiv$—this relationship is summarised in the following theorem.

Theorem 1.3.45

Let p and q be logical formulae.

(a) q can be derived from p if and only if $p \Rightarrow q$ is a tautology;

(b) $p \equiv q$ if and only if $p \Leftrightarrow q$ is a tautology.

Proof

For (a), note that a derivation of q from p is sufficient to establish the truth of $p \Rightarrow q$ by the introduction rule for implication ($\Rightarrow$I), and so if q can be derived from p, then $p \Rightarrow q$ is a tautology. Conversely, if $p \Rightarrow q$ is a tautology, then q can be derived from p using the elimination rule for implication ($\Rightarrow$E) together with the (tautological) assumption that $p \Rightarrow q$ is true.

Now (b) follows from (a), since logical equivalence is defined in terms of derivation in each direction, and $\Leftrightarrow$ is simply the conjunction of two implications. □

Section 1.E

Chapter 1 exercises

1.E.1. For fixed $n \in \mathbb{N}$, let p represent the proposition 'n is even', let q represent the proposition 'n is prime' and let r represent the proposition '$n = 2$'. For each of the following propositional formulae, translate it into plain English and determine whether it is true for all $n \in \mathbb{N}$, true for some values of n and false for some values of n, or false for all $n \in \mathbb{N}$.

(a) $(p \wedge q) \Rightarrow r$

(b) $q \wedge (\neg r) \Rightarrow (\neg p)$

(c) $((\neg p) \vee (\neg q)) \vee (\neg r)$

(d) $(p \wedge q) \wedge (\neg r)$

1.E.2. For each of the following plain English statements, translate it into a symbolic propositional formula. The propositional variables in your formulae should represent the simplest propositions that they can.

(a) Guinea pigs are quiet, but they're loud when they're hungry.

(b) It doesn't matter that 2 is even, it's still a prime number.

(c) $\sqrt{2}$ can't be an integer because it is an irrational number.

1.E.3. Let p and q be propositions, and assume that $p \Rightarrow (\neg q)$ is true and that $(\neg q) \Rightarrow p$ is false. Which of the following are true, and which are false?

(a) q being false is necessary for p to be true.

(b) q being false is sufficient for p to be true.

(c) p being true is necessary for q to be false.

(d) p being true is sufficient for p to be false.

In Questions 1.E.4 to 1.E.7, use the definitions of the logical operators in Section 1.1 to describe what steps should be followed in order to prove the propositional formula in the question; the letters p, q, r and s are propositional variables.

1.E.4. $(p \wedge q) \Rightarrow (\neg r)$

1.E.5. $(p \vee q) \Rightarrow (r \Rightarrow s)$

1.E.6. $(p \Rightarrow q) \Leftrightarrow (\neg p \Rightarrow \neg q)$

1.E.7. $(p \wedge (\neg q)) \vee (q \wedge (\neg p))$

1.E.8. Find a statement in plain English, involving no variables at all, that is equivalent to the logical formula $\forall a \in \mathbb{Q}, \forall b \in \mathbb{Q}, (a < b \Rightarrow \exists c \in \mathbb{R}, [a < c < b \ \wedge \ \neg(c \in \mathbb{Q})])$. Then prove this statement, using the structure of the logical formula as a guide.

1.E.9. Find a purely symbolic logical formula that is equivalent to the following statement, and then prove it: "*No matter which integer you may choose, there will be an integer greater than it.*"

1.E.10. Let X be a set and let $p(x)$ be a predicate. Find a logical formula representing the statement 'there are exactly two elements $x \in X$ such that $p(x)$ is true'. Use the structure of this logical formula to describe how a proof should be structured, and use this structure to prove that there are exactly two real numbers x such that $x^2 = 1$.

1.E.11. Prove that

$$p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge ((\neg p) \Rightarrow (\neg q))$$

How might this logical equivalence help you to prove statements of the form 'p if and only if q'?

1.E.12. Prove using truth tables that $p \Rightarrow q \not\equiv q \Rightarrow p$. Give an example of propositions p and q such that $p \Rightarrow q$ is true but $q \Rightarrow p$ is false.

In Questions 1.E.13 to 1.E.16, find a logical formula whose column in a truth table is as shown.

1.E.13.

p	q	
✓	✓	×
✓	×	✓
×	✓	✓
×	×	×

1.E.14.

p	q	
✓	✓	✓
✓	×	×
×	✓	✓
×	×	×

1.E.15.

p	q	r	
✓	✓	✓	✓
✓	✓	×	✓
✓	×	✓	×
✓	×	×	×
×	✓	✓	×
×	✓	×	×
×	×	✓	✓
×	×	×	✓

1.E.16.

p	q	r	
✓	✓	✓	✓
✓	✓	×	×
✓	×	✓	×
✓	×	×	×
×	✓	✓	✓
×	✓	×	×
×	×	✓	✓
×	×	×	×

1.E.17. A new logical operator $\uparrow$ is defined by the following rules:

(i) If a contradiction can be derived from the assumption that p is true, then $p \uparrow q$ is true;

(ii) If a contradiction can be derived from the assumption that q is true, then $p \uparrow q$ is true;

(iii) If r is any proposition, and if $p \uparrow q$, p and q are all true, then r is true.

This question explores this curious new logical operator.

(a) Prove that $p \uparrow p \equiv \neg p$, and deduce that $((p \uparrow p) \uparrow (p \uparrow p)) \equiv p$.

(b) Prove that $p \vee q \equiv (p \uparrow p) \uparrow (q \uparrow q)$ and $p \wedge q \equiv (p \uparrow q) \uparrow (p \uparrow q)$.

(c) Find a propositional formula using only the logical operator $\uparrow$ that is equivalent to $p \Rightarrow q$.

1.E.18. Let X be $\mathbb{Z}$ or $\mathbb{Q}$, and define a logical formula p by:

$$\forall x \in X, \exists y \in X, (x < y \wedge [\forall z \in X, \neg(x < z \wedge z < y)])$$

Write out $\neg p$ as a maximally negated logical formula. Prove that p is true when $X = \mathbb{Z}$, and p is false when $X = \mathbb{Q}$.

1.E.19. Use Definition 1.2.26 to write out a maximally negated logical formula that is equivalent to $\neg \exists ! x \in X, p(x)$. Describe the strategy that this equivalence suggests for proving that there is not a unique $x \in X$ such that $p(x)$ is true, and use this strategy to prove that, for all $a \in \mathbb{R}$, if $a \neq -1$ then there is not a unique $x \in \mathbb{R}$ such that $x^4 - 2ax^2 + a^2 - 1 = 0$.

1.E.20. Define a new quantifier $\forall !$ such that de Morgan's laws for quantifiers (Theorem 1.3.29) hold with $\forall$ and $\exists$ replaced by $\forall !$ and $\exists !$, respectively.

True–False questions

In Questions 1.E.21 to 1.E.30, determine (with proof) whether the statement is true or false.

1.E.21. Every implication is logically equivalent to its contrapositive.

1.E.22. Every implication is logically equivalent to its converse.

1.E.23. Every propositional formula whose only logical operators are conjunctions and negations is logically equivalent to a propositional formula whose only logical operators are disjunctions and negations.

1.E.24. Every propositional formula whose only logical operators are conjunctions is logically equivalent to a propositional formula whose only logical operators are disjunctions.

1.E.25. The formulae $p \wedge (q \vee r)$ and $(p \wedge q) \vee r$ are logically equivalent.

1.E.26. The formulae $p \vee (q \vee r)$ and $(p \vee q) \wedge (p \vee r)$ are logically equivalent.

1.E.27. The logical formulae $\neg \forall x \in X, \forall y \in Y, p(x,y)$ and $\exists x \in X, \forall y \in Y, p(x,y)$ are logically equivalent.

1.E.28. $\neg \forall x \geqslant 0, \exists y \in \mathbb{R}, y^2 = x$ is logically equivalent to $\forall x < 0, \exists y \notin \mathbb{R}, y^2 \neq x$.

1.E.29. $\neg \forall x \geqslant 0, \exists y \in \mathbb{R}, y^2 = x$ is logically equivalent to $\exists x \geqslant 0, \forall y \in \mathbb{R}, y^2 \neq x$.

1.E.30. $\neg \forall x \geqslant 0, \exists y \in \mathbb{R}, y^2 = x$ is logically equivalent to $\exists x < 0, \forall y \in \mathbb{R}, y^2 = x$.

Always–Sometimes–Never questions

In Questions 1.E.31 to 1.E.35, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

1.E.31. Let p and q be propositions and assume that p is true. Then $p \Rightarrow q$ is true.

1.E.32. Let p and q be propositions and assume that p is false. Then $p \Rightarrow q$ is true.

1.E.33. Let X and Y be sets and let $p(x,y)$ be a predicate with free variables $x \in X$ and $y \in Y$. Then the logical formulae $\forall x \in X$, $\forall y \in Y$, $p(x,y)$ and $\forall y \in Y$, $\forall x \in X$, $p(x,y)$ are logically equivalent.

1.E.34. Let X and Y be sets and let $p(x,y)$ be a predicate with free variables $x \in X$ and $y \in Y$. Then the logical formulae $\forall x \in X$, $\exists y \in Y$, $p(x,y)$ and $\exists y \in Y$, $\forall x \in X$, $p(x,y)$ are logically equivalent.

1.E.35. Let X and Y be sets and let $p(x,y)$ be a predicate with free variables $x \in X$ and $y \in Y$. Then the logical formulae $\forall x \in X$, $\forall y \in Y$, $p(x,y)$ and $\exists x \in X$, $\exists y \in Y$, $\neg p(x,y)$ are logically equivalent.

Chapter 2

Sets

We saw an informal definition of a *set* in Chapter 0, but so far the only sets that we have seen are the number sets ($\mathbb{N}$, $\mathbb{Z}$ and so on).

In Section 2.1, we will study sets in the abstract—in particular, the sets that we study are arbitrary collections of mathematical objects, not just numbers. This is essential for further study in pure mathematics, since most (if not all) areas of pure mathematics concern certain kinds of sets!

In Section 2.2, we will then define some operations that allow us to form new sets out of old sets, and prove some identities of an algebraic nature that are closely related to the rules governing logical operators and quantifiers that we saw in Chapter 1.

Section 2.1

Sets

We will begin this section on a note that is somewhat harrowing, but which hopefully clarifies why there is a need to make our foundational assumptions more precise than we did in Chapter 0.

✎ **Exercise 2.1.1** · *Russell's paradox*

Let $S = \{X \mid X \text{ is a set}\}$ and let $R = \{X \in S \mid X \notin X\}$; thus, S is the set of all sets, and R is the set of all sets that are not elements of themselves.

Determine the truth value of the proposition $R \in R$.

If you completed Exercise 2.1.1 as intended, then you will have noticed that something is wrong: there is a proposition whose truth implies its own falsity, and vice versa.

It goes without saying that this kind of thing should not be allowed to happen. But where did we go wrong? The problem, it turns out, is that we need to be a little pickier about what a 'set' is: we can't just write down an expression in set-builder notation and expect it to be a set, like when we defined $S = \{X \mid X \text{ is a set}\}$.

This problem can be overcome by revising our definition of a set.

✦ **Definition 2.1.2** · *updated from Definition 0.18*

A **set** is a collection of **elements** from a specified **universe of discourse**. The collection of everything in the universe of discourse is called the **universal set**, denoted by $\mathscr{U}$ (LaTeX code: `\mathcal{U}`).

The universal set $\mathscr{U}$ is understood to contain as elements all mathematical objects we are interested in: all the different types of numbers are elements of $\mathscr{U}$, as are the number sets themselves, and everything else we will go on to define in this textbook (power sets, functions, relations, etc.). However, importantly, $\mathscr{U}$ itself is *not* an element of $\mathscr{U}$.

We will avoid referring explicitly to the universal set $\mathscr{U}$ whenever possible, but it will always be there in the background: whenever we refer to an arbitrary mathematical object x without specifying what set it comes from (e.g. '$x = x$ is true for all x'), it will be understood that $x \in \mathscr{U}$. Thus, for example:

- $\forall x,\, p(x)$ really means $\forall x \in \mathscr{U},\, p(x)$ — note that $\forall x \in X,\, p(x) \equiv \forall x,\, (x \in X \Rightarrow p(x))$;
- $\exists x,\, p(x)$ really means $\exists x \in \mathscr{U},\, p(x)$ — note that $\exists x \in X,\, p(x) \equiv \forall x,\, (x \in X \wedge p(x))$.

✎ **Exercise 2.1.3**

Think about why the characterisations of '$\forall x \in X$' in terms of '$\forall x$' and of '$\exists x \in X$' in terms of '$\exists x$' involve different logical operators ($\Rightarrow$ for $\forall$, and $\wedge$ for $\exists$), and verify de Morgan's laws for quantifiers (Theorem 1.3.29) for the expanded forms.

The new definition of a set allows us to be slightly more precise about what we mean when we use set-builder notation to specify a set.

Definition 2.1.4 · *Set-builder notation*

Let $p(x)$ be a statement involving a variable x. The **set-builder notation** $\{x \mid p(x)\}$ denotes the set of all mathematical objects x making the statement $p(x)$ true; thus, for all a, we have

$$a \in \{x \mid p(x)\} \quad \Leftrightarrow \quad a \in \mathscr{U} \wedge p(a)$$

More generally, for all sets X and all expressions $\text{expr}(x)$ involving x as a variable, we define:

- $\{x \in X \mid p(x)\} = \{x \mid x \in X \wedge p(x)\}$; and
- $\{\text{expr}(x) \mid p(x)\} = \{y \mid \exists x, y = \text{expr}(x) \wedge p(x)\}$.

In general, we cannot expect an arbitrary set defined in set-builder notation to be an element of $\mathscr{U}$—for example, $\{x \mid x = x\} = \mathscr{U} \notin \mathscr{U}$. This observation is important for the following exercise.

Exercise 2.1.5 · *Russell's paradox has been resolved!*

Repeat Exercise 2.1.1 with the updated definitions of sets and set-builder notation.

With our foundational crisis resolved, we will almost completely cease to refer to $\mathscr{U}$ from this point onwards. Readers who want to delve deeper should read Section B.1. Readers who want to remain blissfully ignorant can take for granted that all of the mathematical objects we will discuss from this point forward are elements of $\mathscr{U}$, and need not think of $\mathscr{U}$ again for the foreseeable future.

Subsets

We saw in Exercise 0.21(h) that the proposition $\mathbb{N} \in \mathbb{Z}$ is false: the elements of $\mathbb{Z}$ are integers, which are numbers, whereas $\mathbb{N}$ is a set, so $\mathbb{N}$ cannot be an element of $\mathbb{Z}$.

This isn't very satisfying, though: while $\mathbb{Z}$ doesn't contain $\mathbb{N}$ as an *element*, there has to be some sense in which $\mathbb{Z}$ contains $\mathbb{N}$, since every natural number is an integer.

This is what is captured by the notion of a *subset*.

Definition 2.1.6

Let X be a set. A **subset** of X is a set U such that

$$\forall a \in U, a \in X \quad \text{or equivalently:} \quad \forall a, (a \in U \Rightarrow a \in X)$$

We write $U \subseteq X$ (LaTeX code: `\subseteq`) for the assertion that U is a subset of X.

Additionally, the notation $U \nsubseteq X$ (LaTeX code: `\nsubseteq`) means that U is not a subset of X, and the notation $U \subsetneq X$ (LaTeX code: `\subsetneq`) means that U is a **proper subset** of X, that is a subset of X that is not equal to X.

Strategy 2.1.7 · *Proving a subset containment*

In order to prove that a set U is a subset of a set X, it suffices to take an arbitrary element $a \in U$ and prove that $a \in X$.

Aside

The subset symbol '$\subseteq$' is suggestively similar to the inequality symbol '$\leqslant$', so you may be wondering why we don't write '$\subset$'(LaTeX code: `\subset`) (mirroring '$<$') to denote proper subsets. The reason is that the symbol $\subset$ is ambiguous without specifying what it means: some authors use $\subset$ to simply mean *subset* (so that $A \subset B$ means $A \subseteq B$), and others use it to mean *proper subset* (so that $A \subset B$ means $A \subsetneq B$). The symbols $\subseteq$ and $\subsetneq$ are always unambiguous, so they are what we will use in this textbook.

Example 2.1.8

Every set is a subset of itself—that is, $X \subseteq X$ for all sets X. The proof of this is extremely simple: we must prove $\forall x \in X, x \in X$. But then this is trivial: let $x \in X$, then $x \in X$ by assumption. Done!

Example 2.1.9

Let $a, b, c, d \in \mathbb{R}$ with $a < c < d < b$. Then $[c,d] \subseteq (a,b)$. Indeed, let $x \in [c,d]$. Then $c \leqslant x \leqslant d$. But then

$$a < c \leqslant x \leqslant d < b \quad \Rightarrow \quad a < x < b$$

so that $[c,d] \subseteq (a,b)$, as required.

Exercise 2.1.10

Let $a, b, c, d \in \mathbb{R}$ with $a < b$ and $c < d$. Prove that $[a,b) \subseteq (c,d]$ if and only if $a > c$ and $b \leqslant d$.

Example 2.1.11

The number sets from Chapter 0 are related by the following chain of subset inclusions.

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

The following proposition proves a property of subsethood known as *transitivity*—we'll revisit this property in Section 5.1.

Proposition 2.1.12 · *Transitivity of subsethood*

Let X, Y, Z be sets. If $X \subseteq Y$ and $Y \subseteq Z$, then $X \subseteq Z$.

Proof

Suppose that $X \subseteq Y$ and $Y \subseteq Z$. We need to prove $X \subseteq Z$.

So let $a \in X$. Since $X \subseteq Y$, it follows from Definition 2.1.6 that $a \in Y$; and since $Y \subseteq Z$, it follows again from Definition 2.1.6 that $a \in Z$.

Hence $X \subseteq Z$, as required. □

Set equality

This section is all about defining sets, comparing sets, and building new sets from old, and so to make much more progress, we first need to establish what we mean when we say that two sets are *equal*.

Discussion 2.1.13

Let X and Y be sets. What should it mean to say that X and Y are equal? Try to provide a precise definition of equality of sets before reading on.

There are different possible notions of 'sameness' for sets: we might want to say that two sets X and Y are equal when they have quite literally the same definition; or we might want to say that X and Y are equal when they contain the same objects as elements. For instance, suppose X is 'the set of all odd natural numbers' and Y is 'the set of all integers that are differences of consecutive perfect squares'—in this case, the first of these characterisations of equality might lead us to say $X \neq Y$, whereas the second would lead us to say $X = Y$.

Clearly, we have to state our terms at some point. And that point is now.

Axiom 2.1.14 · *Set extensionality*

Let X and Y be sets. Then $X = Y$ if and only if $\forall a$, $(a \in X \Leftrightarrow a \in Y)$, or equivalently, if $X \subseteq Y$ and $Y \subseteq X$.

This characterisation of set equality suggests the following strategy for proving that two sets are equal.

Strategy 2.1.15 · *Proof by double containment*

In order to prove that a set X is equal to a set Y, it suffices to:

- Prove $X \subseteq Y$, i.e. let $a \in X$ be an arbitrary element, and derive $a \in Y$; and then
- Prove $X \supseteq Y$, i.e. let $a \in Y$ be an arbitrary element, and derive $a \in X$.

We often write '$(\subseteq)$' and '$(\supseteq)$' to indicate the direction of the containment being proved.

Example 2.1.16

We prove that $\{x \in \mathbb{R} \mid x^2 \leqslant 1\} = [-1, 1]$ by double containment.

- $(\subseteq)$ Let $a \in \{x \in \mathbb{R} \mid x^2 \leqslant 1\}$. Then $a \in \mathbb{R}$ and $a^2 \leqslant 1$, so that $(1-a)(1+a) = 1 - a^2 \geqslant 0$. It follows that either:
  - $1 - a \geqslant 0$ and $1 + a \geqslant 0$, in which case $a \leqslant 1$ and $a \geqslant -1$, so that $a \in [-1, 1]$.
  - $1 - a \leqslant 0$ and $1 + a \leqslant 0$, in which case $a \geqslant 1$ and $a \leqslant -1$, which is a contradiction since $-1 < 1$.

  So we must have $a \in [-1, 1]$, as required.
- $(\supseteq)$ Let $a \in [-1, 1]$. Then $-1 \leqslant a \leqslant 1$, so $|a| \leqslant 1$, and hence $a^2 = |a|^2 \leqslant 1$, so that $a \in \{x \in \mathbb{R} \mid x^2 \leqslant 1\}$, as required.

✎ **Exercise 2.1.17**
Prove that $\{x \in \mathbb{R} \mid x^2 < x\} = (0,1)$.

The set extensionality axiom has the consequence that sets are independent of the order in which their elements appear in list notation, and they are independent of how many times a single element is written—that is, an element of a set is only 'counted' once, even if it appears multiple times in list notation. This is illustrated in the following exercise.

✎ **Exercise 2.1.18**
Prove by double containment that $\{0,1\} = \{1,0\}$ and $\{0,0\} = \{0\}$.

Inhabitation and emptiness

Another fundamental example of a set is the *empty set*, which is the set with no elements. But we have to be slightly careful about how we use the word 'the', since it implies *uniqueness*, and we don't know (yet) that two sets with no elements are necessarily equal. So first we will define what it means for a set to be empty, and then we'll show that there is exactly one empty set.

✦ **Definition 2.1.19**
A set X is **inhabited** (or **nonempty**) if it has at least one element; otherwise, it is **empty**.

The assertion that X is inhabited is equivalent to the logical formula $\exists a, a \in X$, and the assertion that X is empty is equivalent to the logical formula $\neg\exists a, a \in X$. This suggests the following strategy for proving that a set is inhabited, or that it is empty.

❖ **Strategy 2.1.20** · *Proving that a set is inhabited or empty*
In order to prove a set X is inhabited, it suffices to exhibit an element. In order to prove a set X is empty, assume that X is inhabited—that is, that there is some element $a \in X$—and derive a contradiction.

In other texts, the term *nonempty* is more common than *inhabited*, but there are reasons to prefer latter. Indeed, the statement 'X is non-empty' translates more directly to $\neg(\neg\exists a, a \in X)$, which has an unnecessary double-negative and suggests a proof of inhabitation by contradiction. For this reason, we use the term *inhabited* in this book.

Emptiness may seem like a trivial condition—and it is—but owing to its canonicity, it arises all over the place.

✎ **Example 2.1.21**
The set $A = \{x \in \mathbb{R} \mid x^2 = 2\}$ is inhabited: for example $\sqrt{2} \in \mathbb{R}$ and $\sqrt{2}^2 = 2$, so $\sqrt{2} \in A$. It also happens that $-\sqrt{2} \in A$, but we did not need to show this in order to conclude that A is inhabited.

However, the set $B = \{x \in \mathbb{Q} \mid x^2 = 2\}$ is empty since, if it were inhabited, then there would be an element $x \in B$. This would contradict the fact that $\sqrt{2}$ is irrational (Assumption 0.49, Proposition 4.3.12), since then x would be a rational number such that $x^2 = 2$.

Example 2.1.22

We observed in Example 0.26 that the set $[0]$ is empty; here's a more formal proof. Towards a contradiction, suppose $[0]$ is inhabited. Then there is some $k \in \mathbb{N}$ such that $k < 0$, which is a contradiction. Hence $[0]$ is empty, after all.

Exercise 2.1.23

Let $a, b \in \mathbb{R}$. Prove that $[a,b]$ is empty if and only if $a > b$, and that (a,b) is empty if and only if $a \geqslant b$.

The next exercise is a logical technicality, which is counterintuitive for the same reason that makes the principle of explosion (Axiom 1.1.47) difficult to grasp. However, it is extremely useful for proving facts about the empty set, as we will see soon in Theorem 2.1.25.

Exercise 2.1.24

Let E be an empty set and let $p(x)$ be a predicate with one free variable x with domain of discourse E. Show that the proposition $\forall x \in E,\, p(x)$ is true, and that the proposition $\exists x \in E,\, p(x)$ is false. What does the proposition $\forall x \in E,\, x \neq x$ mean in English? Is it true?

Thanks to the axiom of extensionality (Axiom 2.1.14), any two empty sets must be equal since they both contain the same elements—namely, no elements at all! This is made formal in the following theorem.

Theorem 2.1.25 · *Uniqueness of the empty set*

Let E and E' be sets. If E and E' are empty, then $E = E'$.

Proof

Suppose that E and E' are empty. The assertion that $E = E'$ is equivalent to

$$(\forall a \in E,\, a \in E') \wedge (\forall a \in E',\, a \in E)$$

But $\forall a \in E,\, a \in E'$ and $\forall a \in E',\, a \in E$ are both true by Exercise 2.1.24 since E and E' are empty. So $E = E'$, as claimed. □

Theorem 2.1.25 implies that there is one and only one empty set, thus allowing us to talk about 'the' empty set. It is so important that it has its own symbol, introduced next.

Notation 2.1.26

The **empty set** (also known as the **null set**) is the set with no elements, and is denoted by $\varnothing$ (LaTeX code: `\varnothing`).

Some authors write $\{\}$ instead of $\varnothing$, since $\{\}$ is simply the empty set expressed in list notation.

Exercise 2.1.27

Is it true or false that $\varnothing \in X$ for all sets X? Is it true or false that $\varnothing \subseteq X$ for all sets X?

Exercise 2.1.28

Explain why the sets $\varnothing$, $\{\varnothing\}$ and $\{\{\varnothing\}\}$ are different, and then determine the truth values of each of the following propositions:

(a) $\varnothing \in \varnothing$ (c) $\varnothing \in \{\varnothing\}$ (e) $\{\varnothing\} \subseteq \{\varnothing\}$

(b) $\varnothing \subseteq \varnothing$ (d) $\varnothing \in \{\{\varnothing\}\}$ (f) $\{\varnothing\} \subseteq \{\{\varnothing\}\}$

Find more examples of simple elementhood and subsethood propositions—that is, propositions of the form $A \in B$ or $A \subseteq B$—involving the three sets $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$; you should find at least one true and one false elementhood proposition, and at least one true and one false subsethood proposition.

Power sets

Definition 2.1.29

Let X be a set. The **power set** of X, written $\mathscr{P}(X)$ (LaTeX code: `\mathcal{P}`), is the set of all subsets of X.

Example 2.1.30

There are four subsets of $\{1,2\}$, namely

$$\varnothing, \quad \{1\}, \quad \{2\}, \quad \{1,2\}$$

so $\mathscr{P}(X) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$.

Exercise 2.1.31

Write out the elements of $\mathscr{P}(\{1,2,3\})$.

Exercise 2.1.32

Let X be a set. Show that $\varnothing \in \mathscr{P}(X)$ and $X \in \mathscr{P}(X)$.

Exercise 2.1.33

Write out the elements of $\mathscr{P}(\varnothing)$, $\mathscr{P}(\mathscr{P}(\varnothing))$ and $\mathscr{P}(\mathscr{P}(\mathscr{P}(\varnothing)))$.

Power sets are often a point of confusion because they bring the property of being a *subset* of one set to that of being an *element* of another, in the sense that for all sets U and X we have

$$U \subseteq X \quad \Leftrightarrow \quad U \in \mathscr{P}(X)$$

This distinction looks easy to grasp, but when the sets U and X look alike, it's easy to fall into various traps.

The following exercise is intended to help you overcome similar potential kinds of confusion by means of practice. Try to think precisely about what the definitions involved are.

Exercise 2.1.34

Determine, with proof, whether or not each of the following statements is true.

(a) $\mathscr{P}(\varnothing) \in \{\varnothing\}$;

(b) $\mathscr{P}(\varnothing) \in \mathscr{P}(\mathscr{P}(\varnothing))$;

(c) $\mathscr{P}(\mathscr{P}(\varnothing)) \in \{\varnothing, \{\varnothing, \{\varnothing\}\}\}$.

Repeat the exercise with all instances of '$\in$' replaced by '$\subseteq$'.

Section 2.2

Set operations

In Example 2.1.16 we noted that $[0,\infty)$ is the set of all non-negative real numbers. What if we wanted to talk about the set of all non-negative rational numbers instead? It would be nice if there was some expression in terms of $[0,\infty)$ and $\mathbb{Q}$ to denote this set.

This is where *set operations* come in—they allow us to use previously defined sets to introduce new sets.

Intersection ($\cap$)

The operation of *intersection* allows us to discuss the set of objects that a collection of sets has in common.

Definition 2.2.1 · *Intersection*
The **intersection** of a collection of sets is the set of all objects that are elements of all of the sets in the collection.

The general definition of intersection is a little bit abstract at first sight. We will work our way up to it, but we will begin by examining intersections of collections of just *two* sets. There is a convenient notation in this case.

Notation 2.2.2 · *Pairwise intersection*
Let A and B be sets. We write $A \cap B$ (LaTeX code: `\cap`) to denote the intersection of A and B; that is:

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

Example 2.2.3
The set $[0,\infty) \cap \mathbb{Q}$ is the intersection of $[0,\infty)$ and $\mathbb{Q}$, so is the set of all objects that are elements of both $[0,\infty)$ and $\mathbb{Q}$. The elements of $[0,\infty)$ are the non-negative real numbers, and the elements of $\mathbb{Q}$ are the rational numbers, and so $[0,\infty) \cap \mathbb{Q}$ is the set of all non-negative rational numbers.

Exercise 2.2.4
Prove that $[0,\infty) \cap \mathbb{Z} = \mathbb{N}$.

Exercise 2.2.5
Write down the elements of the set $\{0,1,4,7\} \cap \{1,2,3,4,5\}$.

Exercise 2.2.6
Express $[-2,5) \cap [4,7)$ as a single interval.

Exercise 2.2.7
Let $A \subseteq \mathbb{R}$. Find expressions using the intersection operation for (a) the set of all positive elements of A, and (b) the set of all integer elements of A.

Proposition 2.2.8
Let X and Y be sets. Then $X \subseteq Y$ if and only if $X \cap Y = X$.

Proof
Suppose that $X \subseteq Y$. We prove $X \cap Y = X$ by double containment.

- ($\subseteq$) Suppose $a \in X \cap Y$. Then $a \in X$ and $a \in Y$ by definition of intersection, so in particular we have $a \in X$.
- ($\supseteq$) Suppose $a \in X$. Then $a \in Y$ since $X \subseteq Y$, so that $a \in X \cap Y$ by definition of intersection.

Conversely, suppose that $X \cap Y = X$. To prove that $X \subseteq Y$, let $a \in X$. Then $a \in X \cap Y$ since $X = X \cap Y$, so that $a \in Y$ by definition of intersection, as required. □

Exercise 2.2.9
Let X be a set. Prove that $X \cap \varnothing = \varnothing$ and that $X \cap X = X$.

It is often useful to know when sets have *no* elements in common, because this captures the notion of mutual exclusivity. For example, the set of all even integers and the set of all odd integers have no elements in common, which is to say that evenness and oddness are mutually exclusive properties of integers. This property is called *disjointness*.

Definition 2.2.10
Let A and B be sets. We say A and B are **disjoint** if $A \cap B$ is empty.

Example 2.2.11
The sets $\{0,2,4\}$ and $\{1,3,5\}$ are disjoint, since they have no elements in common.

Exercise 2.2.12
Let $a,b,c,d \in \mathbb{R}$ with $a < b$ and $c < d$. Prove that the open intervals (a,b) and (c,d) are disjoint if and only if $b \leqslant c$ or $d \leqslant a$.

The notation for the intersection of an arbitrary finite collection of sets is not very surprising:

Notation 2.2.13 · *Finitary intersection*
Let $n \in \mathbb{N}$ with $n \geqslant 2$ and let $A_0, A_1, \ldots, A_{n-1}$ be sets. The intersection of $A_0, A_1, \ldots, A_{n-1}$ is denoted by

$$A_0 \cap A_1 \cap \cdots \cap A_{n-1} = \{x \mid x \in A_0 \text{ and } x \in A_1 \text{ and } \cdots \text{ and } x \in A_{n-1}\}$$

or, more concisely, $A_0 \cap A_1 \cap \cdots \cap A_{n-1} = \{x \mid x \in A_i \text{ for all } i \in [n]\}$.

For example, if A, B and C are sets, then

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

Example 2.2.14

The set $[0,\infty) \cap (-\infty,3) \cap \mathbb{Z}$ is the set of all objects that are elements of all three of...

- ... $[0,\infty)$, which is the set of all non-negative real numbers;
- ... $(-\infty,3)$, which is the set of all real numbers that are less than 3; and
- ... $\mathbb{Z}$, which is the set of all integers.

The objects that are in all three of these sets are therefore the integers n that satisfy the inequality $0 \leqslant n < 3$; and so

$$[0,\infty) \cap (-\infty,3) \cap \mathbb{Z} = \{0,1,2\}$$

Exercise 2.2.15

Let $A \subseteq \mathbb{R}$. Find an expression involving the intersection operation for the set of all non-negative rational elements of A.

We will often have occasion to take the intersection of an arbitrary collection of sets (possibly infinitely many). For example, we might want to reason about the set of all real numbers that are common to all of the intervals of the form $[0, 1 + \frac{1}{n})$ for natural number values of $n \geqslant 1$.

When a collection of sets can be defined in terms of a variable with a specified range, we can use that variable as a label (more formally called an *index*) for the sets in the collection, and then we can define and reason about set operations on the collection of sets in terms of the indexing variable. For example, in the previous paragraph we used the variable n, with the range $\{n \in \mathbb{N} \mid n \geqslant 1\}$, as an index for the collection of intervals of the form $[0, 1 + \frac{1}{n})$; we can then reason about the intersection of these intervals in terms of the indexing variable n.

Definition 2.2.16 · *Indexed families of sets*

Let I be a set. An **indexed family of sets** over I is a specification of a set A_i for each $i \in I$; thus the indexed family of sets is described in set-builder notation by $\{A_i \mid i \in I\}$.

The set I is called the **indexing set**, and the elements of I are called **indices**—namely, an element $i \in I$ is the **index** of the corresponding set A_i in the indexed family.

Example 2.2.17

The notation $\left\{[0, 1 + \frac{1}{n}) \;\middle|\; n \geqslant 1\right\}$ expresses the collection of sets of the form $[0, 1 + \frac{1}{n})$ as an indexed family of sets.

Formally, the indexing set for this family of sets is the set $I = \{n \in \mathbb{N} \mid n \geqslant 1\}$—in the notation we used, the fact that n refers to a natural number is left implicit from context. Given an index $n \in I$, the set in the family with index n is the interval $[0, 1 + \frac{1}{n})$.

Notation 2.2.18 · *Indexed intersection*
Let $\{A_i \mid i \in I\}$ be an indexed family of sets, for some indexing set I.

We write $\bigcap_{i\in I} A_i$ (LaTeX code: `\bigcap_{i \in I} A_i`) for the intersection of the sets in the indexed family; that is:

$$\bigcap_{i\in I} A_i \;=\; \{x \mid x \in A_i \text{ for all } i \in I\}$$

Variations of Notation 2.2.18 can be understood from context. For example, if the indexing set is given by $I = \{n \in \mathbb{N} \mid n \geqslant 1\}$, then we might write

$$\bigcap_{n\in I} A_n \;=\; \bigcap_{n\geqslant 1} A_n \;=\; \bigcap_{n=1}^{\infty} A_n$$

Example 2.2.19
We prove by double containment that $\bigcap_{n\geqslant 1} [0, 1 + \frac{1}{n}) = [0, 1]$.

Before we begin, let's unpack the definition of the indexed intersection. As explained in Example 2.2.17, the indexing set in this case is $I = \{n \in \mathbb{N} \mid n \geqslant 1\}$, with the set indexed by an element $n \in I$ being the interval $[0, 1 + \frac{1}{n})$. Therefore, for an arbitrary object x we have

$$x \in \bigcap_{n\geqslant 1} [0, 1 + \tfrac{1}{n}) \quad \Leftrightarrow \quad x \in [0, 1 + \tfrac{1}{n}) \text{ for all natural numbers } n \geqslant 1$$

We will use this in what follows.

- ($\subseteq$) Let $x \in \bigcap_{n\geqslant 1} [0, 1 + \frac{1}{n})$. Then $x \in [0, 1 + \frac{1}{n})$ for all natural numbers $n \geqslant 1$, so that $x \in \mathbb{R}$ and $0 \leqslant x < 1 + \frac{1}{n}$ for all natural numbers $n \geqslant 1$.

  Towards a contradiction, assume that $x > 1$, and define $N = \lceil \frac{1}{x-1} \rceil + 1$. This definition of N ensures that $N \in \mathbb{N}$, that $N \geqslant 1$, and that $N \geqslant \frac{1}{x-1}$. Rearranging the latter inequality reveals that $x \geqslant 1 + \frac{1}{N}$, contradicting the fact that $0 \leqslant x < 1 + \frac{1}{n}$ for all natural numbers $n \geqslant 1$.

  So we must have $x \leqslant 1$. Therefore $0 \leqslant x \leqslant 1$, and so $x \in [0, 1]$, as required.

- ($\supseteq$) Let $x \in [0, 1]$. Then $x \in \mathbb{R}$ and $0 \leqslant x \leqslant 1$.

  Let $n \in \mathbb{N}$ with $n \geqslant 1$. Then $0 \leqslant x \leqslant 1 < 1 + \frac{1}{n}$, so that $x \in [0, 1 + \frac{1}{n})$.

  This proves that $x \in [0, 1 + \frac{1}{n})$ for all natural numbers $n \geqslant 1$, and so $x \in \bigcap_{n\geqslant 1} [0, 1 + \frac{1}{n})$.

By double containment, it follows that $\bigcap_{n\geqslant 1} [0, 1 + \frac{1}{n}) = [0, 1]$.

Exercise 2.2.20
For each $n \in \mathbb{Z}$, let $A_n = \{x \in \mathbb{R} \mid x \neq n\}$. Describe the set $\bigcap_{n\in\mathbb{Z}} A_n$ in plain terms.

Union ($\cup$)

The operation of *union* allows us to discuss the set of objects that appear in at least one of the sets in a collection.

Definition 2.2.21 · *Union*
The **union** of a collection of sets is the set of all objects that are elements of at least one of the sets in the collection.

Just like with intersections, we will begin our study of unions by focusing on unions of pairs of sets, and then we will build up to unions of more general collections of sets.

Notation 2.2.22 · *Pairwise union*
Let A and B be sets. We write $A \cup B$ (LaTeX code: `\cup`) to denote the union of A and B; that is:

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

Example 2.2.23
Let E be the set of even integers and O be the set of odd integers. Since every integer is either even or odd, $E \cup O = \mathbb{Z}$. Note that $E \cap O = \varnothing$, thus $\{E, O\}$ is an example of a *partition* of $\mathbb{Z}$—see Definition 5.2.21.

Exercise 2.2.24
Write down the elements of the set

$$\{0,1,4,7\} \cup \{1,2,3,4,5\}$$

Exercise 2.2.25
Express $[-2,5) \cup [4,7)$ as a single interval.

The union operation allows us to define the following class of sets that will be particularly useful for us when studying counting principles in Section 8.1.

Exercise 2.2.26
Let X and Y be sets. Prove that $X \subseteq Y$ if and only if $X \cup Y = Y$.

There are many algebraic results concerning what happens when different set operations, such as intersection and union, are combined. The next example and exercise are two such results, and there are many more.

Example 2.2.27 · *Distributivity of intersection over union*
Let X, Y, Z be sets. We prove that $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

- ($\subseteq$) Let $x \in X \cap (Y \cup Z)$. Then $x \in X$, and either $x \in Y$ or $x \in Z$. If $x \in Y$ then $x \in X \cap Y$, and if $x \in Z$ then $x \in X \cap Z$. In either case, we have $x \in (X \cap Y) \cup (X \cap Z)$.
- ($\supseteq$) Let $x \in (X \cap Y) \cup (X \cap Z)$. Then either $x \in X \cap Y$ or $x \in X \cap Z$. In both cases we have $x \in X$ by definition of intersection. In the first case we have $x \in Y$, and in the second case we have $x \in Z$; in either case, we have $x \in Y \cup Z$, so that $x \in X \cap (Y \cup Z)$.

Exercise 2.2.28 · *Distributivity of union over intersection*
Let X, Y, Z be sets. Prove that $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$.

The notation for forming unions of arbitrary finite collections of sets works in exactly the same way as for intersections.

Notation 2.2.29 · *Finitary union*
Let $n \in \mathbb{N}$ with $n \geqslant 2$ and let $A_0, A_1, \ldots, A_{n-1}$ be sets. The union of $A_0, A_1, \ldots, A_{n-1}$ is denoted by

$$A_0 \cup A_1 \cup \cdots \cup A_{n-1} = \{x \mid x \in A_0 \text{ or } x \in A_1 \text{ or } \cdots \text{ or } x \in A_{n-1}\}$$

or, more concisely, $A_0 \cup A_1 \cup \cdots \cup A_{n-1} = \{x \mid x \in A_i \text{ for some } i \in [n]\}$.

For example, if A, B and C are sets, then

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

Example 2.2.30
We will prove that $\mathbb{R} = (-\infty, 0) \cup \{0\} \cup (0, \infty)$.

- ($\subseteq$) Let $x \in \mathbb{R}$. Either $x < 0$, in which case $x \in (-\infty, 0)$; or $x = 0$, in which case $x \in \{0\}$; or $x > 0$, in which case $x \in (0, \infty)$. In all three cases, x is an element of at least one of the three sets in the union, so that $x \in (-\infty, 0) \cup \{0\} \cup (0, \infty)$, as required.
- ($\supseteq$) Let $x \in (-\infty, 0) \cup \{0\} \cup (0, \infty)$. Then either $x \in (-\infty, 0)$, or $x \in \{0\}$, or $x \in (0, \infty)$. In all three cases, the definition of the set in question gives immediately that $x \in \mathbb{R}$; so indeed $x \in \mathbb{R}$, as required.

It follows by double containment that $\mathbb{R} = (-\infty, 0) \cup \{0\} \cup (0, \infty)$.

Exercise 2.2.31
Define $X = \{x \in \mathbb{R} \mid x^4 - 5x^2 + 4 > 0\}$. Find an expression for X as a union of intervals.

Exercise 2.2.32
State and prove distributivity laws for general finitary unions and intersections of sets, generalising Example 2.2.27 and Exercise 2.2.28.

The notation for unions of indexed families of sets mirrors that of intersections (Notation 2.2.18).

Notation 2.2.33 · *Indexed union*

Let $\{A_i \mid i \in I\}$ be an indexed family of sets, for some indexing set I.

We write $\bigcup_{i\in I} A_i$ (LaTeX code: `\bigcup_{i \in I} A_i`) for the union of the sets in the collection; that is:

$$\bigcup_{i\in I} A_i \;=\; \{x \mid x \in A_i \text{ for some } i \in I\}$$

Variations of Notation 2.2.33 can be understood from context. For example, if the indexing set is given by $I = \{n \in \mathbb{N} \mid n \geqslant 1\}$, then we might write

$$\bigcup_{n\in I} A_n \;=\; \bigcup_{n=1}^{\infty} A_n$$

Example 2.2.34

Given a set X, for each $n \in \mathbb{N}$ let $\mathscr{P}_n(X)$ refer to the set of all subsets of X that have exactly n elements, and let $\mathscr{F}(X)$ be the set of all finite subsets of X. (We have not formally defined what 'finite' means yet—that is coming in Section 6.1—but an intuitive understanding will suffice for the purposes of this exercise.)

We will prove that $\bigcup_{n\in\mathbb{N}} \mathscr{P}_n(X) = \mathscr{F}(X)$.

- ($\subseteq$) Let $A \in \bigcup_{n\in\mathbb{N}} \mathscr{P}_n(X)$. Then $A \in \mathscr{P}_n(X)$ for some $n \in \mathbb{N}$ by definition of indexed union, and so A is a subset of X with exactly n elements. But then A is a finite subset of X, so that $A \in \mathscr{F}(X)$, as required.
- ($\supseteq$) Let $A \in \mathscr{F}(X)$. Then A is a finite subset of X, and so A the number of elements of A is a natural number, say n. But then $A \in \mathscr{P}_n(X)$, and so $A \in \bigcup_{n\in\mathbb{N}} \mathscr{P}_n(X)$, as required.

It follows by double containment that $\bigcup_{n\in\mathbb{N}} \mathscr{P}_n(X) = \mathscr{F}(X)$.

Exercise 2.2.35

Express $\bigcup_{n\geqslant 1} [0, 1 - \frac{1}{n}]$ as an interval.

Exercise 2.2.36

Prove that $\bigcap_{n\in\mathbb{N}} [n] = \varnothing$ and $\bigcup_{n\in\mathbb{N}} [n] = \mathbb{N}$.

Exercise 2.2.37
Find a family of sets $\{X_n \mid n \in \mathbb{N}\}$ such that all three of the following properties hold: (i) $\bigcup_{n\in\mathbb{N}} X_n = \mathbb{N}$; (ii) $\bigcap_{n\in\mathbb{N}} X_n = \varnothing$; and (iii) $X_i \cap X_j \neq \varnothing$ for all $i, j \in \mathbb{N}$.

Relative complement ($\setminus$)

Definition 2.2.38
Let X and A be sets. The **relative complement** of A in X, denoted $X \setminus A$ (LaTeX code: `\setminus`), is defined by

$$X \setminus A = \{x \in X \mid x \notin A\} = \{x \mid x \in X \text{ and } x \notin A\}$$

The operation $\setminus$ is also known as the **set difference** operation.

Aside
The notation $X \setminus A$ is often read aloud as 'X minus A', and in fact some authors write $X - A$ instead of $X \setminus A$ for the relative complement operation. When using the same symbol for subtraction of numbers as for the relative complement of sets, it is important to avoid accidentally applying algebraic rules that are true for one operation but not the other; for example, it is true that $a - x = b - x$ implies $a = b$ for all $a, b, x \in \mathbb{R}$, but it is false that $A - X = B - X$ implies $A = B$ for all sets A, B, X.

Example 2.2.39
Let E be the set of all even integers. Then $n \in \mathbb{Z} \setminus E$ if and only if n is an integer and n is not an even integer; that is, if and only if n is odd. Thus $\mathbb{Z} \setminus E$ is the set of all odd integers.

Moreover, $n \in \mathbb{N} \setminus E$ if and only if n is a natural number and n is not an even integer. Since the even integers which are natural numbers are precisely the even natural numbers, $\mathbb{N} \setminus E$ is precisely the set of all odd natural numbers.

Exercise 2.2.40
Write down the elements of the set

$$\{0, 1, 4, 7\} \setminus \{1, 2, 3, 4, 5\}$$

Exercise 2.2.41
Express $[-2, 5) \setminus [4, 7)$ and $[4, 7) \setminus [-2, 5)$ as intervals.

Exercise 2.2.42
Let A be the set of all real numbers that are not integers.

(a) Find an expression for A as a relative complement of one set in another.

(b) Find an expression for A as a union of intervals.

Exercise 2.2.43

Let X and A be sets. Prove that $X \setminus (X \setminus A) = X \cap A$. Deduce that $A \subseteq X$ if and only if $X \setminus (X \setminus A) = A$.

Aside

It is common for introductory mathematics textbooks to introduce a set operation called (*absolute*) *complement*: the complement of a set A, typically denoted by A^{c} or $\overline{A}$, is defined to be the set of all objects that are not elements of A. Without additional conditions being imposed, this is problematic because the complement of a set is not generally a set—at least, not within the universe of discourse. For example, the complement of $\varnothing$ is the entire union of discourse.

Because of this, we will not use this absolute notion of complement, and will restrict our attention to *relative* complements.

However, there are contexts where the sets of interest are all subsets of some fixed set. For example, we might want to study subsets of a sample space Ω in the context of probability theory (as in Chapter 11), or subsets of a fixed set X in the context of order theory (as in Section 12.1). In these contexts, we may use the *notation* for absolute complements, but it will formally be the relative complement within the larger set—that is, $A^{\mathsf{c}} = \Omega \setminus A$ in the former context, and $A^{\mathsf{c}} = X \setminus A$ in the latter.

This is not something the reader should worry about in this section, unless they are simultaneously using other resources that use the absolute complement notation.

Comparison with logical operators and quantifiers

The astute reader will have noticed some similarities between set operations and the logical operators and quantifiers that we saw in Chapter 1.

Indeed, this can be summarised in the following table. In each row, the expressions in both columns are equivalent, where p denotes '$a \in X$', q denotes '$a \in Y$', and $r(i)$ denotes '$a \in X_i$'.

sets	logic
$a \notin X$	$\neg p$
$a \in X \cap Y$	$p \wedge q$
$a \in X \cup Y$	$p \vee q$
$a \in \bigcap_{i \in I} X_i$	$\forall i \in I, r(i)$
$a \in \bigcup_{i \in I} X_i$	$\exists i \in I, r(i)$
$a \in X \setminus Y$	$p \wedge (\neg q)$

This translation between logic and set theory does not stop there; in fact, as the following theorem shows, De Morgan's laws for the logical operators (Theorem 1.3.24) and for quantifiers (Theorem 1.3.29) also carry over to the set operations of union and intersection.

Theorem 2.2.44 · *De Morgan's laws for sets*
Given sets X, A, B and a family $\{A_i \mid i \in I\}$, we have

(a) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$;

(b) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$;

(c) $X \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X \setminus A_i)$;

(d) $X \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \setminus A_i)$.

Proof of (a)
We proceed by double containment.

- $(\subseteq)$. Let $x \in X \setminus (A \cup B)$. Then $x \in A$ and $x \notin A \cup B$ by definition of relative complement.

  Since $x \notin A \cup B$, we have $x \notin A$ and $x \notin B$; the fact that this is a conjunction ('and') and not a disjunction ('or') ultimately follows from the definition of intersection and de Morgan's laws for logical operators (Theorem 1.3.24).

  Since $x \in X$ and $x \notin B$ we have $x \in X \setminus A$, and since $x \in X$ and $x \notin B$ we have $x \in X \setminus B$; these are both by definition of relative complement.

  Finally, by definition of intersection, it follows that $x \in (X \setminus A) \cap (X \setminus B)$, as required.

- $(\supseteq)$. Let $x \in (X \setminus A) \cap (X \setminus B)$. Then $x \in X \setminus A$ and $x \in X \setminus B$ by definition of intersection.

  Since $x \in X \setminus A$, we have $x \in X$ and $x \notin A$ by definition of relative complement. Furthermore, since $x \in X \setminus B$, we also have $x \notin B$.

  Since $x \notin A$ and $x \notin B$, it follows that $x \notin A \cup B$ by definition of union (together with de Morgan's laws for logical operators).

  Finally, since $x \in X$ and $x \notin A \cup B$, we have $x \in X \setminus (A \cup B)$ by definition of relative complement, as required.

By double containment, it follows that $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$. □

Exercise 2.2.45
Complete the proof of de Morgan's laws for sets.

Cartesian product ($\times$, $\prod$, A^k)

The set extensionality axiom (Axiom 2.1.14) tells us that two sets are equal when each element of one set is an element of the other. As remarked on Section 2.1, the notion of a set cannot on its own capture the ideas of listing its elements in an order or with repeats—for example, the expressions $\{1,2,3\}$, $\{3,2,1\}$ and $\{1,2,2,3,3,3\}$ both represent the same set. For this we need a new concept, namely that of an *ordered list*.

Definition 2.2.46 · *Ordered lists, a.k.a. tuples*

A (**finite**) **ordered list** is an expression of the form

$$(x_0, x_1, \ldots, x_{k-1})$$

for some $k \in \mathbb{N}$, where each x_i for $i \in [k]$ is some object. The natural number k is called the **length** of the ordered list. The natural number $i \in [k]$ is the **position** (or **index**) of the object x_i in the list.

Ordered lists satisfy the property that

$$(x_0, x_1, \ldots, x_{k-1}) = (y_0, y_1, \ldots, y_{\ell-1}) \quad \Leftrightarrow \quad k = \ell \text{ and } x_i = y_i \text{ for all } i \in [k]$$

That is, two ordered lists are equal if and only if they have the same length, and have the same objects in all positions.

The term k**-tuple** refers to an ordered list of length k. The 0-tuple $()$ is the **empty list**. We will generally blur the distinction between objects x and the corresponding 1-tuple (x). Ordered 2-tuples, 3-tuples, 4-tuples, and so on, are respectively referred to as **ordered pairs**, **ordered triples**, **ordered quadruples**, and so on.

Example 2.2.47

Consider the ordered lists $(1,2,3)$, $(3,2,1)$ and $(1,2,3,3,3)$:

- $(1,2,3)$ and $(3,2,1)$ are ordered triples (3-tuples), and $(1,2,3,3,3)$ is an ordered quintuple (5-tuple).
- $(1,2,3) \neq (1,2,3,3,3)$ since these lists have different lengths.
- $(1,2,3) \neq (3,2,1)$ since they have different objects in position 0 (and in position 2, though simply having a different object in position 0 is sufficient to prove that they are not equal).

The above observations illustrate the difference between *ordered lists* and *list notation for sets*: as mentioned earlier the sets $\{1,2,3\}$, $\{3,2,1\}$ and $\{1,2,3,3,3\}$ are all equal to one another, since they have exactly the same elements (namely the natural numbers 1, 2 and 3).

Exercise 2.2.48

Write down all ordered lists that have all of the elements of $\{1,2,3\}$ as terms and have no repeated terms.

The *cartesian product* operation is used to construct sets of ordered lists, where the sets in the product determine what kind of object the corresponding term in the list can be.

Definition 2.2.49 · *Cartesian product*

Given $k \in \mathbb{N}$, the **cartesian product** of sets $A_0, A_1, \ldots, A_{k-1}$ is defined by

$$\prod_{i \in [k]} A_i \;=\; A_0 \times A_1 \times \cdots \times A_{k-1} \;=\; \{(x_0, x_1, \ldots, x_{k-1}) \mid x_i \in A_i \text{ for all } i \in [k]\}$$

(LaTeX code: `\prod_i \in [k]`)

Thus the elements of $\prod_{i\in[k]} A_i$ are k-tuples, where for each $i \in [k]$ the term in position i is an element of the set A_i.

Example 2.2.50

If you have ever taken calculus, you will probably be familiar with the set $\mathbb{R}\times\mathbb{R}$.

$$\mathbb{R}\times\mathbb{R} = \{(x,y) \mid x,y \in \mathbb{R}\}$$

Formally, this is the set of ordered pairs of real numbers. Geometrically, if we interpret $\mathbb{R}$ as an infinite line, the set $\mathbb{R}\times\mathbb{R}$ is the (real) plane: an element $(x,y) \in \mathbb{R}\times\mathbb{R}$ describes the point in the plane with coordinates (x,y).

We can investigate this further. For example, the following set:

$$\mathbb{R}\times\{0\} = \{(x,0) \mid x \in \mathbb{R}\}$$

is precisely the x-axis. We can describe graphs as subsets of $\mathbb{R}\times\mathbb{R}$. Indeed, the graph of $y = x^2$ is given by

$$G = \{(x,y) \in \mathbb{R}\times\mathbb{R} \mid y = x^2\} = \{(x,x^2) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}\times\mathbb{R}$$

This is an example of a *graph* of a function. This notion will be defined in generality later (Definition 3.1.6).

Exercise 2.2.51

Write down the elements of the set $\{1,2\}\times\{3,4,5\}$.

Exercise 2.2.52

Let X be a set. Prove that $X\times\varnothing = \varnothing$.

Exercise 2.2.53

Let X, Y and Z be sets. Under what conditions is it true that $X\times Y = Y\times X$?

Notation 2.2.54 · *k-fold cartesian product*

Let $k \in \mathbb{N}$ and let A be a set. We write A^k (LaTeX code: `A^k`) to denote the k-fold cartesian product of A; that is

$$A^k = \prod_{i\in[k]} A = \underbrace{A\times A\times\cdots\times A}_{k \text{ times}}$$

Thus $A^k = \{(x_0,x_1,\ldots,x_{k-1}) \mid x_i \in A \text{ for all } i \in [k]\}$ is the set of all k-tuples of elements of A.

In particular, $A^0 = \{()\}$, and by our convention of blurring the distinction between 1-tuples (x) and objects x, we have $A^1 = A$.

Exercise 2.2.55

Consider the following sets:

$$\mathbb{R}^3, \quad \mathbb{R}^2 \times \mathbb{R}, \quad \mathbb{R} \times \mathbb{R}^2, \quad \mathbb{R} \times \mathbb{R} \times \mathbb{R}, \quad (\mathbb{R} \times \mathbb{R}) \times \mathbb{R}, \quad \mathbb{R} \times (\mathbb{R} \times \mathbb{R})$$

Are all of these sets equal? If not, which pairs are (and which are not) equal to each other, and is there another sense in which they are all equivalent?

Section 2.E

Chapter 2 exercises

Sets

2.E.1. Express the following sets in the indicated form of notation.

(a) $\{n \in \mathbb{Z} \mid n^2 < 20\}$ in list notation;

(b) $\{4k+3 \mid k \in \mathbb{N}\}$ in implied list notation;

(c) The set of all odd multiples of six in set-builder notation;

(d) The set $\{1, 2, 5, 10, 17, \ldots, n^2+1, \ldots\}$ in set-builder notation.

2.E.2. Find sets X_n for each $n \in \mathbb{N}$ such that $X_{n+1} \subsetneq X_n$ for all $n \in \mathbb{N}$. Can any of the sets X_n be empty?

2.E.3. Express the set $\mathscr{P}(\{\varnothing, \{\varnothing, \{\varnothing\}\}\})$ in list notation.

2.E.4. Let X be a set and let $U, V \subseteq X$. Prove that U and V are disjoint if and only if $U \subseteq X \setminus V$.

2.E.5. For each of the following statements, determine whether or not it is true for all sets A and X, and prove your claim.

(a) If $X \setminus A = \varnothing$, then $X = A$.

(b) If $X \setminus A = X$, then $A = \varnothing$.

(c) If $X \setminus A = A$, then $A = \varnothing$.

(d) $X \setminus (X \setminus A) = A$.

2.E.6. For each of the following statements, determine whether it is true for all sets X, Y, false for all sets X, Y, or true for some choices of X and Y and false for others.

(a) $\mathscr{P}(X \cup Y) = \mathscr{P}(X) \cup \mathscr{P}(Y)$

(b) $\mathscr{P}(X \cap Y) = \mathscr{P}(X) \cap \mathscr{P}(Y)$

(c) $\mathscr{P}(X \times Y) = \mathscr{P}(X) \times \mathscr{P}(Y)$

(d) $\mathscr{P}(X \setminus Y) = \mathscr{P}(X) \setminus \mathscr{P}(Y)$

2.E.7. Let F be a set whose elements are all sets. Prove that if $\forall A \in F,\ \forall x \in A,\ x \in F$, then $F \subseteq \mathscr{P}(F)$.

Questions 2.E.9 to 2.E.14 concern the *symmetric difference* of sets, defined below.

Definition 2.E.8

The **symmetric difference** of sets X and Y is the set $X \triangle Y$ (LaTeX code: `\triangle`) defined by

$$X \triangle Y = \{a \mid a \in X \text{ or } a \in Y \text{ but not both}\}$$

2.E.9. Prove that $X \triangle Y = (X \setminus Y) \cup (Y \setminus X) = (X \cup Y) \setminus (X \cap Y)$ for all sets X and Y.

2.E.10. Let X be a set. Prove that $X \triangle X = \varnothing$ and $X \triangle \varnothing = X$.

2.E.11. Let X and Y be sets. Prove that $X = Y$ if and only if $X \triangle Y = \varnothing$.

2.E.12. Prove that sets X and Y are disjoint if and only if $X \triangle Y = X \cup Y$.

2.E.13. Prove that $X \triangle (Y \triangle Z) = (X \triangle Y) \triangle Z$ for all sets X, Y and Z.

2.E.14. Prove that $X \cap (Y \triangle Z) = (X \cap Y) \triangle (X \cap Z)$ for all sets X, Y and Z.

Definition 2.E.15
A subset $U \subseteq \mathbb{R}$ is **open** if, for all $a \in U$, there exists $\delta > 0$ such that $(a-\delta, a+\delta) \subseteq U$.

In Questions Questions 2.E.16 to 2.E.19 you will prove some elementary facts about open subsets of $\mathbb{R}$.

2.E.16. For each of the following subsets of $\mathbb{R}$, determine (with proof) whether it is open:

(a) $\varnothing$;
(b) $(0,1)$;
(c) $(0,1]$;
(d) $\mathbb{Z}$;
(e) $\mathbb{R} \setminus \mathbb{Z}$;
(f) $\mathbb{Q}$.

2.E.17. Prove that a subset $U \subseteq \mathbb{R}$ is open if and only if, for all $a \in U$, there exist $u, v \in \mathbb{R}$ such that $u < a < v$ and $(u, v) \subseteq U$.

2.E.18. In this question you will prove that the intersection of finitely many open sets is open, but the intersection of infinitely many open sets might not be open.

(a) Let $n \geqslant 1$ and suppose $U_1, U_2, \ldots, U_n$ are open subsets of $\mathbb{R}$. Prove that the intersection $U_1 \cap U_2 \cap \cdots \cap U_n$ is open.

(b) Prove that $(0, 1 + \frac{1}{n})$ is open for all $n \geqslant 1$, but that $\bigcap_{n \geqslant 1} (0, 1 + \frac{1}{n})$ is not open.

2.E.19. Prove that a subset $U \subseteq \mathbb{R}$ is open if and only if it can be expressed as a union of open intervals—more precisely, $U \subseteq \mathbb{R}$ is open if and only if, for some indexing set I, there exist real numbers a_i, b_i for each $i \in I$, such that $U = \bigcup_{i \in I} (a_i, b_i)$.

2.E.20. Let $\{A_i \mid i \in \mathbb{N}\}$ and $\{B_j \mid j \in \mathbb{N}\}$ be families of sets such that:

(i) For all $i \in \mathbb{N}$, there exists some $j \geqslant i$ such that $B_j \subseteq A_i$; and

(ii) For all $j \in \mathbb{N}$, there exists some $i \geqslant j$ such that $A_i \subseteq B_j$.

Prove that $\bigcap_{i \in \mathbb{N}} A_i = \bigcap_{j \in \mathbb{N}} B_j$.

True–False questions

In Questions 2.E.21 to 2.E.26, determine (with proof) whether the statement is true or false.

2.E.21. If E is a set and $\neg \forall x, x \in E$, then $E = \varnothing$.

2.E.22. If E is a set and $\forall x, \neg x \in E$, then $E = \varnothing$.

2.E.23. $\{1,2,3\} = \{1,2,1,3,2,1\}$

2.E.24. $\varnothing \in \varnothing$

2.E.25. $\varnothing \in \{\varnothing\}$

2.E.26. $\varnothing \in \{\{\varnothing\}\}$.

Always–Sometimes–Never questions

In Questions 2.E.27 to 2.E.42, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

2.E.27. Let a and b be arbitrary objects. Then $\{a,b\} = \{b,a\}$.

2.E.28. Let a, b and c be arbitrary objects. Then $\{a,b\} = \{a,c\}$.

2.E.29. Let X and Y be sets and suppose there is some a such that $a \in X \Leftrightarrow a \in Y$. Then $X = Y$.

2.E.30. Let X and Y be sets and suppose there is some a such that $a \in X$ but $a \notin Y$. Then $X = Y$.

2.E.31. Let X and Y be sets with $X \neq Y$. Then $\forall a, a \in X \Rightarrow a \notin Y$.

2.E.32. Let X and Y be sets with $X \cap Y = \varnothing$. Then $\forall a, a \in X \Rightarrow a \notin Y$.

2.E.33. Let X and Y be sets with $X \neq Y$. Then $\exists a, a \in X \wedge a \notin Y$.

2.E.34. Let X and Y be sets of sets and suppose that $X \in Y$. Then $X \subseteq Y$.

2.E.35. Let X be a set. Then $\varnothing \in \mathscr{P}(X)$.

2.E.36. Let X be a set. Then $\{\varnothing\} \in \mathscr{P}(X)$.

2.E.37. Let X be a set. Then $\{\varnothing, X\} \subseteq \mathscr{P}(X)$.

2.E.38. Let X and Y be sets such that $\mathscr{P}(X) = \mathscr{P}(Y)$. Then $X = Y$.

2.E.39. Let X and Y be sets such that $X \setminus Y = \varnothing$. Then $X = Y$.

2.E.40. Let X, Y and Z be sets such that $X \setminus Y = Z$ and $Y \subseteq Z$. Then $X = Y \cup Z$.

2.E.41. Let X and Y be sets. Then $X \cup Y = X \cap Y$.

2.E.42. Let X be a set and let $A \subseteq X$. Then $A \in X$.

Chapter 3

Functions

If sets are the bread of mathematics, then functions are the butter. As we will see in subsequent chapters of this book, functions are central in almost every area of mathematical study: they allow us to transform elements of one set into elements of another, to define 'infinity', and to capture real-world notions such as probability and randomness in the abstract.

It is likely (but not assumed) that you have seen functions before, such as real-valued functions in calculus or linear transformations in linear algebra. However, we will study functions in the abstract; the 'inputs' and 'outputs' of our functions need not be numbers, vectors or points in space; they can be anything at all—in fact, the inputs or outputs to our functions might themselves be functions!

We introduce the notion of a function abstractly in Section 3.1. Much of our time will be spent developing basic notions involving functions, including graphs, composition, images and preimages.

We will zoom in on two properties in particular in Section 3.2, namely *injectivity* and *surjectivity*. These properties allow us to compare the sizes of sets we will use them extensively in Chapters 8 and 10 for doing exactly that.

Section 3.1

Functions

One way of studying interactions between sets is by studying *functions* between them, which we will define informally in Definition 3.1.1. Functions are mathematical objects which assign to each element of one set exactly one element of another. Almost every branch of mathematics studies functions, be it directly or indirectly, and almost every application of mathematics arises from a translation of the abstract notion of a function to the real world. Just one example of this is the theory of computation—functions provide precisely the language necessary to describe the deterministic input-output behaviour of algorithms.

You might have come across the notion of a function before now. In schools, functions are often introduced as being like *machines*—they have inputs and outputs, and on a given input they always return the same output. For instance, there is a function which takes integers as inputs and gives integers as outputs, which on the input x returns the integer $x+3$.

This characterisation of functions, however, is clearly not precise enough for the purposes of mathematical proof. A next approximation to a precise definition of a function might look something like this:

Definition 3.1.1

A **function** f from a set X to a set Y is a specification of elements $f(x) \in Y$ for $x \in X$, such that

$$\forall x \in X, \exists! y \in Y, y = f(x)$$

Given $x \in X$, the (unique!) element $f(x) \in Y$ is called the **value** of f at x.

The set X is called the **domain** (or **source**) of f, and Y is called the **codomain** (or **target**) of f. We write $f : X \to Y$ (LaTeX code: `f : X \to Y`) to denote the assertion that f is a function with domain X and codomain Y.

This is better—we're now talking about sets, and not mysterious 'machines'.

Moreover, this definition establishes a close relationship between functions and the $\exists!$ quantifier: indeed, to say that f assigns to each element of X a unique element of Y is to say precisely that

$$\forall x \in X, \exists! y \in Y, y = f(x)$$

Conversely, any true proposition of the form $\forall x \in X, \exists! y \in Y, p(x,y)$ defines a function $f : X \to Y$: the function f assigns to each $x \in X$ the unique $y \in Y$ such that $p(x,y)$ is true. In other words, $\forall x \in X, p(x, f(x))$ is true!

We can use this to generate some examples of functions.

Example 3.1.2

Example 1.2.27 said that every positive real number has a unique positive square root; we proved this in Example 1.2.30. What this means is that there is a function $r : (0,\infty) \to (0,\infty)$ defined

by letting $r(x)$ be the (unique) positive square root of x, for each $x \in (0,\infty)$. That is, we have a function r defined by $r(x) = \sqrt{x}$.

Exercise 3.1.3

Recall Exercise 1.2.31. Which of the statements (a), (b) or (c) is of the form $\forall x \in X, \exists! y \in Y, p(x,y)$? For each statement of this form, determine the domain and codomain of the corresponding function, and write an expression defining this function.

Specifying a function

Just like with sets, there are many ways to specify a function $f : X \to Y$, but when we do so, we must be careful that what we write really *does* define a function!

This correctness of specification is known as *well-definedness*, and ultimately amounts to verifying that the condition $\forall x \in X, \exists! y \in Y, f(x) = y$ holds for the specification of f. Namely *totality*, *existence* and *uniqueness*:

- **Totality.** A value $f(x)$ should be specified for each $x \in X$—this corresponds to the '$\forall x \in X$' quantifier in the definition of functions.
- **Existence.** For each $x \in X$, the specified value $f(x)$ should actually exist, and should be an element of Y—this corresponds to the *existence* part of the '$\exists! y \in Y$' quantifier in the definition of functions.
- **Uniqueness.** For each $x \in X$, the specified value $f(x)$ should refer to only one element of Y—this corresponds to the *uniqueness* part of the '$\exists! y \in Y$' quantifier in the definition of functions.

When specifying a function, you should justify each of these components of well-definedness unless they are extremely obvious. You will probably find that, in most cases, the only component in need of justification is uniqueness, but keep all three in mind.

Lists. If X is finite, then we can specify a function $f : X \to Y$ by simply listing the values of f at all possible elements $x \in X$. For example, we can define a function

$$f : \{1,2,3\} \to \{\mathsf{red}, \mathsf{yellow}, \mathsf{green}, \mathsf{blue}, \mathsf{purple}\}$$

by declaring

$$f(1) = \mathsf{red}, \quad f(2) = \mathsf{purple}, \quad f(3) = \mathsf{green}$$

Note that the function is at this point completely specified: we know its values at all elements of the domain $\{1,2,3\}$. It doesn't matter that some of the elements of the codomain (yellow and blue) are unaccounted for—all that matters is that each element of the domain is associated with exactly one element of the codomain.

Unfortunately, most of the sets that we work with will be infinite, or of an unspecified finite size; in these cases, simply writing a list of values isn't sufficient. Fortunately for us, there are other ways of specifying functions.

Formulae. In many cases, particularly when the domain X and codomain Y are number sets, we can define a function by giving a formula for the value of $f(x)$ for each $x \in X$. For example, we can define a function $f : \mathbb{R} \to \mathbb{R}$ by letting

$$f(x) = x^2 + 3 \text{ for all } x \in \mathbb{R}$$

By cases. It will at times be convenient to define a function using different specifications for different elements of the domain. A very simple example is the *absolute value function* $|-| : \mathbb{R} \to \mathbb{R}$, defined for $x \in \mathbb{R}$

$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x \leqslant 0 \end{cases}$$

Here we have split into two cases based on the conditions $x \geqslant 0$ and $x \leqslant 0$.

When specifying a function $f : X \to Y$ by cases, it is important that the conditions be:

- **exhaustive**: given $x \in X$, at least one of the conditions on X must hold; and
- **compatible**: if any $x \in X$ satisfies more than one condition, the specified value must be the same no matter which condition is picked.

For the absolute value function defined above, these conditions are satisfied. Indeed, for $x \in \mathbb{R}$, it is certainly the case that $x \geqslant 0$ or $x \leqslant 0$, so the conditions are exhaustive. Moreover, given $x \in \mathbb{R}$, if both $x \geqslant 0$ and $x \leqslant 0$, then $x = 0$—so we need to check that the specification yields the same value when $x = 0$ regardless of which condition we pick. The $x \geqslant 0$ condition yields the value 0, and the $x \leqslant 0$ condition yields the value -0, which is equal to 0—so the conditions are compatible. We could have used $x < 0$ instead of $x \leqslant 0$; in this case the conditions are *mutually exclusive*, so certainly compatible because they do not overlap.

Algorithms. You might, on first exposure to functions, have been taught to think of a function as a *machine* which, when given an *input*, produces an *output*. This 'machine' is defined by saying what the possible inputs and outputs are, and then providing a list of instructions (an *algorithm*) for the machine to follow, which on any input produces an output—and, moreover, if fed the same input, the machine always produces the same output.

For example, we might instruct a machine to take rational numbers as inputs and give rational numbers as outputs, and to follow the following sequence of steps on a given input

$$\text{multiply by } 2 \to \text{add } 5 \to \text{square the result} \to \text{divide by } 6$$

This 'machine' defines a function $M : \mathbb{Q} \to \mathbb{Q}$ which, in equation form, is specified by

$$M(x) = \frac{(2x+5)^2}{6} \text{ for all } x \in \mathbb{Q}$$

In our more formal set-up, therefore, we can define a function $M : I \to O$ by specifying:

- a set I of all **inputs**;
- a set O of potential **outputs**; and

- a deterministic[a] algorithm which describes how an input $x \in I$ is transformed into an output $M(x) \in O$.

That is, the domain is the set I of all possible 'inputs', the codomain is a set O containing all the possible 'outputs', and the function M is a rule specifying how an input is associated with the corresponding output.

For now, we will use algorithmic specifications of functions only sparingly—this is because it is much harder to make formal what is meant by an 'algorithm', and it is important to check that a given algorithm is deterministic.

Function equality

In Section 2.1 we discussed how there may be many different possible ways of characterising equality of sets. This matter was resolved by declaring that two sets are equal if and only if they have the same elements (this was Axiom 2.1.14).

A similar matter arises for functions. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$ for all $x \in \mathbb{R}$, and the function $g : \mathbb{R} \to \mathbb{R}$, defined by letting $g(x)$ be the result of taking x, multiplying it by three, dividing the result by four, dividing the result by six, and then multiplying the result by sixteen. It so happens that $g(x) = 2x$ for all $x \in \mathbb{R}$ as well, but that is not how g is defined; moreover, if f and g were implemented as algorithms, then it would take longer to compute the values of g than it would take to compute the values of f.

Should we consider f and g to be *equal*? If we are only interested in whether f and g have the same values on each argument, then the answer should be 'yes'; if we are interested in the algorithmic behaviour of f and g, then the answer should be 'no'.

We resolve this dilemma with the following axiom. By adopting this axiom, we are stating that the functions f and g discussed above are equal.

Axiom 3.1.4 · *Function extensionality*
Let $f : X \to Y$ and $g : A \to B$ be functions. Then $f = g$ if and only if the following conditions hold:

(i) $X = A$ and $Y = B$; and

(ii) $f(x) = g(x)$ for all $x \in X$.

Strategy 3.1.5 · *Proving two functions are equal*
Given functions $f, g : X \to Y$ with the same domain and codomain, in order to prove that $f = g$, it suffices to prove that $f(x) = g(x)$ for all $x \in X$.

A consequence of Axiom 3.1.4 is that, for fixed sets X and Y, a function $X \to Y$ is uniquely determined by its input-output pairs. This set is called the *graph* of the function; the proof of the equivalence between functions and their graphs is the content of Theorem 3.1.9.

[a] The word 'deterministic' just means that the algorithm always produces the same output on a single input.

Definition 3.1.6

Let $f : X \to Y$ be a function. The **graph** of f is the subset $\mathrm{Gr}(f) \subseteq X \times Y$ (LaTeX code: `\mathrm{Gr}`) defined by

$$\mathrm{Gr}(f) = \{(x, f(x)) \mid x \in X\} = \{(x, y) \in X \times Y \mid y = f(x)\}$$

Example 3.1.7

Given a (sufficiently well-behaved) function $f : \mathbb{R} \to \mathbb{R}$, we can represent $\mathrm{Gr}(f) \subseteq \mathbb{R} \times \mathbb{R}$ by plotting it on a pair of axes using Cartesian coordinates in the usual way. For example, if f is defined by $f(x) = \frac{x}{2}$ for all $x \in \mathbb{R}$, then its graph

$$\mathrm{Gr}(f) = \left\{ \left(x, \frac{x}{2}\right) \;\middle|\; x \in \mathbb{R} \right\}$$

can be represented by graph plot in Figure 3.1.

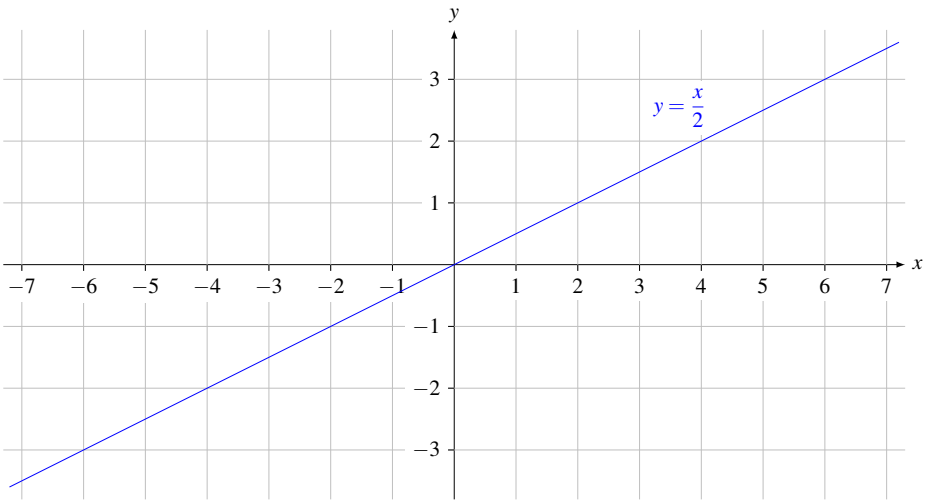


Figure 3.1: Graph of the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \frac{x}{2}$ for all $x \in \mathbb{R}$

Exercise 3.1.8

Find a function $f : \mathbb{Z} \to \mathbb{Z}$ whose graph is equal to the set

$$\{\ldots, (-2, -5), (-1, -2), (0, 1), (1, 4), (2, 7), (3, 10), \ldots\}$$

Theorem 3.1.9 below provides a way of verifying that a function is well-defined by characterising their graphs.

Theorem 3.1.9

Let X and Y be sets. A subset $G \subseteq X \times Y$ is the graph of a function if and only if

$$\forall x \in X, \exists! y \in Y, (x, y) \in G$$

Proof
($\Rightarrow$). Suppose $G \subseteq X \times Y$ is the graph of a function, say $G = \mathrm{Gr}(f)$ for some $f : X \to Y$. Then for each $x \in X$, it follows from well-definedness of f that $f(x)$ is the unique element $y \in Y$ for which $(x,y) \in G$. That is, $(x, f(x)) \in G$, and if $y \in Y$ with $(x,y) \in G$, then $y = f(x)$.

($\Leftarrow$). Suppose $G \subseteq X \times Y$ satisfies $\forall x \in X, \exists! y \in Y, (x,y) \in G$. Define a function $f : X \to Y$ by, for each $x \in X$, defining the value $f(x)$ to be the unique element $y \in Y$ for which $(x,y) \in G$. Well-definedness of f is then immediate from our assumption of the existence and uniqueness of such a value of y for each $x \in X$. □

Example 3.1.10

The set G defined by

$$G = \{(1, \mathsf{red}), (2, \mathsf{red}), (3, \mathsf{green})\}$$

is the graph of a function $f : \{1,2,3\} \to \{\mathsf{red}, \mathsf{green}, \mathsf{blue}\}$. The function f is defined by

$$f(1) = \mathsf{red}, \quad f(2) = \mathsf{red}, \quad f(3) = \mathsf{green}$$

However, G is *not* the graph of a function $\{1,2,3,4\} \to \{\mathsf{red}, \mathsf{green}, \mathsf{blue}\}$, since G contains no elements of the form $(4,y)$ for $y \in \{\mathsf{red}, \mathsf{green}, \mathsf{blue}\}$. Moreover, the set G' defined by

$$G' = \{(1, \mathsf{red}), (2, \mathsf{red}), (2, \mathsf{blue}), (3, \mathsf{green})\}$$

does not define the graph of a function $\{1,2,3\} \to \{\mathsf{red}, \mathsf{green}, \mathsf{blue}\}$, since there is not a *unique* element of the form $(2,y)$ in G'—rather, there are two of them!

Exercise 3.1.11

For each of the following specifications of sets X, Y, G, determine whether or not G is the graph of a function from X to Y.

(a) $X = \mathbb{R}, Y = \mathbb{R}, G = \{(a, a^2) \mid a \in \mathbb{R}\}$;

(b) $X = \mathbb{R}, Y = \mathbb{R}, G = \{(a^2, a) \mid a \in \mathbb{R}\}$;

(c) $X = [0,\infty), Y = [0,\infty), G = \{(a^2, a) \mid a \in [0,\infty)\}$;

(d) $X = \mathbb{Q}, Y = \mathbb{Q}, G = \{(x,y) \in \mathbb{Q} \times \mathbb{Q} \mid xy = 1\}$.

(e) $X = \mathbb{Q}, Y = \mathbb{Q}, G = \{(a,a) \mid a \in \mathbb{Z}\}$;

Aside

In light of Theorem 3.1.9, some people choose to define functions $X \to Y$ as particular subsets of $X \times Y$—that is, they identify functions with their graphs. This is particularly useful when studying the logical foundations of mathematics. We avoid this practice here, because it is not conceptually necessary, and it would preclude other possible ways of encoding functions.

There is a class of functions called *identity functions* that, despite being very simple, are so important that we will give them a numbered definition!

Definition 3.1.12

Let X be a set. The **identity function** on X is the function $\mathrm{id}_X : X \to X$ (LaTeX code: `\mathrm{id}_X`) defined by $\mathrm{id}_X(x) = x$ for all $x \in X$.

You should convince yourself that the specification of id_X given in Definition 3.1.12 is well-defined.

Another interesting example of a function is the *empty function*, which is useful in coming up with counterexamples and proving combinatorial identities (see Section 8.1).

Definition 3.1.13

Let X be a set. The **empty function** with codomain X is the (unique!) function $\varnothing \to X$. It has no values, since there are no elements of its domain.

Again, you should convince yourself that this specification is well-defined. Conceptually, convincing yourself of this is not easy; but writing down the proof of well-definedness is extremely easy—you will find that there is simply nothing to prove!

Example 3.1.14

Define $f : \mathbb{R} \to \mathbb{R}$ by the equation $f(x)^2 = x$ for all $x \in \mathbb{R}$. This is not well-defined for a few reasons. First, if $x < 0$ then there is no real number y such that $y^2 = x$, so for $x < 0$ there are no possible values of $f(x)$ in the codomain of f, so *existence* fails. Second, if $x > 0$ then there are in fact *two* real numbers y such that $y^2 = x$, namely the positive square root $\sqrt{x}$ and the negative square root $-\sqrt{x}$. The specification of f does not indicate which of these values to take, so *uniqueness* fails.

Notice that the function $r : (0,\infty) \to (0,\infty)$ from Example 3.1.2 *is* (well-)defined by the equation $r(x)^2 = x$ for all $x \in (0,\infty)$. This illustrates why it is very important to specify the domain and codomain when defining a function.

Exercise 3.1.15

Which of the following specifications of functions are well-defined?

(a) $g : \mathbb{Q} \to \mathbb{Q}$ defined by the equation $(x+1)g(x) = 1$ for all $x \in \mathbb{Q}$;

(b) $h : \mathbb{N} \to \mathbb{Q}$ defined by $(x+1)h(x) = 1$ for all $x \in \mathbb{N}$;

(c) $k : \mathbb{N} \to \mathbb{N}$ defined by $(x+1)k(x) = 1$ for all $x \in \mathbb{N}$;

(d) $\ell : \mathbb{N} \to \mathbb{N}$ defined by $\ell(x) = \ell(x)$ for all $x \in \mathbb{N}$.

Exercise 3.1.16

Find a condition on sets X and Y such that the specification of a function $i : X \cup Y \to \{0,1\}$ given by

$$i(z) = \begin{cases} 0 & \text{if } z \in X \\ 1 & \text{if } z \in Y \end{cases}$$

to be well-defined.

Composition of functions

In our section on sets, we talked about various operations that can be performed on sets—union, intersection, and so on. There are also operations on functions, by far the most important of which is *composition*. To understand how composition works, let's revisit the algorithmically defined function $M : \mathbb{Q} \to \mathbb{Q}$ from page 110:

$$\text{multiply by 2} \to \text{add 5} \to \text{square the result} \to \text{divide by 6}$$

The function M is, in some sense, a *sequence* of functions, performed one-by-one until the desired result is reached. This is precisely *composition of functions*.

Definition 3.1.17

Let $f : X \to Y$ and $g : Y \to Z$ be functions. The **composite** of f and g is the function $g \circ f : X \to Z$ (LaTeX code: `g \circ f`) (read 'g composed with f' or 'g after f' or even just '$g\ f$') defined by

$$(g \circ f)(x) = g(f(x)) \text{ for all } x \in X$$

Intuitively, $g \circ f$ is the function resulting from first applying f, and then applying g, to the given input.

Common error

Function composition is in some sense written 'backwards': in the expression $g \circ f$, the function which is applied *first* is written *last*—there is a good reason for this: the argument to the function is written after the function! However, this mis-match often trips students up on their first exposure to function composition, so be careful!

Example 3.1.18

The function M from page 110 can be defined as the composite

$$M = ((k \circ h) \circ g) \circ f$$

where

- $f : \mathbb{Q} \to \mathbb{Q}$ is defined by $f(x) = 2x$ for all $x \in \mathbb{Q}$;
- $g : \mathbb{Q} \to \mathbb{Q}$ is defined by $g(x) = x + 5$ for all $x \in \mathbb{Q}$;
- $h : \mathbb{Q} \to \mathbb{Q}$ is defined by $h(x) = x^2$ for all $x \in \mathbb{Q}$;
- $k : \mathbb{Q} \to \mathbb{Q}$ is defined by $k(x) = \frac{x}{6}$ for all $x \in \mathbb{Q}$.

Exercise 3.1.19

Let $f, g, h, k : \mathbb{Q} \to \mathbb{Q}$ be as in Example 3.1.18. Compute equations defining the following composites:

(a) $f \circ g$;

(b) $g \circ f$;

(c) $((f \circ g) \circ h) \circ k$;

(d) $f \circ (g \circ (h \circ k))$;

(e) $(g \circ g) \circ (g \circ g)$.

Example 3.1.20

Let $f : X \to Y$ be any function. Then

$$\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X$$

To see this, let $x \in X$. Then

$$\begin{aligned}
(\mathrm{id}_Y \circ f)(x) &= \mathrm{id}_Y(f(x)) && \text{by definition of composition} \\
&= f(x) && \text{by definition of } \mathrm{id}_Y \\
&= f(\mathrm{id}_X(x)) && \text{by definition of } \mathrm{id}_X \\
&= (f \circ \mathrm{id}_X)(x) && \text{by definition of composition}
\end{aligned}$$

Equality of the three functions in question follows.

Exercise 3.1.21

Prove that composition of functions is *associative*, that is, if $f : X \to Y$, $g : Y \to Z$ and $h : Z \to W$ are functions, then

$$h \circ (g \circ f) = (h \circ g) \circ f : X \to W$$

As a consequence of associativity, when we want to compose more than two functions, it doesn't matter what order we compose the functions in. As such, we can just write $h \circ g \circ f$.

Exercise 3.1.22

Let $f : X \to Y$ and $g : Z \to W$ be functions, and suppose that $Y \subsetneq Z$. Note that there is a function $h : X \to W$ defined by $h(x) = g(f(x))$ for all $x \in X$. Write h as a composite of functions involving f and g.

Characteristic functions

A class of functions that are particularly useful for proving results about sets are *characteristic functions*.

Definition 3.1.23

Let X be a set and let $A \subseteq X$. The **characteristic function** of A is the function $\chi_A : X \to [2]$ (LaTeX code: `\chi_{A}`), where $[2] = \{0, 1\}$ (see Definition 0.25), defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Example 3.1.24

Consider the subset $A = \{1,3,5\} \subseteq [6]$. Then the values of the characteristic function $\chi_A : [6] \to [2]$ are given by

$$\chi_A(0) = 0 \quad \chi_A(1) = 1 \quad \chi_A(2) = 0$$
$$\chi_A(3) = 1 \quad \chi_A(4) = 0 \quad \chi_A(5) = 1$$

Theorem 3.1.25

Let X be a set and let $A, B \subseteq X$. Then $A = B$ if and only if $\chi_A = \chi_B$.

Proof

- ($\Rightarrow$) Assume $A = B$ and let $x \in X$. Then

$$\begin{aligned} \chi_A(x) = 1 &\Leftrightarrow x \in A && \text{by definition of } \chi_A \\ &\Leftrightarrow x \in B && \text{since } A = B \\ &\Leftrightarrow \chi_B(x) = 1 && \text{by definition of } \chi_B \end{aligned}$$

  Likewise $\chi_A(x) = 0$ if and only if $\chi_B(x) = 0$, so that $\chi_A = \chi_B$ by function extensionality.

- ($\Leftarrow$) Assume $\chi_A = \chi_B$ and let $x \in A$. Then

$$\begin{aligned} x \in A &\Leftrightarrow \chi_A(x) = 1 && \text{by definition of } \chi_A \\ &\Leftrightarrow \chi_B(x) = 1 && \text{since } \chi_A = \chi_B \\ &\Leftrightarrow x \in B && \text{by definition of } \chi_B \end{aligned}$$

  so $A = B$ by set extensionality. □

Strategy 3.1.26 · *Proving set identities using characteristic functions*

In order to prove that two subsets A and B of a set X are equal, it suffices to prove that $\chi_A = \chi_B$.

Theorem 3.1.27 · *Characteristic function identities*

Let X be a set and let $A, B \subseteq X$. Then

(x) $\chi_{A\cap B}(x) = \chi_A(x)\chi_B(x)$ for all $x \in X$;

(x) $\chi_{A\cup B}(x) = \chi_A(x) + \chi_B(x) - \chi_A(x)\chi_B(x)$ for all $x \in X$;

(x) $\chi_{X\setminus A}(x) = 1 - \chi_A(x)$ for all $x \in X$.

Proof of (x)

Let $x \in X$. Since the only values that $\chi_A(x)$ and $\chi_B(x)$ can take are 0 and 1, we have

$$\chi_A(x)\chi_B(x) = \begin{cases} 1 & \text{if } \chi_A(x) = 1 \text{ and } \chi_B(x) = 1 \\ 0 & \text{otherwise} \end{cases}$$

But $\chi_A(x) = 1$ if and only if $x \in A$ and $\chi_B(x) = 1$ if and only if $x \in B$, so that

$$\chi_A(x)\chi_B(x) = \begin{cases} 1 & \text{if } x \in A \cap B \\ 0 & \text{if } x \notin A \cap B \end{cases}$$

This is exactly to say that $\chi_A(x)\chi_B(x) = \chi_{A\cap B}(x)$, as required. □

Exercise 3.1.28
Prove parts (b) and (c) of Theorem 3.1.27.

Theorem 3.1.27 can be used in conjunction with Strategy 3.1.26 to prove set theoretic identities using their characteristic functions.

Example 3.1.29
In Example 2.2.27 we proved that $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ for all sets X, Y and Z. We prove this again using characteristic functions, considering X, Y and Z as subsets of a universal set $\mathscr{U}$.

So let $a \in U$. Then

$$\begin{aligned}
&\chi_{X\cap(Y\cup Z)}(a) \\
&= \chi_X(a)\chi_{Y\cup Z}(a) && \text{by Theorem 3.1.27(a)} \\
&= \chi_X(a)(\chi_Y(a) + \chi_Z(a) - \chi_Y(a)\chi_Z(a)) && \text{by Theorem 3.1.27(b)} \\
&= \chi_X(a)\chi_Y(a) + \chi_X(a)\chi_Z(a) - \chi_X(a)\chi_Y(a)\chi_Z(a) && \text{rearranging} \\
&= \chi_X(a)\chi_Y(a) + \chi_X(a)\chi_Z(a) - \chi_X(a)^2\chi_Y(a)\chi_Z(a) && \text{since } \chi_X(a)^2 = \chi_X(a) \\
&= \chi_{X\cap Y}(a) + \chi_{X\cap Z}(a) - \chi_{X\cap Y}(a)\chi_{X\cap Z}(a) && \text{by Theorem 3.1.27(a)} \\
&= \chi_{(X\cap Y)\cup(X\cap Z)}(a) && \text{by Theorem 3.1.27(b)}
\end{aligned}$$

Using Strategy 3.1.26, it follows that $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$.

Exercise 3.1.30
Use characteristic functions to prove de Morgan's laws for pairwise unions and intersections (Theorem 2.2.44).

Images and preimages

Definition 3.1.31
Let $f : X \to Y$ be a function and let $U \subseteq X$. The **image of** U **under** f is the subset $f[U] \subseteq Y$ (also written $f_*(U)$ (LaTeX code: `f_*`) or even just $f(U)$) is defined by

$$f[U] = \{f(x) \mid x \in U\} = \{y \in Y \mid \exists x \in U, y = f(x)\}$$

That is, $f[U]$ is the set of values that the function f takes when given inputs from U.

The **image of** f is the image of the entire domain, i.e. the set $f[X]$.

Example 3.1.32

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. The image of f is the set $[0,\infty)$ of all nonnegative real numbers. Let's prove this:

- $(f[\mathbb{R}] \subseteq [0,\infty))$. Let $y \in f[\mathbb{R}]$. Then $y = x^2$ for some $x \in \mathbb{R}$. But $x^2 \geqslant 0$, so we must have $y \in [0,\infty)$, as required.
- $([0,\infty) \subseteq f[\mathbb{R}])$. Let $y \in [0,\infty)$. Then $\sqrt{y} \in \mathbb{R}$, and $y = (\sqrt{y})^2 = f(\sqrt{y})$. Hence $y \in f[\mathbb{R}]$, as required.

We have shown by double containment that $f[\mathbb{R}] = [0,\infty)$.

Exercise 3.1.33

For each of the following functions f and subsets U of their domain, describe the image $f[U]$.

(a) $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = 3n$, with $U = \mathbb{N}$;

(b) $f : X \to X \times X$ (where X is any set) defined by $f(x) = (x,x)$ with $U = X$;

(c) $f : \{a,b,c\} \to \{1,2,3\}$ defined by $f(a) = 1$, $f(b) = 3$ and $f(c) = 1$, with $U = \{a,b,c\}$.

Exercise 3.1.34

Prove that $f[\varnothing] = \varnothing$ for all functions f.

Example 3.1.35

Let $f : X \to Y$ be a function and let $U,V \subseteq X$. Then $f[U \cap V] \subseteq f[U] \cap f[V]$. To see this, let $y \in f[U \cap V]$. Then $y = f(x)$ for some $x \in U \cap V$. By definition of intersection, $x \in U$ and $x \in V$. Since $x \in U$ and $y = f(x)$, we have $y \in f[U]$; likewise, since $x \in V$, we have $y \in f[V]$. But then by definition of intersection, we have $y \in f[U] \cap f[V]$.

Exercise 3.1.36

Let $f : X \to Y$ be a function and let $U,V \subseteq X$. We saw in Example 3.1.35 that $f[U \cap V] \subseteq f[U] \cap f[V]$. Determine which of the following is true, and for each, provide a proof of its truth or falsity:

(a) $f[U] \cap f[V] \subseteq f[U \cap V]$;

(b) $f[U \cup V] \subseteq f[U] \cup f[V]$;

(c) $f[U] \cup f[V] \subseteq f[U \cup V]$.

Definition 3.1.37

Let $f : X \to Y$ be a function and let $V \subseteq Y$. The **preimage of V under f** is the subset $f^{-1}[V]$ (LaTeX code: `f^{-1}`) (also written $f^*(V)$ (LaTeX code: `f^*`), or just $f^{-1}(V)$) is defined by

$$f^{-1}[V] = \{x \in X \mid f(x) \in V\} = \{x \in X \mid \exists y \in V, f(x) = y\}$$

That is, $f^{-1}[V]$ is the set of all the elements of its domain X that the function f sends to elements of V.

Example 3.1.38

Let $f : \mathbb{Z} \to \mathbb{Z}$ be the function defined by $f(x) = x^2$ for all $x \in X$. Then

- $f^{-1}[\{1,4,9\}] = \{-3,-2,-1,1,2,3\}$;
- $f^{-1}[\{1,2,3,4,5,6,7,8,9\}] = \{-3,-2,-1,1,2,3\}$ too, since the other elements of $[9]$ are not perfect squares, and hence not of the form $f(x)$ for $x \in \mathbb{Z}$;
- $f^{-1}[\mathbb{N}] = \mathbb{Z}$, since for any $x \in \mathbb{Z}$ we have $f(x) \geqslant 0$, so that $f(x) \in \mathbb{N}$.

Example 3.1.39

Let $f : X \to Y$ be a function, let $U \subseteq X$ and let $V \subseteq Y$. Then $f[U] \subseteq V$ if and only if $U \subseteq f^{-1}[V]$. The proof is as follows.

($\Rightarrow$). Suppose $f[U] \subseteq V$; we'll prove $U \subseteq f^{-1}[V]$. So fix $x \in U$. Then $f(x) \in f[U]$ by definition of image. But then $f(x) \in V$ by our assumption that $f[U] \subseteq V$, and so $x \in f^{-1}[V]$ by definition of preimage. Since x was arbitrarily chosen from U, it follows that $U \subseteq f^{-1}[V]$.

($\Leftarrow$). Suppose $U \subseteq f^{-1}[V]$; we'll prove $f[U] \subseteq V$. So fix $y \in f[U]$. Then $y = f(x)$ for some $x \in U$ by definition of image. But then $x \in f^{-1}[V]$ by our assumption that $U \subseteq f^{-1}[V]$, and so $f(x) \in V$ by definition of preimage. But $y = f(x)$, so $y \in V$, and since y was arbitrarily chosen, it follows that $f[U] \subseteq V$.

The following exercise demonstrates that preimages interact very nicely with the basic set operations (intersection, union and relative complement):

Exercise 3.1.40

Let $f : X \to Y$ be a function. Prove that $f^{-1}[\varnothing] = \varnothing$ and $f^{-1}[Y] = X$.

Exercise 3.1.41

Let X be a set. Prove that every function $f : X \to \{0,1\}$ is the characteristic function of the subset $f^{-1}[\{1\}] \subseteq X$.

Section 3.2

Injections and surjections

To motivate some of the definitions to come, look at the dots (•) and stars (⋆) below. Are there more dots or more stars?

Pause for a second and think about how you knew the answer to this question.

Indeed, there are more dots than stars. There are a couple of ways to arrive at this conclusion:

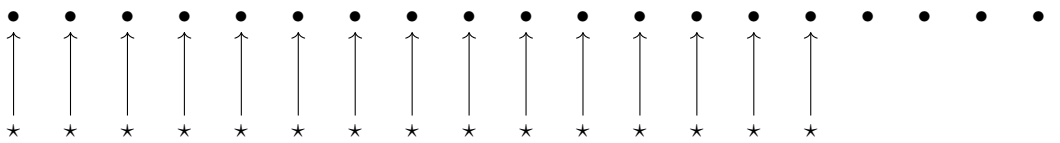
(i) You could count the number of dots, count the number of stars, and then compare the two numbers; or

(ii) You could notice that the dots and the stars are evenly spaced, but that the line of dots is longer than the line of stars.

It is likely that you chose method (ii). In fact, it is likely that you haven't even counted the number of dots or the number of stars yet—and you don't need to! We can conclude that there are more dots than stars by simply pairing up dots with stars—we eventually run out of stars, and there are still dots left over, so there must have been more dots than stars.

Injectivity

One way of formalising this act of pairing up stars with dots mathematically is to define a function $f : S \to D$ from the set S of stars to the set D of dots, where the value of f at each star is the dot that it is paired with. We of course must do this in such a way that each dot is paired with at most one star:



It is a property of this function—called *injectivity*—that allows us to deduce that there are more dots than stars.

Intuitively, a function $f : X \to Y$ is injective if it puts the elements of X in one-to-one correspondence with the elements of a subset of Y—just like how the stars are in one-to-one correspondence with a subset of the dots in the example above.

Definition 3.2.1

A function $f : X \to Y$ is **injective** (or **one-to-one**) if

$$\forall a, b \in X,\ f(a) = f(b) \Rightarrow a = b$$

An injective function is said to be an **injection**.

Strategy 3.2.2 · *Proving a function is injective*

In order to prove that a function $f : X \to Y$ is injective, it suffices to fix $a, b \in X$, assume that $f(a) = f(b)$, and then derive $a = b$.

By contraposition, $f : X \to Y$ being injective is equivalent to saying, for all $a, b \in X$, if $a \neq b$, then $f(a) \neq f(b)$. This is usually less useful for *proving* that a function is injective, but it does provide a good intuition—it says that f sends distinct inputs to distinct outputs.

The following is a very simple example from elementary arithmetic:

Example 3.2.3

Define $f : \mathbb{Z} \to \mathbb{Z}$ by letting $f(x) = 2n + 1$ for all $n \in \mathbb{Z}$. We'll prove that f is injective. Fix $m, n \in \mathbb{Z}$, and assume that $f(m) = f(n)$. By definition of f, we have $2m + 1 = 2n + 1$. Subtracting 1 yields $2m = 2n$, and dividing by 2 yields $m = n$. Hence f is injective.

The following example is slightly more sophisticated.

Proposition 3.2.4

Let $f : X \to Y$ and $g : Y \to Z$ be functions. If f and g are injective, then $g \circ f$ is injective.

Proof

Suppose that f and g are injective and let $a, b \in X$. We need to prove that

$$(g \circ f)(a) = (g \circ f)(b) \quad \Rightarrow \quad a = b$$

So assume $(g \circ f)(a) = (g \circ f)(b)$. By definition of function composition, this implies that $g(f(a)) = g(f(b))$. By injectivity of g, we have $f(a) = f(b)$; and by injectivity of f, we have $a = b$. □

Exercise 3.2.5

Let $f : X \to Y$ and $g : Y \to Z$ be functions. Prove that if $g \circ f$ is injective, then f is injective.

Exercise 3.2.6

Write out what it means to say a function $f : X \to Y$ is *not* injective, and say how you would prove that a given function is not injective. Give an example of a function which is not injective, and use your proof technique to write a proof that it is not injective.

✎ **Exercise 3.2.7**
For each of the following functions, determine whether it is injective or not injective.

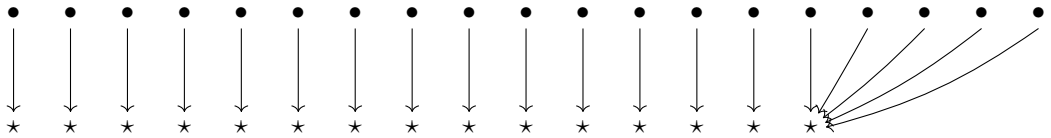
- $f : \mathbb{N} \to \mathbb{Z}$, defined by $f(n) = n^2$ for all $n \in \mathbb{N}$.
- $g : \mathbb{Z} \to \mathbb{N}$, defined by $g(n) = n^2$ for all $n \in \mathbb{Z}$.
- $h : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, defined by $h(x, y, z) = 2^x \cdot 3^y \cdot 5^z$ for all $x, y, z \in \mathbb{N}$.

✎ **Exercise 3.2.8**
Let $a, b \in \mathbb{R}$ with $b \neq 0$, and define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = a + bt$ for all $t \in \mathbb{R}$. Prove that f is injective.

Surjectivity

Let's revisit the rows of dots and stars that we saw earlier. Beforehand, we made our idea that there are more dots than stars formal by proving the existence of an injection $f : S \to D$ from the set S of stars to the set D of dots.

However, we could have drawn the same conclusion instead from defining a function $D \to S$, which in some sense *covers* the stars with dots—that is, every star is paired up with at least one dot.



This property is called *surjectivity*—a function $f : X \to Y$ is surjective if every element of Y is a value of f. This is made precise in Definition 3.2.9.

✦ **Definition 3.2.9**
A function $f : X \to Y$ is **surjective** (or **onto**) if

$$\forall y \in Y,\ \exists x \in X,\ f(x) = y$$

A surjective function is said to be a **surjection**.

❖ **Strategy 3.2.10**
To prove that a function $f : X \to Y$ is surjective, it suffices to take an arbitrary element $y \in Y$ and, in terms of y, find an element $x \in X$ such that $f(x) = y$.

In order to find x, it is often useful to start from the equation $f(x) = y$ and derive some possible values of x. But be careful—in order to complete the proof, it is necessary to verify that the equation $f(x) = y$ is true for the chosen value of x.

Example 3.2.11
Fix $n \in \mathbb{N}$ with $n > 0$, and define a function $r : \mathbb{Z} \to [n]$ by letting $r(a)$ be the remainder of a when divided by n (see Theorem 0.42). This function is surjective, since for each $k \in [n]$ we have $r(k) = k$.

Exercise 3.2.12
For each of the following pairs of sets (X, Y), determine whether the function $f : X \to Y$ defined by $f(x) = 2x + 1$ is surjective.

(a) $X = \mathbb{Z}$ and $Y = \{x \in \mathbb{Z} \mid x \text{ is odd}\}$;

(b) $X = \mathbb{Z}$ and $Y = \mathbb{Z}$;

(c) $X = \mathbb{Q}$ and $Y = \mathbb{Q}$;

(d) $X = \mathbb{R}$ and $Y = \mathbb{R}$.

Exercise 3.2.13
Let $f : X \to Y$ be a function. Find a subset $V \subseteq Y$ and a surjection $g : X \to V$ agreeing with f (that is, such that $g(x) = f(x)$ for all $x \in X$).

Exercise 3.2.14
Let $f : X \to Y$ be a function. Prove that f is surjective if and only if $Y = f[X]$

Exercise 3.2.15
Let $f : X \to Y$ be a function. Prove that there is a set Z and functions

$$p : X \to Z \quad \text{and} \quad i : Z \to Y$$

such that p is surjective, i is injective, and $f = i \circ p$.

Exercise 3.2.16
Let $f : X \to \mathscr{P}(X)$ be a function. By considering the set $B = \{x \in X \mid x \notin f(x)\}$, prove that f is not surjective.

Bijectivity

Bijective functions formalise the idea of putting sets into one-to-one correspondence—each element of one set is paired with exactly one element of another.

Definition 3.2.17
A function $f : X \to Y$ is **bijective** if it is injective and surjective. A bijective function is said to be a **bijection**.

Proof tip
To prove that a function f is bijective, prove that it is injective and surjective.

Example 3.2.18
Let $D \subseteq \mathbb{Q}$ be the set of *dyadic rational numbers*, that is

$$D = \left\{ x \in \mathbb{Q} \;\middle|\; x = \frac{a}{2^n} \text{ for some } a \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\}$$

Let $k \in \mathbb{N}$, and define $f : D \to D$ by $f(x) = \frac{x}{2^k}$. We will prove that f is a bijection.

- (**Injectivity**) Fix $x, y \in D$ and suppose that $f(x) = f(y)$. Then $\frac{x}{2^k} = \frac{y}{2^k}$, so that $x = y$, as required.
- (**Surjectivity**) Fix $y \in D$. We need to find $x \in D$ such that $f(x) = y$. Well certainly if $2^k y \in D$ then we have
$$f(2^k y) = \frac{2^k y}{2^k} = y$$
so it suffices to prove that $2^k y \in D$. Since $y \in D$, we must have $y = \frac{a}{2^n}$ for some $n \in \mathbb{N}$.
  - ◇ If $k \leqslant n$ then $n - k \in \mathbb{N}$ and so $2^k y = \frac{a}{2^{n-k}} \in D$.
  - ◇ If $k > n$ then $k - n > 0$ and $2^k y = 2^{k-n} a \in \mathbb{Z}$; but $\mathbb{Z} \subseteq D$ since if $a \in \mathbb{Z}$ then $a = \frac{a}{2^0}$. So again we have $2^k y \in D$.

  In both cases we have $2^k y \in D$; and $f(2^k y) = y$, so that f is surjective.

Since f is both injective and surjective, it is bijective.

Exercise 3.2.19
Let X be a set. Prove that the identity function $\mathrm{id}_X : X \to X$ is a bijection.

Exercise 3.2.20
Let $f : X \to Y$ and $g : Y \to Z$ be bijections. Prove that $g \circ f$ is a bijection.

Inverses

Our next goal is to characterise injections, surjections and bijections in terms of other functions, called *inverses*.

Recall Definition 3.2.1, which says that a function $f : X \to Y$ is injective if, for all $a, b \in X$, if $f(a) = f(b)$ then $a = b$.

Exercise 3.2.21
Let $f : X \to Y$ be a function. Prove that f is injective if and only if

$$\forall y \in f[X], \exists! x \in X, y = f(x)$$

Thinking back to Section 3.1, you might notice that this means that the logical formula '$y = f(x)$' defines a function $f[X] \to X$—specifically, if f is injective then there is a function $g : f[X] \to X$ which is (well-)defined by specifying $x = g(f(x))$ for all $x \in X$. Thinking of f as an *encoding* function, we then have that g is the corresponding *decoding* function—decoding is possible by injectivity of f. (If f were not injective then distinct elements of X might have the same encoding, in which case we're stuck if we try to decode them!)

Exercise 3.2.22

Define a function $e : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $e(m,n) = 2^m \cdot 3^n$. Prove that e is injective. We can think of e as encoding *pairs* of natural numbers as single natural numbers—for example, the pair $(4,1)$ is encoded as $2^4 \cdot 3^1 = 48$. For each of the following natural numbers k, find the pairs of natural numbers encoded by e as k:

$$1 \qquad 24 \qquad 7776 \qquad 59049 \qquad 396718580736$$

In Exercise 3.2.22, we were able to decode any natural number of the form $2^m \cdot 3^n$ for $m,n \in \mathbb{N}$. This process of decoding yields a function

$$d : \{k \in \mathbb{N} \mid k = 2^m \cdot 3^n \text{ for some } m,n \in \mathbb{N}\} \to \mathbb{N} \times \mathbb{N}$$

What would happen if we tried to decode a natural number not of the form $2^m \cdot 3^n$ for $m,n \in \mathbb{N}$, say 5 or 100? Well… it doesn't really matter! All we need to be true is that $d(e(m,n)) = (m,n)$ for all $(m,n) \in \mathbb{N} \times \mathbb{N}$; the value of d on other natural numbers is irrelevant.

Definition 3.2.23

Let $f : X \to Y$ be a function. A **left inverse** (or **post-inverse**) for f is a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$.

Example 3.2.24

Let $e : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be as in Exercise 3.2.22. Define a function $d : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ by

$$d(k) = \begin{cases} (m,n) & \text{if } k = 2^m \cdot 3^n \text{ for some } m,n \in \mathbb{N} \\ (0,0) & \text{otherwise} \end{cases}$$

Note that d is well-defined by the fundamental theorem of arithmetic (Theorem 7.2.12). Moreover, given $m,n \in \mathbb{N}$, we have

$$d(e(m,n)) = d(2^m \cdot 3^n) = (m,n)$$

and so d is a left inverse for e.

Exercise 3.2.25

Let $f : X \to Y$ be a function. Prove that if f has a left inverse, then f is injective.

Exercise 3.2.25 gives us a new strategy for proving that a function is injective.

Strategy 3.2.26 · *Proving a function is injective by finding a left inverse*
In order to prove that a function $f : X \to Y$ is injective, it suffices to find a function $g : Y \to X$ such that $g(f(x)) = x$ for all $x \in X$.

It would be convenient if the converse to Exercise 3.2.25 were true—and it is, provided that we impose the condition that the domain of the function be inhabited.

Proposition 3.2.27
Let $f : X \to Y$ be a function. If f is injective and X is inhabited, then f has a left inverse.

Proof
Suppose that f is injective and X is inhabited. Fix $x_0 \in X$—note that this element exists since X is inhabited—and define $g : Y \to X$ as follows.

$$g(y) = \begin{cases} x & \text{if } y = f(x) \text{ for some } x \in X \\ x_0 & \text{otherwise} \end{cases}$$

The only part of the specification of g that might cause it to fail to be well-defined is the case when $y = f(x)$ for some $x \in X$. The reason why g is well-defined is precisely injectivity of f: if $y = f(x)$ for some $x \in X$, then the value of $x \in X$ for which $y = f(x)$ is unique. (Indeed, if $a \in X$ satisfied $y = f(a)$, then we'd have $a = x$ by injectivity of f.)

So g is indeed well-defined. To see that g is a left inverse for f, let $x \in X$. Letting $y = f(x)$, we see that y falls into the first case in the specification of g, so that $g(f(x)) = g(y) = a$ for the value of $a \in X$ for which $y = f(a)$—but as noted above, we have $a = x$ by injectivity of f. □

Exercise 3.2.28
Let $f : X \to Y$ be a function with left inverse $g : Y \to X$. Prove that g is a surjection.

We established a relationship between injections and left inverses in Exercise 3.2.25 and Proposition 3.2.27, so it might come as no surprise that there is a relationship between surjections and *right* inverses.

Definition 3.2.29
Let $f : X \to Y$ be a function. A **right inverse** (or **pre-inverse**) for f is a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.

Example 3.2.30
Define $f : \mathbb{R} \to [0, \infty)$ by $f(x) = x^2$. Note that f is surjective, since for each $y \in [0, \infty)$ we have $\sqrt{y} \in \mathbb{R}$ and $f(\sqrt{y}) = y$. However f is not injective; for instance

$$f(-1) = 1 = f(1)$$

Here are three right inverses for f:

- The positive square root function $g : [0,\infty) \to \mathbb{R}$ defined by $g(y) = \sqrt{y}$ for all $y \in [0,\infty)$. Indeed, for each $y \in [0,\infty)$ we have
$$f(g(y)) = f(\sqrt{y}) = (\sqrt{y})^2 = y$$
- The negative square root function $h : [0,\infty) \to \mathbb{R}$ defined by $h(y) = -\sqrt{y}$ for all $y \in [0,\infty)$. Indeed, for each $y \in [0,\infty)$ we have
$$f(h(y)) = f(-\sqrt{y}) = (-\sqrt{y})^2 = y$$
- The function $k : [0,\infty) \to \mathbb{R}$ defined by
$$k(y) = \begin{cases} \sqrt{y} & \text{if } 2n \leqslant y < 2n+1 \text{ for some } n \in \mathbb{N} \\ -\sqrt{y} & \text{otherwise} \end{cases}$$
Note that k is well-defined, and again $f(k(y)) = y$ for all $y \in [0,\infty)$ since no matter what value $k(y)$ takes, it is equal to either $\sqrt{y}$ or $-\sqrt{y}$.

There are many more right inverses for f—in fact, there are infinitely many more!

Exercise 3.2.31
Let $f : X \to Y$ be a function. Prove that if f has a right inverse, then f is surjective.

Strategy 3.2.32 · *Proving a function is surjective by finding a right inverse*
In order to prove that a function $f : X \to Y$ is surjective, it suffices to find a function $g : Y \to X$ such that $f(g(y)) = y$ for all $y \in Y$.

Interlude: the axiom of choice

It would be convenient if the converse to Exercise 3.2.31 were true—that is, if $f : X \to Y$ is surjective, then it has a right inverse. Let's examine what a proof of this fact would entail. The fact that $f : X \to Y$ is surjective can be expressed as
$$\forall y \in Y, \exists x \in X, f(x) = y$$
A right inverse would be a function $g : Y \to X$, so by Definition 3.1.1, it must satisfy the following condition
$$\forall y \in Y, \exists! x \in X, g(y) = x$$

The temptation is therefore to construct $g : Y \to X$ as follows. Let $y \in Y$. By definition of surjectivity, there exists some $x \in X$ such that $f(x) = y$—define $g(y)$ to be such an element x. Then we have $f(g(y)) = f(x) = y$, as required.

There is an extremely subtle—but important—issue with this construction.

By choosing $g(y)$ to be a fixed element of X such that $f(x) = y$, we are assuming ahead of time that there is a mechanism for choosing, for each $y \in Y$, a unique element of $f^{-1}[\{y\}]$ to be the value of

$g(y)$. In other words we are assuming that some statement $R(x,y)$ satisfies the property

$$\forall y \in Y, \exists! x \in X, [x \in f^{-1}[\{y\}] \wedge R(x,y)]$$

But by Definition 3.1.1, this assumption is saying exactly that there exists a function $Y \to X$ that associates to each $y \in Y$ an element $x \in X$ such that $f(x) = y$.

To state this in plainer terms: we tried to prove that there exists a right inverse for f by assuming that there exists a right inverse for f. Evidently, this is not a valid proof strategy.

Surprisingly, it turns out that neither the assumption that every surjection has a right inverse, nor the assumption that there exists a surjection with no right inverse, leads to a contradiction. As such, the assertion that every surjection has a right inverse is *provably unprovable*, although the proof that it is unprovable is far beyond the scope of this textbook.

Nonetheless, the construction of a right inverse that we gave above didn't *feel* like we were abusing the fabric of mathematics and logic.

The essence of the proof is that if a statement of the form $\forall a \in A, \exists b \in B, p(a,b)$ is true, then we should be able to define a function $h : A \to B$ such that $p(a,h(a))$ is true for all $a \in A$: the function h 'chooses' for each $a \in A$ a particular element $b = h(a) \in B$ such that $p(a,b)$ is true.

What makes this possible is to *axiom of choice*, stated precisely below.

Axiom 3.2.33 · *Axiom of choice*
Let $\{X_i \mid i \in I\}$ be a family of inhabited sets. Then there is a function $h : I \to \bigcup_{i \in I} X_i$ such that $h(i) \in X_i$ for each $i \in I$.

There are reasons to keep track of the axiom of choice:

- The axiom of choice is perhaps the *strangest* assumption that we make—most of the other axioms that we have stated have been 'evidently true', but this is not the case for the axiom of choice;
- There are fields of mathematics which require the translation of results about sets into results about other kinds of objects—knowing whether the axiom of choice is necessary to prove a result tells us whether this is possible;
- The axiom of choice is highly non-constructive: if a proof of a result that does not use the axiom of choice is available, it usually provides more information than a proof of the same result that does use the axiom of choice.

With this in mind, when we need to invoke the axiom of choice to prove a result, we will mark the result with the letters **AC**. This can be freely ignored on first reading, but readers may find it useful when using this book as a reference at a later date.

Proposition$^{\text{AC}}$ **3.2.34**
Let X and Y be sets and let $p(x,y)$ be a logical formula with free variables $x \in X$ and $y \in Y$. If $\forall x \in X, \forall y \in Y, p(x,y)$ is true, then there exists a function $h : X \to Y$ such that $\forall x \in X, p(x,h(x))$ is true.

Proof
For each $a \in X$, define $Y_a = \{b \in Y \mid p(a,b)\}$. Note that Y_a is inhabited for each $a \in X$ by the assumption that $\forall x \in X, \exists y \in Y, p(x,y)$ is true. Since $Y_a \subseteq Y$ for each $a \in X$, by the axiom of choice there exists a function $h : X \to Y$ such that $h(a) \in Y_a$ for all $a \in X$. But then $p(a,h(a))$ is true for each $a \in X$ by definition of the sets Y_a. □

In light of Proposition 3.2.34, the axiom of choice manifests itself in proofs as the following proof strategy.

Strategy^{AC} 3.2.35 · *Making choices*
If an assumption in a proof has the form $\forall x \in X, \exists y \in Y, p(x,y)$, then we may make a choice, for each $a \in A$, of a particular element $b = b_a \in B$ for which $p(a,b)$ is true.

Back to inverses

We now return to the converse of Exercise 3.2.31.

Proposition^{AC} 3.2.36
Every surjection has a right inverse.

Proof
Let $f : X \to Y$ be a surjection, and define $g : Y \to X$ as follows. Given $y \in Y$, define $g(y)$ to be a particular choice of $x \in X$ such that $f(x) = y$—note that there exists such an element $x \in X$ since f is surjective, so g exists by Strategy 3.2.35. But then by definition of g we have $f(g(y)) = y$ for all $y \in Y$, so that g is a surjection. □

It seems logical that we might be able to classify bijections as being those functions which have a left inverse and a right inverse. We can actually say something stronger—the left and right inverse can be taken to be the same function! (In fact, Proposition 3.2.42 establishes that they are necessarily the same function.)

Definition 3.2.37
Let $f : X \to Y$ be a function. A (**two-sided**) **inverse** for f is a function $g : Y \to X$ which is both a left inverse and a right inverse for f.

It is customary to simply say ‘inverse’ rather than ‘two-sided inverse’.

Example 3.2.38
Let D be the set of dyadic rational numbers, as defined in Example 3.2.18. There, we defined a function $f : D \to D$ defined by $f(x) = \frac{x}{2^k}$ for all $x \in D$, where k is some fixed natural number. We find an inverse for f.

Define $g : D \to D$ by $g(x) = 2^k x$. Then

- g is a left inverse for f. To see this, note that for all $x \in D$ we have

$$g(f(x)) = g\left(\frac{x}{2^k}\right) = 2^k \cdot \frac{x}{2^k} = x$$

- g is a right inverse for f. To see this, note that for all $y \in D$ we have

$$f(g(y)) = f(2^k y) = \frac{2^k y}{2^k} = y$$

Since g is a left inverse for f and a right inverse for f, it is a two-sided inverse for f.

Exercise 3.2.39

The following functions have two-sided inverses. For each, find its inverse and prove that it is indeed an inverse.

(a) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \dfrac{2x+1}{3}$ for all $x \in \mathbb{R}$;

(b) $g : \mathscr{P}(\mathbb{N}) \to \mathscr{P}(\mathbb{N})$ defined by $g(X) = \mathbb{N} \setminus X$ for all $X \in \mathscr{P}(\mathbb{N})$;

(c) $h : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $h(m,n) = 2^m(2n+1) - 1$ for all $m, n \in \mathbb{N}$.

In light of the correspondences between injections and left inverses, and surjections and right inverses, it may be unsurprising that there is a correspondence between *bijections* and *two-sided inverses*.

Exercise 3.2.40

Let $f : X \to Y$ be a function. Then f is bijective if and only if f has an inverse.

Strategy 3.2.41 · *Proving a function is bijective by finding an inverse*

In order to prove that a function $f : X \to Y$ is bijective, it suffices to find a function $g : Y \to X$ such that $g(f(x)) = x$ for all $x \in X$ and $f(g(y)) = y$ for all $y \in Y$.

When proving a function $f : X \to Y$ is bijective by finding an inverse $g : Y \to X$, it is important to check that g is *both* a left inverse *and* a right inverse for f. If you only prove that g is a left inverse for f, for example, then you have only proved that f is injective!

It turns out that if a function has both a left and a right inverse, then they must be equal. This is the content of the following proposition.

Proposition 3.2.42

Let $f : X \to Y$ be a function and suppose $\ell : Y \to X$ is a left inverse for f and $r : Y \to X$ is a right inverse for f. Then $\ell = r$.

Proof
The proof is deceptively simple. Let $y \in Y$. Then:

$$\begin{aligned} \ell(y) &= \ell(f(r(y))) && \text{since } f \circ r = \mathrm{id}_Y \\ &= r(y) && \text{since } \ell \circ f = \mathrm{id}_X \end{aligned}$$

so $\ell = r$ by function extensionality. □

There is some intuition behind why the left and right inverses of a function $f : X \to Y$ should be equal if they both exist.

- A left inverse $\ell : Y \to X$ exists only if f is injective. It looks at each element $y \in Y$ and, if it is in the image of f, returns the (unique) value $x \in X$ for which $f(x) = y$.
- A right inverse $r : Y \to X$ exists only if f is surjective. It looks at each element $y \in Y$ and picks out one of the (possibly many) values $x \in X$ for which $f(x) = y$.

When f is a bijection, every element of Y is in the image of f (by surjectivity), and is a value of f at a unique element of X (by injectivity), and so the left and right inverses are forced to return the same value on each input—hence they are equal.

It follows from Proposition 3.2.42 that, for any function $f : X \to Y$, any two inverses for f are equal—that is, every bijective function has a *unique* inverse!

Notation 3.2.43
Let $f : X \to Y$ be a function. Write $f^{-1} : Y \to X$ to denote the (unique) inverse for f, if it exists.

Proposition 3.2.44
Let $f : X \to Y$ be a bijection. A function $g : Y \to X$ is a left inverse for f if and only if it is a right inverse for f.

Proof
We will prove the two directions separately.

- ($\Rightarrow$) Suppose $g : Y \to X$ is a left inverse for f—that is, $g(f(x)) = x$ for all $x \in X$. We prove that $f(g(y)) = y$ for all $y \in Y$, thus establishing that g is a right inverse for f. So let $y \in Y$. Since f is a bijection, it is in particular a surjection, so there exists $x \in X$ such that $y = f(x)$. But then

$$\begin{aligned} f(g(y)) &= f(g(f(x))) && \text{since } y = f(x) \\ &= f(x) && \text{since } g(f(x)) = x \\ &= y && \text{since } y = f(x) \end{aligned}$$

  So indeed g is a right inverse for f.

- ($\Leftarrow$) Suppose $g : Y \to X$ is a right inverse for f—that is, $f(g(y)) = y$ for all $y \in Y$. We prove that $g(f(x)) = x$ for all $x \in X$, thus establishing that g is a left inverse for f. So let $x \in X$. Letting $y = f(x)$, we have $f(g(y)) = y$ since g is a right inverse for f. This says precisely that $f(g(f(x))) = f(x)$, since $y = f(x)$. By injectivity of f, we have $g(f(x)) = x$, as required. □

✎ **Exercise 3.2.45**
Let $f : X \to Y$ be a bijection. Prove that $f^{-1} : Y \to X$ is a bijection.

✎ **Exercise 3.2.46**
Let $f : X \to Y$ and $g : Y \to Z$ be bijections. Prove that $g \circ f : X \to Z$ is a bijection, and write an expression for its inverse in terms of f^{-1} and g^{-1}.

✎ **Exercise 3.2.47**
Let $f : X \to A$ and $g : Y \to B$ be bijections. Prove that there is a bijection $X \times Y \to A \times B$, and describe its inverse.

At the beginning of this section we motivated the definitions of injections, surjections and bijections by using them to compare two quantities (of dots and stars)—however, as you might have noticed, we have not yet actually proved that this intuition aligns with reality. For example, how do we know that if there is an injection $f : X \to Y$ then Y has at least as many elements as X?

Answering this seemingly simple question is surprisingly difficult and has different answers depending on whether the sets involved are finite or infinite. In fact, the words 'finite', 'infinite' and 'size' are themselves defined in terms of injections, surjections and bijections! We therefore leave this task to future sections.

In Section 6.1, we define what it means for a set to be finite and what the size of a finite set is (Definition 6.1.1), and then prove that the sizes of finite sets can be compared by finding an injection, surjection or bijection between them Theorem 6.1.6.

Comparing the sizes of infinite sets, and even defining what 'size' means for infinite sets, is another can of worms entirely and leads to some fascinating mathematics. For example, we can prove some counterintuitive results, such as the set $\mathbb{N}$ of natural numbers and the set $\mathbb{Q}$ of rational numbers have the same size. The journey down this rabbit hole begins in Chapter 10.

Section 3.E

Chapter 3 exercises

3.E.1. For each of the following equations, determine whether there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that, for all $x, y \in \mathbb{R}$, the equation holds if and only if $y = f(x)$.

(a) $x + y = 1$

(b) $x^2 + y^2 = 1$

(c) $x = 0$

(d) $y = 0$

(e) $(1 + x^2)y = 1$

(f) $(1 - x^2)y = 0$

3.E.2. Let X be a set. Prove that

$$\forall a \in X, \exists! U \in \mathscr{P}(X), (a \in U \wedge \exists! x \in X, x \in U)$$

Give an explicit description of the function $X \to \mathscr{P}(X)$ that is suggested by this logical formula.

3.E.3. Show that there is only one function whose codomain is empty. What is its domain?

Definition 3.E.4

A function $f : \mathbb{R} \to \mathbb{R}$ is **even** if $f(-x) = f(x)$ for all $x \in \mathbb{R}$, and it is **odd** if $f(-x) = -f(x)$ for all $x \in \mathbb{R}$.

3.E.5. Let $n \in \mathbb{N}$. Prove that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^n$ for all $x \in \mathbb{R}$ is even if and only if n is even, and odd if and only if n is odd.

3.E.6. Prove that there is a unique function $f : \mathbb{R} \to \mathbb{R}$ that is both even and odd.

3.E.7. Let $U \subseteq \mathbb{R}$, and let $\chi_U : \mathbb{R} \to [2]$ be the indicator function of U.

(a) Prove that χ_U is an even function if and only if $U = \{-u \mid u \in U\}$;

(b) Prove that χ_U is an odd function if and only if $\mathbb{R} \setminus U = \{-u \mid u \in U\}$.

3.E.8. Prove that for every function $f : \mathbb{R} \to \mathbb{R}$, there is a unique even function $g : \mathbb{R} \to \mathbb{R}$ and a unique odd function $h : \mathbb{R} \to \mathbb{R}$ such that $f(x) = g(x) + h(x)$ for all $x \in \mathbb{R}$.

3.E.9. Let $\{\theta_n : [n] \to [n] \mid n \in \mathbb{N}\}$ be a family of functions such that $f \circ \theta_m = \theta_n \circ f$ for all $f : [m] \to [n]$. Prove that $\theta_n = \mathrm{id}_{[n]}$ for all $n \in \mathbb{N}$.

3.E.10. Let X be a set and let $U, V \subseteq X$. Describe the indicator function $\chi_{U \triangle V}$ of the symmetric difference of U and V (see Definition 2.E.8) in terms of χ_U and χ_V.

3.E.11. [This question assumes some familiarity with integral calculus.] A lie that is commonly told to calculus students is that

$$\int \frac{1}{x}\, dx = \log(|x|) + c$$

where log denotes the natural logarithm function and c is an arbitrary real constant. Prove that it is actually the case that

$$\int \frac{1}{x}\, dx = \log(|x|) + a\chi_U(x) + b\chi_V(x)$$

where a and b are arbitrary real constants, and U and V are particular (inhabited) subsets of $\mathbb{R} \setminus \{0\}$ that you should determine.

Images and preimages

In Questions 3.E.12 to 3.E.15, find the image $f[U]$ of the subset U of the domain of the function f described in the question.

3.E.12. $f : \mathbb{R} \to \mathbb{R}$; $f(x) = \sqrt{1+x^2}$ for all $x \in \mathbb{R}$; $U = \mathbb{R}$.

3.E.13. $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$; $f(a,b) = a + 2b$ for all $(a,b) \in \mathbb{Z} \times \mathbb{Z}$; $U = \{1\} \times \mathbb{Z}$.

3.E.14. $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$; $f(0) = \varnothing$ and $f(n+1) = f(n) \cup \{n\}$ for all $n \in \mathbb{N}$; $U = \mathbb{N}$.

3.E.15. $f : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$ (where $\mathbb{R}^{\mathbb{R}}$ is the set of all functions $\mathbb{R} \to \mathbb{R}$); $f(h)(x) = h(|x|)$ for all $h \in \mathbb{R}^{\mathbb{R}}$ and all $x \in \mathbb{R}$; $U = \mathbb{R}^{\mathbb{R}}$.

In Questions 3.E.16 to 3.E.18, find the preimage $f^{-1}[V]$ of the subset V of the codomain of the function f described in the question.

3.E.16. $f : \mathbb{R} \to \mathbb{R}$; $f(x) = \sqrt{1+x^2}$ for all $x \in \mathbb{R}$; $V = (-5,5]$.

3.E.17. $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$; $f(a,b) = a + 2b$ for all $(a,b) \in \mathbb{Z} \times \mathbb{Z}$; $V = \{n \in \mathbb{Z} \mid n \text{ is odd}\}$.

3.E.18. $f : \mathscr{P}(\mathbb{N}) \times \mathscr{P}(\mathbb{N}) \to \mathscr{P}(\mathbb{N})$; $f(A,B) = A \cap B$ for all $(A,B) \in \mathscr{P}(\mathbb{N}) \times \mathscr{P}(\mathbb{N})$; $V = \{\varnothing\}$.

3.E.19. Let $f : X \to Y$ be a function. For each of the following statements, either prove it is true or find a counterexample.

(a) $U \subseteq f^{-1}[f[U]]$ for all $U \subseteq X$;

(b) $f^{-1}[f[U]] \subseteq U$ for all $U \subseteq X$;

(c) $V \subseteq f[f^{-1}[V]]$ for all $V \subseteq Y$;

(d) $f[f^{-1}[V]] \subseteq V$ for all $V \subseteq Y$.

3.E.20. Let $f : X \to Y$ be a function, let A be a set, and let $p : X \to A$ and $i : A \to Y$ be functions such that the following conditions hold:

(i) i is injective;

(ii) $i \circ p = f$; and

(iii) If $q : X \to B$ and $j : B \to Y$ are functions such that j is injective and $j \circ q = f$, then there is a unique function $u : A \to B$ such that $j \circ u = i$.

Prove that there is a unique bijection $v : A \to f[X]$ such that $i(a) = v(a)$ for all $a \in f[X]$.

3.E.21. Let $f : X \to Y$ be a function and let $U, V \subseteq Y$. Prove that:

(a) $f^{-1}[U \cap V] = f^{-1}[U] \cap f^{-1}[V]$;

(b) $f^{-1}[U \cup V] = f^{-1}[U] \cup f^{-1}[V]$; and

(c) $f^{-1}[Y \setminus U] = X \setminus f^{-1}[U]$.

Thus preimages preserve the basic set operations.

3.E.22. Let $f : X \to Y$ and $g : Y \to Z$ be functions.

(a) Prove that $(g \circ f)[U] = g[f[U]]$ for all $U \subseteq X$;

(b) Prove that $(g \circ f)^{-1}[W] = f^{-1}[g^{-1}[W]]$ for all $W \subseteq Z$.

3.E.23. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Define

$$S = \{(x,y,z) \in X \times Y \times Z \mid y = f(x) \text{ and } z = g(y)\}$$

and let $p : S \to X \times Z$ be the function defined by $p(x,y,z) = (x,z)$ for all $(x,y,z) \in S$. Prove that $p[S] = \mathrm{Gr}(g \circ f)$.

Injections, surjections and bijections

3.E.24. (a) Prove that, for all functions $f : X \to Y$ and $g : Y \to Z$, if $g \circ f$ is bijective, then f is injective and g is surjective.

(b) Find an example of a function $f : X \to Y$ and a function $g : Y \to Z$ such that $g \circ f$ is bijective, f is not surjective and g is not injective.

3.E.25. For each of the following pairs (U,V) of subsets of $\mathbb{R}$, determine whether the specification '$f(x) = x^2 - 4x + 7$ for all $x \in U$' defines a function $f : U \to V$ and, if it does, determine whether f is injective and whether f is surjective.

(a) $U = \mathbb{R}$ and $V = \mathbb{R}$;

(b) $U = (1,4)$ and $V = [3,7)$;

(c) $U = [3,4)$ and $V = [4,7)$;

(d) $U = (3,4]$ and $V = [4,7)$;

(e) $U = [2,\infty)$ and $V = [3,\infty)$;

(f) $U = [2,\infty)$ and $V = \mathbb{R}$.

3.E.26. For each of the following pairs of sets X and Y, find (with proof) a bijection $f : X \to Y$.

(a) $X = \mathbb{Z}$ and $Y = \mathbb{N}$;

(b) $X = \mathbb{R}$ and $Y = (-1,1)$;

(c) $X = [0,1]$ and $Y = (0,1)$;

(d) $X = [a,b]$ and $Y = (c,d)$, where $a,b,c,d \in \mathbb{R}$ with $a < b$ and $c < d$.

3.E.27. Prove that the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $f(a,b) = \binom{a+b+1}{2} + b$ for all $(a,b) \in \mathbb{N} \times \mathbb{N}$ is a bijection.

3.E.28. Let $e : X \to X$ be a function such that $e \circ e = e$. Prove that there exist a set Y and functions $f : X \to Y$ and $g : Y \to X$ such that $g \circ f = e$ and $f \circ g = \mathrm{id}_Y$.

True–False questions

In Questions 3.E.29 to 3.E.36, determine (with proof) whether the statement is true or false.

3.E.29. The set $G = \{(x,y) \in \mathbb{R} \times \mathbb{R} \mid x^2 = y^2\}$ is the graph of a function.

3.E.30. The set $G = \{(x,y) \in \mathbb{Z} \times \mathbb{N} \mid x^2 = y^2\}$ is the graph of a function.

3.E.31. Every function with empty domain has an empty codomain.

3.E.32. Every function with empty codomain has an empty domain.

3.E.33. The image of a function is a subset of its domain.

3.E.34. Given a function $f : X \to Y$, the assignment $V \mapsto f^{-1}[V]$ defines a function $\mathscr{P}(Y) \to \mathscr{P}(X)$.

3.E.35. Let f, g and h be composable functions. If $h \circ g \circ f$ is injective, then g is injective.

3.E.36. Every left inverse is surjective and every right inverse is injective.

Always–Sometimes–Never questions

In Questions 3.E.37 to 3.E.48, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

3.E.37. Let $f : X \to Y$ be a function and let $U \subseteq X$. Then there is a function $g : U \to Y$ defined by $g(x) = f(x)$ for all $x \in U$.

3.E.38. Let $f : X \to Y$ be a function and let $V \subseteq Y$. Then there is a function $g : X \to V$ defined by $g(x) = f(x)$ for all $x \in X$.

3.E.39. Let X be a set. Then there is a unique function $X \to \{0\}$.

3.E.40. Let X be a set. Then there is a unique function $X \to \varnothing$.

3.E.41. Let X and Y be sets and let $G \subseteq X \times Y$. Then G is the graph of a function $f : X \to Y$.

3.E.42. Let $f : \{1,2,3\} \to \{1,2\}$, let $G = \mathrm{Gr}(f)$ and let $G' = \{(y,x) \mid (x,y) \in G\}$. Then G' is the graph of a function $\{1,2\} \to \{1,2,3\}$.

3.E.43. Let $f : X \to Y$ be a function and let $U \subseteq X$ be inhabited. Then $f[U]$ is inhabited.

3.E.44. Let $f : X \to Y$ be a function and let $V \subseteq Y$ be inhabited. Then $f^{-1}[V]$ is inhabited.

3.E.45. Let $f : \{1,2\} \to \{1,2,3\}$ be a function. Then f is injective.

3.E.46. Let $f : \{1,2\} \to \{1,2,3\}$ be a function. Then f is surjective.

3.E.47. Let $a,b,c,d \in \mathbb{R}$ and define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x,y) = (ax+by, cx+dy)$ for all $(x,y) \in \mathbb{R}^2$. Then f is injective if and only if f is surjective.

3.E.48. Let $U,V \subseteq \mathbb{R}$ and suppose that $x^2 \in V$ for all $x \in U$. The function $f : U \to V$ defined by $f(x) = x^2$ is injective.

Chapter 4

Mathematical induction

Section 4.1

Peano's axioms

The purpose of this section is to forget everything we think we know about the natural numbers, and reconstruct our former knowledge (and more!) using the following fundamental property:

Every natural number can be obtained in a unique way by starting from zero and adding one some finite number of times.

This is slightly imprecise—it is not clear what is meant by 'adding one some finite number of times', for example. Worse still, we are going to define what 'finite' means in terms of natural numbers in Section 6.1, so we'd better not refer to finiteness in our definition of natural numbers!

The following definition captures precisely the properties that we need in order to characterise the idea of $\mathbb{N}$ that we have in our minds. To begin with, $\mathbb{N}$ should be a set. Whatever the elements of this set $\mathbb{N}$ actually *are*, we will think about them as being natural numbers. One of the elements, in particular, should play the role of the natural number 0—this will be the *zero element* $z \in \mathbb{N}$; and there should be a notion of 'adding one'—this will be the *successor function* $s : \mathbb{N} \to \mathbb{N}$. Thus given an element $n \in \mathbb{N}$, though of as a natural number, we think about the element $s(n)$ as the natural number '$n+1$'. Note that this is strictly for the purposes of intuition: we will define '$+$' and '1' in terms of z and s, not vice versa.

Definition 4.1.1

A **notion of natural numbers** is a set $\mathbb{N}$, together with an element $z \in \mathbb{N}$, called a **zero element**, and a function $s : \mathbb{N} \to \mathbb{N}$ called a **successor function**, satisfying the following properties:

(i) $z \notin s[\mathbb{N}]$; that is, $z \neq s(n)$ for any $n \in \mathbb{N}$.

(ii) s is injective; that is, for all $m, n \in \mathbb{N}$, if $s(m) = s(n)$, then $m = n$.

(iii) $\mathbb{N}$ is generated by z and s; that is, for all sets X, if $z \in X$ and for all $n \in \mathbb{N}$ we have $n \in X \Rightarrow s(n) \in X$, then $\mathbb{N} \subseteq X$.

The properties (i), (ii) and (iii) are called **Peano's axioms**.

Note that Definition 4.1.1 does not specify what $\mathbb{N}$, z and s actually are; it just specifies the properties that they must satisfy. It turns out that it doesn't really matter what notion of natural numbers we use, since any two notions are essentially the same. We will not worry about the specifics here—that task is left to Section B.2: a particular notion of natural numbers is defined in Construction B.2.5, and the fact that all notions of natural numbers are 'essentially the same' is made precise and proved in Theorem B.2.8.

We can define all the concepts involving natural numbers that we are familiar with, and prove all the properties that we take for granted, just from the element $z \in \mathbb{N}$ and the successor function $s : \mathbb{N} \to \mathbb{N}$.

For instance, we define '0' to mean z, define '1' to mean $s(z)$, define '2' to mean $s(s(z))$, and so on. Thus '12' is defined to mean

$$s(s(s(s(s(s(s(s(s(s(s(s(z))))))))))))$$

From now on, then, let's write 0 instead of z for the zero element of $\mathbb{N}$. It would be nice if we could write '$n+1$' instead of $s(n)$, but we must first define what '$+$' means. In order to do this, we need a way of defining expressions involving natural numbers; this is what the *recursion theorem* allows us to do.

Theorem 4.1.2 · *Recursion theorem*

Let X be a set. For all $a \in X$ and all $h : \mathbb{N} \times X \to X$, there is a unique function $f : \mathbb{N} \to X$ such that $f(0) = a$ and $f(s(n)) = h(n, f(n))$ for all $n \in \mathbb{N}$.

Proof

Let $a \in X$ and $h : \mathbb{N} \times X \to X$. We prove existence and uniqueness of f separately.

- Define $f : \mathbb{N} \to X$ by specifying $f(0) = a$ and $f(s(n)) = h(n, f(n))$. Since h is a function and s is injective, existence and uniqueness of $x \in X$ such that $f(n) = x$ is guaranteed, provided that $f(n)$ is defined, so we need only verify totality.

  So let $D = \{n \in \mathbb{N} \mid f(n) \text{ is defined}\}$. Then:

  - $0 \in D$, since $f(0)$ is defined to be equal to a.
  - Let $n \in \mathbb{N}$ and suppose $n \in D$. Then $f(n)$ is defined and $f(s(n)) = h(n, f(n))$, so that $f(s(n))$ is defined, and hence $s(n) \in D$.

  By condition (iii) of Definition 4.1.1, we have $\mathbb{N} \subseteq D$, so that $f(n)$ is defined for all $n \in \mathbb{N}$, as required.

- To see that f is unique, suppose $g : \mathbb{N} \to X$ were another function such that $g(0) = a$ and $g(s(n)) = h(n, g(n))$ for all $n \in \mathbb{N}$.

  To see that $f = g$, let $E = \{n \in \mathbb{N} \mid f(n) = g(n)\}$. Then

  - $0 \in E$, since $f(0) = a = g(0)$.
  - Let $n \in \mathbb{N}$ and suppose that $n \in E$. Then $f(n) = g(n)$, and so

    $$f(s(n)) = h(n, f(n)) = h(n, g(n)) = g(s(n))$$

    and so $s(n) \in E$.

  Again, condition (iii) of Definition 4.1.1 is satisfied, so that $\mathbb{N} \subseteq E$. It follows that $f(n) = g(n)$ for all $n \in \mathbb{N}$, and so $f = g$.

Thus we have established the existence and uniqueness of a function $f : \mathbb{N} \to X$ such that $f(0) = a$ and $f(s(n)) = h(n, f(n))$ for all $n \in \mathbb{N}$. □

The recursion theorem allows us to define expressions involving natural numbers *by recursion*; this is Strategy 4.1.3.

Strategy 4.1.3 · *Definition by recursion*

In order to specify a function $f : \mathbb{N} \to X$, it suffices to define $f(0)$ and, for given $n \in \mathbb{N}$, assume that $f(n)$ has been defined, and define $f(s(n))$ in terms of n and $f(n)$.

Example 4.1.4

We can use recursion to define addition on the natural numbers as follows.

For fixed $m \in \mathbb{N}$, we can define a function $\text{add}_m : \mathbb{N} \to \mathbb{N}$ by recursion by:

$$\text{add}_m(0) = m \quad \text{and} \quad \text{add}_m(s(n)) = s(\text{add}_m(n)) \text{ for all } n \in \mathbb{N}$$

In more familiar notation, for $m, n \in \mathbb{N}$, define the expression '$m+n$' to mean $\text{add}_m(n)$. Another way of expressing the recursive definition of $\text{add}_m(n)$ is to say that, for each $m \in \mathbb{N}$, we are defining $m+n$ by recursion on n as follows:

$$m+0 = m \quad \text{and} \quad m+s(n) = s(m+n) \text{ for all } n \in \mathbb{N}$$

We can use the recursive definition of addition to prove familiar equations between numbers. The following proposition is a proof that $2+2=4$. This may seem silly, but notice that the expression '$2+2=4$' is actually shorthand for the following:

$$\text{add}_{s(s(0))}(s(s(0))) = s(s(s(s(0))))$$

We must therefore be careful to apply the definitions in its proof.

Proposition 4.1.5

$2+2=4$

Proof

We use the recursive definition of addition.

$$\begin{aligned}
2+2 &= 2+s(1) && \text{since } 2 = s(1) \\
&= s(2+1) && \text{by definition of } + \\
&= s(2+s(0)) && \text{since } 1 = s(0) \\
&= s(s(2+0)) && \text{by definition of } + \\
&= s(s(2)) && \text{by definition of } + \\
&= s(3) && \text{since } 3 = s(2) \\
&= 4 && \text{since } 4 = s(3)
\end{aligned}$$

as required. □

The following result allows us to drop the notation '$s(n)$' and just write '$n+1$' instead.

Proposition 4.1.6

For all $n \in \mathbb{N}$, we have $s(n) = n+1$.

Proof

Let $n \in \mathbb{N}$. Then by the recursive definition of addition we have

$$n+1 = n+s(0) = s(n+0) = s(n)$$

as required. □

In light of Proposition 4.1.6, we will now abandon the notation $s(n)$, and write $n+1$ instead.

We can define the arithmetic operations of multiplication and exponentiation by recursion, too.

Example 4.1.7

Fix $m \in \mathbb{N}$. Define $m \cdot n$ for all $n \in \mathbb{N}$ by recursion on n as follows:

$$m \cdot 0 = 0 \quad \text{and} \quad m \cdot (n+1) = (m \cdot n) + m \text{ for all } n \in \mathbb{N}$$

Formally, what we have done is define a function $\text{mult}_m : \mathbb{N} \to \mathbb{N}$ recursively by $\text{mult}_m(z) = z$ and $\text{mult}_m(s(n)) = \text{add}_{\text{mult}_m(n)}(m)$ for all $n \in \mathbb{N}$. But the definition we provided is easier to understand.

Proposition 4.1.8

$2 \cdot 2 = 4$

Proof

We use the recursive definitions of addition and recursion.

$$\begin{aligned}
2 \cdot 2 &= 2 \cdot (1+1) && \text{since } 2 = 1+1 \\
&= (2 \cdot 1) + 2 && \text{by definition of } \cdot \\
&= (2 \cdot (0+1)) + 2 && \text{since } 1 = 0+1 \\
&= ((2 \cdot 0) + 2) + 2 && \text{by definition of } \cdot \\
&= (0+2) + 2 && \text{by definition of } \cdot \\
&= (0 + (1+1)) + 2 && \text{since } 2 = 1+1 \\
&= ((0+1)+1) + 2 && \text{by definition of } + \\
&= (1+1) + 2 && \text{since } 1 = 0+1 \\
&= 2+2 && \text{since } 2 = 1+1 \\
&= 4 && \text{by Proposition 4.1.5}
\end{aligned}$$

as required. □

Exercise 4.1.9

Given $m \in \mathbb{N}$, define m^n for all $n \in \mathbb{N}$ by recursion on n, and prove that $2^2 = 4$ using the recursive definitions of exponentiation, multiplication and addition.

We could spend the rest of our lives doing long computations involving recursively defined arithmetic operations, so at this point we will stop, and return to taking for granted the facts that we know about arithmetic operations.

There are, however, a few more notions that we need to define by recursion so that we can use them in our proofs later.

Definition 4.1.10
Given $n \in \mathbb{N}$, the **sum** of n real numbers $a_1, a_2, \ldots, a_n$ is the real number $\sum_{k=1}^{n} a_k$ defined by recursion on $n \in \mathbb{N}$ by

$$\sum_{k=1}^{0} a_k = 0 \quad \text{and} \quad \sum_{k=1}^{n+1} a_k = \left(\sum_{k=1}^{n} a_k \right) + a_{n+1} \text{ for all } n \in \mathbb{N}$$

Definition 4.1.11
Given $n \in \mathbb{N}$, the **product** of n real numbers $a_1, a_2, \ldots, a_n$ is the real number $\prod_{k=1}^{n} a_k$ defined by recursion on $n \in \mathbb{N}$ by

$$\prod_{k=1}^{0} a_k = 1 \quad \text{and} \quad \prod_{k=1}^{n+1} a_k = \left(\prod_{k=1}^{n} a_k \right) \cdot a_{n+1} \text{ for all } n \in \mathbb{N}$$

Example 4.1.12
Let $x_i = i^2$ for each $i \in \mathbb{N}$. Then

$$\sum_{i=1}^{5} x_i = 1 + 4 + 9 + 16 + 25 = 55$$

and

$$\prod_{i=1}^{5} x_i = 1 \cdot 4 \cdot 9 \cdot 16 \cdot 25 = 14400$$

Exercise 4.1.13
Let $x_1, x_2 \in \mathbb{R}$. Working strictly from the definitions of indexed sum and indexed product, prove that

$$\sum_{i=1}^{2} x_i = x_1 + x_2 \quad \text{and} \quad \prod_{i=1}^{2} x_i = x_1 \cdot x_2$$

Binomials and factorials

Definition 4.1.14 · *to be redefined in Definition 8.1.10*
Let $n \in \mathbb{N}$. The **factorial** of n, written $n!$, is defined recursively by

$$0! = 1 \quad \text{and} \quad (n+1)! = (n+1) \cdot n! \text{ for all } n \geqslant 0$$

Put another way, we have

$$n! = \prod_{i=1}^{n} i$$

for all $n \in \mathbb{N}$—recall Definition 4.1.11 to see why these definitions are really just two ways of wording the same thing.

Definition 4.1.15 · *to be redefined in Definition 8.1.4*

Let $n, k \in \mathbb{N}$. The **binomial coefficient** $\binom{n}{k}$ (LaTeX code: `\binom{n}{k}`) (read 'n choose k') is defined by recursion on n *and* on k by

$$\binom{n}{0} = 1, \quad \binom{0}{k+1} = 0, \quad \binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$$

This definition gives rise to an algorithm for computing binomial coefficients: they fit into a diagram known as **Pascal's triangle**, with each binomial coefficient computed as the sum of the two lying above it (with zeroes omitted):

$$\begin{array}{ccccccccccccc}
 & & & & & \binom{0}{0} & & & & & & & \\
 & & & & \binom{1}{0} & & \binom{1}{1} & & & & & & \\
 & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & & \\
 & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & & \\
 & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & & \\
\binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & & &
\end{array}
\quad = \quad
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & &
\end{array}$$

Exercise 4.1.16

Write down the next two rows of Pascal's triangle.

Section 4.2

Weak induction

Just as recursion exploited the structure of the natural numbers to *define expressions* involving natural numbers, induction exploits the very same structure to *prove results* about natural numbers.

Theorem 4.2.1 · *Weak induction principle*

Let $p(n)$ be logical formula with free variable $n \in \mathbb{N}$, and let $n_0 \in \mathbb{N}$. If

(i) $p(n_0)$ is true; and

(ii) For all $n \geqslant n_0$, if $p(n)$ is true, then $p(n+1)$ is true;

then $p(n)$ is true for all $n \geqslant n_0$.

Proof

Define $X = \{n \in \mathbb{N} \mid p(n_0+n) \text{ is true}\}$; that is, given a natural number n, we have $n \in X$ if and only if $p(n_0+n)$ is true. Then

- $0 \in X$, since $n_0+0 = n_0$ and $p(n_0)$ is true by (i).
- Let $n \in \mathbb{N}$ and assume $n \in X$. Then $p(n_0+n)$ is true. Since $n_0+n \geqslant n_0$ and $p(n_0+n)$ is true, we have $p(n_0+n+1)$ is true by (ii). But then $n+1 \in X$.

So by Definition 4.1.1(iii) we have $\mathbb{N} \subseteq X$. Hence $p(n_0+n)$ is true for all $n \in \mathbb{N}$. But this is equivalent to saying that $p(n)$ is true for all $n \geqslant n_0$. □

Strategy 4.2.2 · *Proof by (weak) induction*

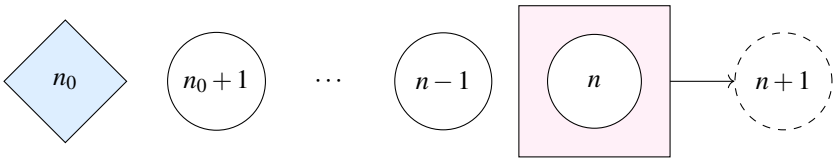
In order to prove a proposition of the form $\forall n \geqslant n_0,\ p(n)$, it suffices to prove that:

- $p(n_0)$ is true; and
- For all $n \geqslant n_0$, if $p(n)$ is true, then $p(n+1)$ is true.

Some terminology has evolved for proofs by induction:

- The proof of $p(n_0)$ is called the **base case**;
- The proof of $\forall n \geqslant n_0,\ (p(n) \Rightarrow p(n+1))$ is called the **induction step**;
- In the induction step, the assumption $p(n)$ is called the **induction hypothesis**;
- In the induction step, the proposition $p(n+1)$ is called the **induction goal**.

The following diagram illustrates the weak induction principle.



To interpret this diagram:

- The shaded diamond represents the base case $p(n_0)$;
- The square represents the induction hypothesis $p(n)$;
- The dashed circle represents the induction goal $p(n+1)$;
- The arrow represents the implication we must prove in the induction step.

We will use analogous diagrams to illustrate the other induction principles in this section.

Proposition 4.2.3
Let $n \in \mathbb{N}$. Then $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$

Proof
We proceed by induction on $n \geqslant 0$.

- (**Base case**) We need to prove $\displaystyle\sum_{k=1}^{0} k = \frac{0(0+1)}{2}$.

  This is true, since $\dfrac{0(0+1)}{2} = 0$, and $\displaystyle\sum_{k=1}^{0} k = 0$ by Definition 4.1.10.

- (**Induction step**) Let $n \geqslant 0$ and suppose that $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$; this is the induction hypothesis.

  We need to prove that $\displaystyle\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}$; this is the induction goal.

  We proceed by calculation:

$$\begin{aligned}
\sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + (n+1) && \text{by Definition 4.1.10} \\
&= \frac{n(n+1)}{2} + (n+1) && \text{by induction hypothesis} \\
&= (n+1)\left(\frac{n}{2} + 1\right) && \text{factorising} \\
&= \frac{(n+1)(n+2)}{2} && \text{rearranging}
\end{aligned}$$

The result follows by induction. □

Before moving on, let's reflect on the proof of Proposition 4.2.3 to highlight some effective ways of writing a proof by induction.

- We began the proof by indicating that it was a proof by induction. While it is clear in this section that most proofs will be by induction, that will not always be the case, so it is good practice to indicate the proof strategy at hand.
- The base case and induction step are clearly labelled in the proof. This is not strictly *necessary* from a mathematical perspective, but it helps the reader to navigate the proof and to identify what the goal is at each step.
- We began the induction step by writing, 'Let $n \geqslant n_0$ and suppose that [*... induction hypothesis goes here...*]'. This is typically how your induction step should begin, since the proposition being proved in the induction step is of the form $\forall n \geqslant n_0, (p(n) \Rightarrow \cdots)$.
- Before proving anything in the base case or induction step, we wrote out what it was that we were trying to prove in that part of the proof. This is helpful because it helps to remind us (and the person reading the proof) what we are aiming to achieve.

Look out for these features in the proof of the next proposition, which is also by induction on $n \geqslant 0$.

Proposition 4.2.4
The natural number $n^3 - n$ is divisible by 3 for all $n \in \mathbb{N}$.

Proof
We proceed by induction on $n \geqslant 0$.

- (**Base case**) We need to prove that $0^3 - 0$ is divisible by 3. Well
$$0^3 - 0 = 0 = 3 \times 0$$
so $0^3 - 0$ is divisible by 3.
- (**Induction step**) Let $n \in \mathbb{N}$ and suppose that $n^3 - n$ is divisible by 3. Then $n^3 - n = 3k$ for some $k \in \mathbb{Z}$.
We need to prove that $(n+1)^3 - (n+1)$ is divisible by 3; in other words, we need to find some integer ℓ such that
$$(n+1)^3 - (n+1) = 3\ell$$
We proceed by computation.

$$\begin{aligned}
&(n+1)^3 - (n+1) \\
&= (n^3 + 3n^2 + 3n + 1) - n - 1 && \text{expand brackets} \\
&= n^3 - n + 3n^2 + 3n + 1 - 1 && \text{rearrange terms} \\
&= n^3 - n + 3n^2 + 3n && \text{since } 1 - 1 = 0 \\
&= 3k + 3n^2 + 3n && \text{by induction hypothesis} \\
&= 3(k + n^2 + n) && \text{factorise}
\end{aligned}$$

Thus we have expressed $(n+1)^3 - (n+1)$ in the form 3ℓ for some $\ell \in \mathbb{Z}$; specifically, $\ell = k + n^2 + n$.

The result follows by induction. □

Exercise 4.2.5
Prove by induction that $\sum_{k=0}^{n} 2^k = 2^{n+1} - 1$ for all $n \in \mathbb{N}$.

The following proposition has a proof by induction in which the base case is not zero.

Proposition 4.2.6
For all $n \geqslant 4$, we have $3n < 2^n$.

Proof
We proceed by induction on $n \geqslant 4$.

- (**Base case**) $p(4)$ is the statement $3 \cdot 4 < 2^4$. This is true, since $12 < 16$.
- (**Induction step**) Suppose $n \geqslant 4$ and that $3n < 2^n$. We want to prove $3(n+1) < 2^{n+1}$. Well,

$$\begin{aligned} 3(n+1) &= 3n+3 && \text{expanding} \\ &< 2^n + 3 && \text{by induction hypothesis} \\ &< 2^n + 2^4 && \text{since } 3 < 16 = 2^4 \\ &\leqslant 2^n + 2^n && \text{since } n \geqslant 4 \\ &= 2 \cdot 2^n && \text{simplifying} \\ &= 2^{n+1} && \text{simplifying} \end{aligned}$$

So we have proved $3(n+1) < 2^{n+1}$, as required.

The result follows by induction. □

Note that the proof in Proposition 4.2.6 says nothing about the truth or falsity of $p(n)$ for $n = 0, 1, 2, 3$. In order to assert that these cases are false, you need to show them individually; indeed:

- $3 \times 0 = 0$ and $2^0 = 1$, hence $p(0)$ is true;
- $3 \times 1 = 3$ and $2^1 = 2$, hence $p(1)$ is false;
- $3 \times 2 = 6$ and $2^2 = 4$, hence $p(2)$ is false;
- $3 \times 3 = 9$ and $2^3 = 8$, hence $p(3)$ is false.

So we deduce that $p(n)$ is true when $n = 0$ or $n \geqslant 4$, and false when $n \in \{1, 2, 3\}$.

Sometimes a 'proof' by induction might appear to be complete nonsense. The following is a classic example of a 'fail by induction':

Example 4.2.7

The following argument supposedly proves that every horse is the same colour.

- (**Base case**) Suppose there is just one horse. This horse is the same colour as itself, so the base case is immediate.
- (**Induction step**) Suppose that every collection of n horses is the same colour. Let X be a set of $n+1$ horses. Removing the first horse from X, we see that the last n horses are the same colour by the induction hypothesis. Removing the last horse from X, we see that the first n horses are the same colour. Hence all the horses in X are the same colour.

By induction, we're done.

Exercise 4.2.8

Write down the statement $p(n)$ that Example 4.2.7 attempted to prove for all $n \geqslant 1$. Convince yourself that the proof of the base case is correct, then write down—with quantifiers—exactly the proposition that the induction step is meant to prove. Explain why the argument in the induction step failed to prove this proposition.

There are several ways to avoid situations like that of Example 4.2.7 by simply putting more thought into writing the proof. Some tips are:

- State $p(n)$ explicitly. In the statement 'all horses are the same colour' it is not clear exactly what the induction variable is. However, we could have said:

  Let $p(n)$ be the statement 'every set of n horses has the same colour'.

- Refer explicitly to the base case n_0 in the induction step. In Example 4.2.7, our induction hypothesis simply stated 'assume every set of n horses has the same colour'. Had we instead said:

  Let $n \geqslant 1$ and assume every set of n horses has the same colour.

  We may have spotted the error in what was to come.

What follows are a couple more examples of proofs by weak induction.

Proposition 4.2.9

For all $n \in \mathbb{N}$, we have $\displaystyle\sum_{k=0}^{n} k^3 = \left(\sum_{k=0}^{n} k\right)^2$.

Proof

We proved in Proposition 4.2.6 that $\displaystyle\sum_{k=0}^{n} k = \frac{n(n+1)}{2}$ for all $n \in \mathbb{N}$, thus it suffices to prove that

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}$$

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

We proceed by induction on $n \geqslant 0$.

- (**Base case**) We need to prove that $0^3 = \dfrac{0^2(0+1)^2}{4}$. This is true since both sides of the equation are equal to 0.
- (**Induction step**) Fix $n \in \mathbb{N}$ and suppose that $\displaystyle\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}$. We need to prove that $\displaystyle\sum_{k=0}^{n+1} k^3 = \frac{(n+1)^2(n+2)^2}{4}$. This is true since:

$$\begin{aligned}
\sum_{i=0}^{n+1} k^3 &= \sum_{i=0}^{n} k^3 + (n+1)^3 && \text{by definition of sum} \\
&= \frac{n^2(n+1)^2}{4} + (n+1)^3 && \text{by induction hypothesis} \\
&= \frac{n^2(n+1)^2 + 4(n+1)^3}{4} && (\textit{algebra}) \\
&= \frac{(n+1)^2(n^2+4(n+1))}{4} && (\textit{algebra}) \\
&= \frac{(n+1)^2(n+2)^2}{4} && (\textit{algebra})
\end{aligned}$$

By induction, the result follows. □

In the next proposition, we prove the correctness of a well-known formula for the sum of an *arithmetic progression* of real numbers.

Proposition 4.2.10
Let $a, d \in \mathbb{R}$. Then

$$\sum_{k=0}^{n} (a+kd) = \frac{(n+1)(2a+nd)}{2}$$

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

Proof
We proceed by induction on $n \geqslant 0$.

- (**Base case**) We need to prove that $\displaystyle\sum_{k=0}^{0} (a+kd) = \frac{(0+1)(2a+0d)}{2}$. This is true, since

$$\sum_{k=0}^{0} (a+kd) = a+0d = a = \frac{2a}{2} = \frac{1\cdot(2a)}{2} = \frac{(0+1)(2a+0d)}{2}$$

- (**Induction step**) Fix $n \in \mathbb{N}$ and suppose that $\displaystyle\sum_{k=0}^{n} (a+kd) = \frac{(n+1)(2a+nd)}{2}$. We need to prove:

$$\sum_{k=0}^{n+1} (a+kd) = \frac{(n+2)(2a+(n+1)d)}{2}$$

This is true, since

$$\begin{aligned}
&\sum_{k=0}^{n+1}(a+kd)\\
&=\sum_{k=0}^{n}(a+kd)+(a+(n+1)d) && \text{by definition of sum}\\
&=\frac{(n+1)(2a+nd)}{2}+(a+(n+1)d) && \text{by induction hypothesis}\\
&=\frac{(n+1)(2a+nd)+2a+2(n+1)d}{2} && (\textit{algebra})\\
&=\frac{(n+1)\cdot 2a+(n+1)\cdot nd+2a+2(n+1)d}{2} && (\textit{algebra})\\
&=\frac{2a(n+1+1)+(n+1)(nd+2d)}{2} && (\textit{algebra})\\
&=\frac{2a(n+2)+(n+1)(n+2)d}{2} && (\textit{algebra})\\
&=\frac{(n+2)(2a+(n+1)d)}{2} && (\textit{algebra})
\end{aligned}$$

By induction, the result follows. □

The following exercises generalises Exercise 4.2.5 to prove the correctness of a formula for the sum of a *geometric progression* of real numbers.

Exercise 4.2.11

Let $a, r \in \mathbb{R}$ with $r \neq 1$. Then

$$\sum_{k=0}^{n} ar^k = \frac{a(1-r^{n+1})}{1-r}$$

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

So far in this section most of the results that we have proved by induction have concerned numbers, and particularly identities involving natural numbers. But the scope of proof by induction is not limited to such results—the next exercise, for example, concerns sets and bijections.

Exercise 4.2.12

Let $n \in \mathbb{N}$ and for each $k \in [n+1]$ let X_k be a set. Prove by induction on n that there is a bijection

$$\prod_{k\in[n+1]} X_k \to \left(\prod_{k\in[n]} X_k\right) \times X_n.$$

Binomials and factorials

Proof by induction turns out to be a very useful way of proving facts about binomial coefficients $\binom{n}{k}$ and factorials $n!$.

Example 4.2.13
We prove that $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ by induction on n.

- (**Base case**) We need to prove $\binom{0}{0} = 1$ and $2^0 = 1$. These are both true by the definitions of binomial coefficients and exponents.
- (**Induction step**) Fix $n \geqslant 0$ and suppose that

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n$$

We need to prove

$$\sum_{i=0}^{n+1} \binom{n+1}{i} = 2^{n+1}$$

This is true, since

$$\begin{aligned}
&\sum_{i=0}^{n+1} \binom{n+1}{i} \\
&= \binom{n+1}{0} + \sum_{i=1}^{n+1} \binom{n+1}{i} && \text{splitting the sum} \\
&= 1 + \sum_{j=0}^{n} \binom{n+1}{j+1} && \text{letting } j = i-1 \\
&= 1 + \sum_{j=0}^{n} \left(\binom{n}{j} + \binom{n}{j+1} \right) && \text{by Definition 4.1.15} \\
&= 1 + \sum_{j=0}^{n} \binom{n}{j} + \sum_{j=0}^{n} \binom{n}{j+1} && \text{separating the sums}
\end{aligned}$$

Now $\sum_{j=0}^{n} \binom{n}{j} = 2^n$ by the induction hypothesis. Moreover, reindexing the sum using $k = j+1$ yields

$$\sum_{j=0}^{n} \binom{n}{j+1} = \sum_{k=1}^{n+1} \binom{n}{k} = \sum_{k=1}^{n} \binom{n}{k} + \binom{n}{n+1}$$

By the induction hypothesis, we have

$$\sum_{k=1}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{0} = 2^n - 1$$

and $\binom{n}{n+1} = 0$, so that $\sum_{j=0}^{n} \binom{n}{j+1} = 2^n - 1$.

Putting this together, we have

$$\begin{aligned}1+\sum_{j=0}^{n}\binom{n}{j}+\sum_{j=0}^{n}\binom{n}{j+1} &= 1+2^n+(2^n-1)\\ &= 2\cdot 2^n\\ &= 2^{n+1}\end{aligned}$$

so the induction step is finished.

By induction, we're done.

Theorem 4.2.14

Let $n,k\in\mathbb{N}$. Then

$$\binom{n}{k}=\begin{cases}\dfrac{n!}{k!(n-k)!} & \text{if } k\leqslant n\\ 0 & \text{if } k>n\end{cases}$$

Proof

We proceed by induction on n.

- (**Base case**) When $n=0$, we need to prove that $\binom{0}{k}=\frac{0!}{k!(-k)!}$ for all $k\leqslant 0$, and that $\binom{0}{k}=0$ for all $k>0$.

  If $k\leqslant 0$ then $k=0$, since $k\in\mathbb{N}$. Hence we need to prove

  $$\binom{0}{0}=\frac{0!}{0!0!}$$

  But this is true since $\binom{0}{0}=1$ and $\frac{0!}{0!0!}=\frac{1}{1\times 1}=1$.

  If $k>0$ then $\binom{0}{k}=0$ by Definition 4.1.15.

- (**Induction step**) Fix $n\in\mathbb{N}$ and suppose that $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ for all $k\leqslant n$ and $\binom{n}{k}=0$ for all $k>n$.

  We need to prove that, for all $k\leqslant n+1$, we have

  $$\binom{n+1}{k}=\frac{(n+1)!}{k!(n+1-k)!}$$

  and that $\binom{n+1}{k}=0$ for all $k>n+1$.

  So fix $k\in\mathbb{N}$. There are four possible cases: either (i) $k=0$, or (ii) $0<k\leqslant n$, or (iii) $k=n+1$, or (iv) $k>n+1$. In cases (i), (ii) and (iii), we need to prove the factorial formula for $\binom{n+1}{k}$; in case (iv), we need to prove that $\binom{n+1}{k}=0$.

  (i) Suppose $k=0$. Then $\binom{n+1}{0}=1$ by Definition 4.1.15, and

  $$\frac{(n+1)!}{k!(n+1-k)!}=\frac{(n+1)!}{0!(n+1)!}=1$$

  since $0!=1$. So $\binom{n+1}{0}=\frac{(n+1)!}{0!(n+1)!}$.

(ii) If $0 < k \leqslant n$ then $k = \ell + 1$ for some natural number $\ell < n$. Then $\ell + 1 \leqslant n$, so we can use the induction hypothesis to apply factorial formula to both $\binom{n}{\ell}$ and $\binom{n}{\ell+1}$. Hence

$$\begin{aligned}
&\binom{n+1}{k} \\
&= \binom{n+1}{\ell+1} && \text{since } k = \ell+1 \\
&= \binom{n}{\ell} + \binom{n}{\ell+1} && \text{by Definition 4.1.15} \\
&= \frac{n!}{\ell!(n-\ell)!} + \frac{n!}{(\ell+1)!(n-\ell-1)!} && \text{by induction hypothesis}
\end{aligned}$$

Now note that

$$\frac{n!}{\ell!(n-\ell)!} = \frac{n!}{\ell!(n-\ell)!} \cdot \frac{\ell+1}{\ell+1} = \frac{n!}{(\ell+1)!(n-\ell)!} \cdot (\ell+1)$$

and

$$\frac{n!}{(\ell+1)!(n-\ell-1)!} = \frac{n!}{(\ell+1)!(n-\ell-1)!} \cdot \frac{n-\ell}{n-\ell} = \frac{n!}{(\ell+1)!(n-\ell)!} \cdot (n-\ell) \qquad \square$$

Piecing this together, we have

$$\begin{aligned}
&\frac{n!}{\ell!(n-\ell)!} + \frac{n!}{(\ell+1)!(n-\ell-1)!} \\
&= \frac{n!}{(\ell+1)!(n-\ell)!} \cdot [(\ell+1) + (n-\ell)] \\
&= \frac{n!(n+1)}{(\ell+1)!(n-\ell)!} \\
&= \frac{(n+1)!}{(\ell+1)!(n-\ell)!}
\end{aligned}$$

so that $\binom{n+1}{\ell+1} = \frac{(n+1)!}{(\ell+1)!(n-\ell)!}$. Now we're done; indeed,

$$\frac{(n+1)!}{(\ell+1)!(n-\ell)!} = \frac{(n+1)!}{k!(n+1-k)!}$$

since $k = \ell + 1$.

(iii) If $k = n+1$, then

$$\begin{aligned}
\binom{n+1}{k} &= \binom{n+1}{n+1} && \text{since } k = n+1 \\
&= \binom{n}{n} + \binom{n}{n+1} && \text{by Definition 4.1.15} \\
&= \frac{n!}{n!0!} + 0 && \text{by induction hypothesis} \\
&= 1
\end{aligned}$$

and $\frac{(n+1)!}{(n+1)!0!} = 1$, so again the two quantities are equal.

(iv) If $k > n+1$, then $k = \ell + 1$ for some $\ell > n$, and so by Definition 4.1.15 and the induction hypothesis we have

$$\binom{n+1}{k} = \binom{n+1}{\ell+1} \overset{\text{IH}}{=} \binom{n}{\ell} + \binom{n}{\ell+1} = 0+0 = 0$$

On first reading, this proof is long and confusing, especially in the induction step where we are required to split into four cases. We will give a much simpler proof in Section 8.1 (see Theorem 8.1.43), where we prove the statement *combinatorially* by putting the elements of two sets in one-to-one correspondence.

We can use Theorem 4.2.14 to prove useful identities involving binomial coefficients.

Example 4.2.15

Let $n,k,\ell \in \mathbb{N}$ with $\ell \leqslant k \leqslant n$ then

$$\binom{n}{k}\binom{k}{\ell} = \binom{n}{\ell}\binom{n-\ell}{k-\ell}$$

Indeed:

$$\begin{aligned}
&\binom{n}{k}\binom{k}{\ell} \\
&= \frac{n!}{k!(n-k)!} \cdot \frac{k!}{\ell!(k-\ell!)} && \text{by Theorem 4.2.14} \\
&= \frac{n!k!}{k!\ell!(n-k)!(k-\ell)!} && \text{combine fractions} \\
&= \frac{n!}{\ell!(n-k)!(k-\ell)!} && \text{cancel } k! \\
&= \frac{n!(n-\ell)!}{\ell!(n-k)!(k-\ell)!(n-\ell)!} && \text{multiply by } \frac{(n-\ell)!}{(n-\ell)!} \\
&= \frac{n!}{\ell!(n-\ell!)} \cdot \frac{(n-\ell)!}{(k-\ell)!(n-k)!} && \text{separate fractions} \\
&= \frac{n!}{\ell!(n-\ell!)} \cdot \frac{(n-\ell)!}{(k-\ell)!((n-\ell)-(k-\ell))!} && \text{rearranging} \\
&= \binom{n}{\ell}\binom{n-\ell}{k-\ell} && \text{by Theorem 4.2.14}
\end{aligned}$$

Exercise 4.2.16

Prove that $\binom{n}{k} = \binom{n}{n-k}$ for all $n,k \in \mathbb{N}$ with $k \leqslant n$.

A very useful application of binomial coefficients in elementary algebra is to the binomial theorem.

Theorem 4.2.17 · *Binomial theorem*

Let $n \in \mathbb{N}$ and $x, y \in \mathbb{R}$. Then

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Proof

In the case when $y = 0$ we have $y^{n-k} = 0$ for all $k < n$, and so the equation reduces to

$$x^n = x^n y^{n-n}$$

which is true, since $y^0 = 1$. So for the rest of the proof, we will assume that $y \neq 0$.

We will now reduce to the case when $y = 1$; and extend to arbitrary $y \neq 0$ afterwards.

We prove $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$ by induction on n.

- (**Base case**) $(1+x)^0 = 1$ and $\binom{0}{0} x^0 = 1 \cdot 1 = 1$, so the statement is true when $n = 0$.
- (**Induction step**) Fix $n \in \mathbb{N}$ and suppose that

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

We need to show that $(1+x)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} x^k$. Well,

$$
\begin{aligned}
&(1+x)^{n+1} \\
&= (1+x)(1+x)^n && \text{by laws of indices} \\
&= (1+x)\cdot \sum_{k=0}^{n} \binom{n}{k} x^k && \text{by induction hypothesis} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^k + x\cdot \sum_{k=0}^{n} \binom{n}{k} x^k && \text{by expanding } (x+1) \\
&= \sum_{k=0}^{n} \binom{n}{k} x^k + \sum_{k=0}^{n} \binom{n}{k} x^{k+1} && \text{distributing } x \\
&= \sum_{k=0}^{n} \binom{n}{k} x^k + \sum_{k=1}^{n+1} \binom{n}{k-1} x^k && k \to k-1 \text{ in second sum} \\
&= \binom{n}{0} x^0 + \sum_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1} \right) x^k + \binom{n}{n} x^{n+1} && \text{splitting the sums} \\
&= \binom{n}{0} x^0 + \sum_{k=1}^{n} \binom{n+1}{k} x^k + \binom{n}{n} x^{n+1} && \text{by Definition 4.1.15} \\
&= \binom{n+1}{0} x^0 + \sum_{k=1}^{n} \binom{n+1}{k} x^k + \binom{n+1}{n+1} x^{n+1} && \text{see } (*) \text{ below} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k
\end{aligned}
$$

The step labelled $(*)$ holds because

$$\binom{n}{0} = 1 = \binom{n+1}{0} \qquad \text{and} \qquad \binom{n}{n} = 1 = \binom{n+1}{n+1}$$

By induction, we've shown that $(1+x)^n = \sum_{i=0}^{n} \binom{n}{k} x^k$ for all $n \in \mathbb{N}$.

When $y \neq 0$ is not necessarily equal to 1, we have that

$$(x+y)^n = y^n \cdot \left(1 + \frac{x}{y}\right)^n = y^n \cdot \sum_{k=0}^{n} \binom{n}{k} \left(\frac{x}{y}\right)^k = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

The middle equation follows by what we just proved; the leftmost and rightmost equations are simple algebraic rearrangements. □

Example 4.2.18

In Example 4.2.13 we saw that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

This follows quickly from the binomial theorem, since

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot 1^k \cdot 1^{n-k} = \sum_{k=0}^{n} \binom{n}{k}$$

Likewise, in Exercise 4.2.19 you proved that the alternating sum of binomial coefficients is zero; that is, for $n \in \mathbb{N}$, we have

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$$

The proof is greatly simplified by applying the binomial theorem. Indeed, by the binomial theorem, we have

$$0 = 0^n = (-1+1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} = \sum_{k=0}^{n} (-1)^k \binom{n}{k}$$

Both of these identities can be proved much more elegantly, quickly and easily using *enumerative combinatorics*. This will be the topic covered in Section 8.1.

Exercise 4.2.19

Use the binomial theorem to prove that

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$$

Section 4.3

Strong induction

Consider the following example, which we will attempt to prove by induction.

Example 4.3.1

Define a sequence recursively by

$$b_0 = 1 \quad \text{and} \quad b_{n+1} = 1 + \sum_{k=0}^{n} b_k \text{ for all } n \in \mathbb{N}$$

We will attempt to prove by induction that $b_n = 2^n$ for all $n \in \mathbb{N}$.

- (**Base case**) By definition of the sequence we have $b_0 = 1 = 2^0$. So far so good.
- (**Induction step**) Fix $n \in \mathbb{N}$, and suppose that $b_n = 2^n$. We need to show that $b_{n+1} = 2^{n+1}$.

  Well, $b_{n+1} = 1 + \sum_{k=0}^{n} b_k = \ldots$ uh oh.

Here's what went wrong. If we could replace each b_k by 2^k in the sum, then we'd be able to complete the proof. However we cannot justify this substitution: our induction hypothesis only gives us information about b_n, not about a general term b_k for $k < n$.

The *strong* induction principle looks much like the weak induction principle, except that its induction hypothesis is more powerful. Despite its name, strong induction is no stronger than weak induction; the two principles are equivalent. In fact, we'll prove the strong induction principle *by weak induction*!

Theorem 4.3.2 · *Strong induction principle*

Let $p(n)$ be a logical formula with free variable $n \in \mathbb{N}$ and let $n_0 \in \mathbb{N}$. If

(i) $p(n_0)$ is true; and

(ii) For all $n \geqslant n_0$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then $p(n+1)$ is true;

then $p(n)$ is true for all $n \geqslant n_0$.

Proof

For each $n \geqslant n_0$, let $q(n)$ be the assertion that $p(k)$ is true for all $n_0 \leqslant k \leqslant n$.

Notice that $q(n)$ implies $p(n)$ for all $n \geqslant n_0$, since given $n \geqslant n_0$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then in particular $p(k)$ is true when $k = n$.

So it suffices to prove $q(n)$ is true for all $n \geqslant n_0$. We do so by weak induction.

- (**Base case**) $q(n_0)$ is equivalent to $p(n_0)$, since the only natural number k with $n_0 \leqslant k \leqslant n_0$ is n_0 itself; hence $q(n_0)$ is true by condition (i).

- (**Induction step**) Let $n \geqslant n_0$ and suppose $q(n)$ is true. Then $p(k)$ is true for all $n_0 \leqslant k \leqslant n$.

  We need to prove that $q(n+1)$ is true—that is, that $p(k)$ is true for all $n_0 \leqslant k \leqslant n+1$. But we know $p(k)$ is true for all $n_0 \leqslant k \leqslant n$—this is the induction hypothesis—and then $p(n+1)$ is true by condition (ii). So we have that $p(k)$ is true for all $n_0 \leqslant k \leqslant n+1$ after all.

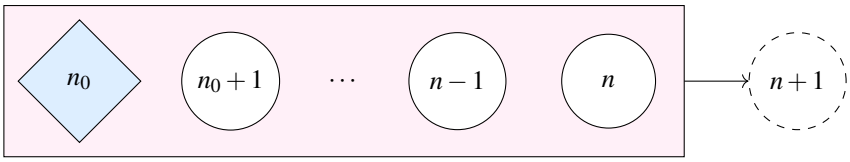
By induction, $q(n)$ is true for all $n \geqslant n_0$. Hence $p(n)$ is true for all $n \geqslant n_0$. □

❖ **Strategy 4.3.3** · *Proof by strong induction*

In order to prove a proposition of the form $\forall n \geqslant n_0$, $p(n)$, it suffices to prove that:

- (**Base case**) $p(n_0)$ is true; and
- (**Induction step**) For all $n \geqslant n_0$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then $p(n+1)$ is true.

Like with weak induction, we can illustrate how strong induction works diagrammatically. The induction hypothesis, represented by the large square, now encompasses $p(k)$ for all $n_0 \leqslant k \leqslant n$, where $p(n_0)$ is the base case.



Observe that the only difference from weak induction is the induction hypothesis.

- **Weak induction step:** Fix $n \geqslant n_0$, assume $p(n)$ is true, derive $p(n+1)$;
- **Strong induction step:** Fix $n \geqslant n_0$, assume $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, derive $p(n+1)$.

We now use strong induction to complete the proof of Example 4.3.1.

✎ **Example 4.3.4** · *Example 4.3.1 revisited*

Define a sequence recursively by

$$b_0 = 1 \quad \text{and} \quad b_{n+1} = 1 + \sum_{k=0}^{n} b_k \text{ for all } n \in \mathbb{N}$$

We will prove by strong induction that $b_n = 2^n$ for all $n \in \mathbb{N}$.

- (**Base case**) By definition of the sequence we have $b_0 = 1 = 2^0$.
- (**Induction step**) Fix $n \in \mathbb{N}$, and suppose that $b_k = 2^k$ for all $k \leqslant n$. We need to show that

$b_{n+1} = 2^{n+1}$. This is true, since

$$\begin{aligned} b_{n+1} &= 1 + \sum_{k=0}^{n} b_k && \text{by the recursive formula for } b_{n+1} \\ &= 1 + \sum_{k=0}^{n} 2^k && \text{by the induction hypothesis} \\ &= 1 + (2^{n+1} - 1) && \text{by Exercise 4.2.5} \\ &= 2^{n+1} \end{aligned}$$

By induction, it follows that $b_n = 2^n$ for all $n \in \mathbb{N}$.

The following theorem adapts the strong induction principle to proofs where we need to refer to a *fixed* number of previous steps in our induction step.

Theorem 4.3.5 · *Strong induction principle (multiple base cases)*
Let $p(n)$ be a logical formula with free variable $n \in \mathbb{N}$ and let $n_0 < n_1 \in \mathbb{N}$. If

(i) $p(n_0), p(n_0+1), \ldots, p(n_1)$ are all true; and

(ii) For all $n \geqslant n_1$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then $p(n+1)$ is true;

then $p(n)$ is true for all $n \geqslant n_0$.

Proof
The fact that $p(n)$ is true for all $n \geqslant n_0$ follows from strong induction. Indeed:

- (**Base case**) $p(n_0)$ is true by (i);
- (**Induction step**) Fix $n \geqslant n_0$ and assume $p(k)$ is true for all $n_0 \leqslant k \leqslant n$. Then:
  - ◇ If $n < n_1$, then $n+1 \leqslant n_1$, so that $p(n)$ is true by (i);
  - ◇ If $n \geqslant n_1$, then $p(n+1)$ is true by (ii).

  In both cases we see that $p(n+1)$ is true, as required.

Thus by strong induction, we have that $p(n)$ is true for all $n \geqslant n_0$. □

Strategy 4.3.6 · *Proof by strong induction with multiple base cases*
In order to prove a statement of the form $\forall n \geqslant n_0,\ p(n)$, it suffices to prove that:

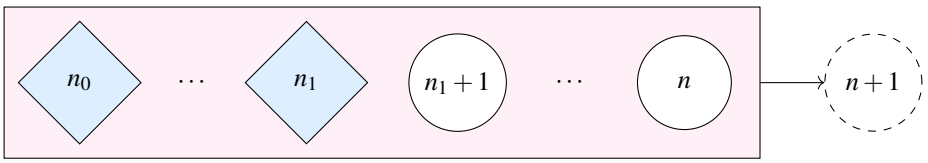
- (**Base cases**) $p(k)$ for all $k \in \{n_0, n_0+1, \ldots, n_1\}$, where $n_1 > n_0$; and
- (**Induction step**) For all $n \geqslant n_1$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then $p(n+1)$ is true.

This kind of strong induction differs from the usual kind in two main ways:

- There are multiple base cases $p(n_0), p(n_0+1), \ldots, p(n_1)$, not just one.

- The induction step refers to both the least base case n_0 and the greatest base case n_1: the variable n in the induction step is taken to be greater than or equal to n_1, while the induction hypothesis assumes $p(k)$ for all $n_0 \leqslant k \leqslant n$.

The following diagram illustrates how strong induction with multiple base cases works.



Note the difference in quantification of variables in the induction step between regular strong induction and strong induction with multiple base cases:

- **One base case.** Fix $n \geqslant \boxed{n_0}$ and assume $p(k)$ is true for all $\boxed{n_0} \leqslant k \leqslant n$.
- **Multiple base cases.** Fix $n \geqslant \boxed{n_1}$ and assume $p(k)$ is true for all $\boxed{n_0} \leqslant k \leqslant n$.

Getting the quantification of the variables n and k in the strong induction step is crucial, since the quantification affects what may be assumed about n and k.

The need for multiple base cases often arises when proving results about recursively defined sequences, where the definition of a general term depends on the values of a fixed number of previous terms.

Example 4.3.7

Define the sequence

$$a_0 = 0, \quad a_1 = 1, \quad a_n = 3a_{n-1} - 2a_{n-2} \text{ for all } n \geqslant 2$$

We find and prove a general formula for a_n in terms of n. Writing out the first few terms in the sequence establishes a pattern that we might attempt to prove:

n	0	1	2	3	4	5	6	7	8
a_n	0	1	3	7	15	31	63	127	255

It appears that $a_n = 2^n - 1$ for all $n \geqslant 0$. We prove this by strong induction, taking the cases $n = 0$ and $n = 1$ as our base cases.

- (**Base cases**) By definition of the sequence, we have:
  - $a_0 = 0 = 2^0 - 1$; and
  - $a_1 = 1 = 2^1 - 1$;

  so the claim is true when $n = 0$ and $n = 1$.
- (**Induction step**) Fix $n \geqslant 1$ and assume that $a_k = 2^k - 1$ for all $0 \leqslant k \leqslant n$. We need to prove that $a_{n+1} = 2^{n+1} - 1$.

Well since $n \geqslant 1$, we have $n+1 \geqslant 2$, so we can apply the recursive formula to a_{n+1}. Thus

$$\begin{aligned}
a_{n+1} &= 3a_n - 2a_{n-1} && \text{by definition of } a_{n+1} \\
&= 3(2^n - 1) - 2(2^{n-1} - 1) && \text{by induction hypothesis} \\
&= 3 \cdot 2^n - 3 - 2 \cdot 2^{n-1} + 2 && \text{expanding} \\
&= 3 \cdot 2^n - 3 - 2^n + 2 && \text{using laws of indices} \\
&= 2 \cdot 2^n - 1 && \text{simplifying} \\
&= 2^{n+1} - 1 && \text{using laws of indices}
\end{aligned}$$

So the result follows by induction.

The following exercises have proofs by strong induction with multiple base cases.

Exercise 4.3.8
Define a sequence recursively by $a_0 = 4$, $a_1 = 9$ and $a_n = 5a_{n-1} - 6a_{n-2}$ for all $n \geqslant 2$. Prove that $a_n = 3 \cdot 2^n + 3^n$ for all $n \in \mathbb{N}$.

Exercise 4.3.9
The *Tribonacci sequence* is the sequence $t_0, t_1, t_2, \ldots$ defined by

$$t_0 = 0, \quad t_1 = 0, \quad t_2 = 1, \quad t_n = t_{n-1} + t_{n-2} + t_{n-3} \text{ for all } n \geqslant 3$$

Prove that $t_n \leqslant 2^{n-3}$ for all $n \geqslant 3$.

Exercise 4.3.10
The *Frobenius coin problem* asks when a given amount of money can be obtained from coins of given denominations. For example, a value of 7 dubloons cannot be obtained using only 3 dubloon and 5 dubloon coins, but for all $n \geqslant 8$, a value of n dubloons *can* be obtained using only 3 dubloon and 5 dubloon coins. Prove this.

Well-ordering principle

The underlying reason why we can perform induction and recursion on the natural numbers is because of the way they are ordered. Specifically, the natural numbers satisfy the *well-ordering principle*: if a set of natural numbers has at least one element, then it has a least element. This sets the natural numbers apart from the other number sets; for example, $\mathbb{Z}$ has no least element, since if $a \in \mathbb{Z}$ then $a - 1 \in \mathbb{Z}$ and $a - 1 < a$.

Theorem 4.3.11 · *Well-ordering principle*
Every inhabited set of natural numbers has a least element.

Proof
We will prove by strong induction that, for all $n \in \mathbb{N}$, every subset $X \subseteq \mathbb{N}$ with $n \in X$ has a least element.

- (**Base case**) Let $X \subseteq \mathbb{N}$ and suppose that $0 \in X$. For all $n \in X$ we have $n \in \mathbb{N}$, since $X \subseteq \mathbb{N}$, and therefore $n \geqslant 0$. But then 0 is the least element of X.
- (**Induction step**) Let $n \in \mathbb{N}$ and suppose that, for all $k \in \mathbb{N}$, every subset $X \subseteq \mathbb{N}$ with $k \in X$ has a least element.

  Let $X \subseteq \mathbb{N}$ and suppose that $n+1 \in X$. One of two cases must occur:

  - ◇ Suppose that for all $k \leqslant n$ we have $k \notin X$. Then $n+1$ is the least element of X, since any smaller element of X would need to be less than or equal to n.
  - ◇ Suppose that there exists $k \leqslant n$ such that $k \in X$. Then by the induction hypothesis, X has a least element.

  In both cases we see that X has a least element, as required.

Finally, note that the fact that we proved by induction implies the statement in the theorem, since an inhabited set of natural numbers is precisely a subset $X \subseteq \mathbb{N}$ such that $n \in X$ for some $n \in \mathbb{N}$.□

We are at last in a position to prove that $\sqrt{2}$ is irrational, a fact that we stated as an assumption in Assumption 0.49. The following proof is a classic application of the well-ordering principle.

Proposition 4.3.12
The number $\sqrt{2}$ is irrational.

To prove Proposition 4.3.12 we will use the following lemma, which uses the well-ordering principle to prove that fractions can be 'cancelled to lowest terms'.

Lemma 4.3.13
Let q be a positive rational number. There is a pair of nonzero natural numbers a, b such that $q = \frac{a}{b}$ and such that the only natural number which divides both a and b is 1.

Proof
First note that we can express q as the ratio of two nonzero natural numbers, since q is a positive rational number. By the well-ordering principle, there is a *least* natural number a such that $q = \frac{a}{b}$ for some positive $b \in \mathbb{N}$.

Suppose that some natural number d other than 1 divides both a and b. Note that $d \neq 0$, since if $d = 0$ then that would imply $a = 0$. Since $d \neq 1$, it follows that $d \geqslant 2$.

Since d divides a and b, there exist natural numbers a', b' such that $a = a'd$ and $b = b'd$. Moreover, $a', b' > 0$ since $a, b, d > 0$. It follows that

$$q = \frac{a}{b} = \frac{a'd}{b'd} = \frac{a'}{b'}$$

But $d \geqslant 2$, and hence

$$a' = \frac{a}{d} \leqslant \frac{a}{2} < a$$

contradicting minimality of a. Hence our assumption that some natural number d other than 1 divides both a and b was false—it follows that the only natural number dividing both a and b is 1. □

We are now ready to prove that $\sqrt{2}$ is irrational.

Proof of Proposition 4.3.12
Suppose $\sqrt{2}$ is rational. Since $\sqrt{2} > 0$, this means that we can write

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

where a and b are both positive natural numbers. By Lemma 4.3.13, we may assume that the only natural number dividing a and b is 1.

Multiplying the equation $\sqrt{2} = \frac{a}{b}$ and squaring yields

$$a^2 = 2b^2$$

Hence a^2 is even. By Proposition 1.1.44, a is even, so we can write $a = 2c$ for some $c > 0$. Then $a^2 = (2c)^2 = 4c^2$, and hence

$$4c^2 = 2b^2$$

Dividing by 2 yields

$$2c^2 = b^2$$

and hence b^2 is even. By Proposition 1.1.44 again, b is even.

But if a and b are both even, the natural number 2 divides both a and b. This contradicts the fact that the only natural number dividing both a and b is 1. Hence our assumption that $\sqrt{2}$ is rational is incorrect, and $\sqrt{2}$ is irrational. □

Writing tip

In the proof of Proposition 4.3.12 we could have separately proved that a^2 being even implies a is even, and that b^2 being even implies b is even. This would have required us to repeat the same proof twice, which is somewhat tedious! Proving auxiliary results separately (as in Lemma 4.3.13) and then quoting them in later theorems can improve the readability of the main proof, particularly when the auxiliary results are particularly technical. Doing so also helps emphasise the important steps.

Exercise 4.3.14

What goes wrong in the proof of Proposition 4.3.12 if we try instead to prove that $\sqrt{4}$ is irrational?

Exercise 4.3.15
Prove that $\sqrt{3}$ is irrational.

The well-ordering principle gives rise to a useful strategy for deriving contradictions, by using the contradictory assumption to build a set of natural numbers with no least element.

Definition 4.3.16
A set $X \subseteq \mathbb{R}$ has **infinite descent** if, for all $x \in X$, there exists $y \in X$ with $y < x$.

The reason for the term 'infinite descent' is that, if an inhabited set $X \subseteq \mathbb{R}$ has infinite descent, then we can construct an *infinite*, *descending* sequence of elements of X. To see this, note that since X is inhabited, there is an element $x_0 \in X$; therefore there exists $x_1 \in X$ with $x_1 < x_0$; and so there exists an element $x_2 \in X$ with $x_2 < x_1$; and so on. In this way, we obtain a sequence $x_0, x_1, x_2, \ldots \in X$ with

$$x_0 > x_1 > x_2 > \cdots$$

Example 4.3.17
There are many examples of subsets $X \subseteq \mathbb{R}$ with infinite descent. For example, $\mathbb{Z}$ has infinite descent, since for all $n \in \mathbb{Z}$ we have $n-1 \in \mathbb{Z}$ and $n-1 < n$. Starting with 0, we obtain an infinite descending sequence of integers $0 > -1 > -2 > -3 > \cdots$.

Exercise 4.3.18
Prove that the interval $(0,1)$ has infinite descent.

The well-ordering principle is equivalent to the assertion that no inhabited subset of $\mathbb{N}$ has infinite descent—so if we construct an inhabited subset of $\mathbb{N}$ with infinite descent, then we have a contradiction. This gives rise to the following proof strategy.

Strategy 4.3.19 · *Proof by infinite descent*
In order to derive a contradiction from an assumption p, it suffices use p to prove that there is an inhabited subset $X \subseteq \mathbb{N}$ with infinite descent.

Our first example of proof by infinite descent will be used in the proof of the existence of *quotients* and *remainders*, which forms part of the division theorem—we stated the theorem without proof in Theorem 0.42 and will prove it fully in Theorem 7.1.1.

Example 4.3.20 · *Existence of quotients and remainders*
Let $a, b \in \mathbb{Z}$ with $b \neq 0$. We will prove that there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$ and $0 \leqslant r < |b|$.

First, note that since $qb = (-q)(-b)$, we may assume that $b > 0$ — otherwise, replace b by $-b$ and q by $-q$ in what follows.

Define

$$X = \{n \in \mathbb{N} \mid a = mb + n \text{ for some } m \in \mathbb{Z}\} \subseteq \mathbb{N}$$

Note that X is inhabited. Indeed:

- Suppose $a \geqslant 0$. Then $a \in \mathbb{N}$ and $a = 0 \cdot b + a$, so that $a \in X$.
- Suppose $a < 0$. Since $b > 0$ we have $b \geqslant 1$, so that $1 - b \leqslant 0$ and hence $a(1-b) \geqslant 0$. So $a(1-b) \in \mathbb{N}$, and moreover $a = ba + a(1-b)$, so setting $m = b$ and $n = a(1-b)$ in the definition of X, we see that $a(1-b) \in X$.

By the well-ordering principle, X has a minimal element $r \in X$, and by definition of X there exists $q \in \mathbb{Z}$ such that $a = qb + r$. Since $r \in X$, we have $r \in \mathbb{N}$, so that $r \geqslant 0$. It remains to prove that $r < b$.

Towards a contradiction, assume that $r \geqslant b$. We will prove that X, which we already know is an inhabited subset of $\mathbb{N}$, has infinite descent.

So let $n \in X$, and let $m \in \mathbb{Z}$ be such that $a = mn + b$. Since $r \in X$ is minimal, we have $n \geqslant r \geqslant b$, and so $n - b \geqslant 0$. Therefore $n - b \in \mathbb{N}$, and furthermore

$$a = mb + n = (m+1)b + (n-b)$$

so that $n - b \in X$. Since $b > 0$, we also have $n - b < n$.

Thus we have proved that, for all $n \in X$, we have $n - b \in X$ and $n - b < n$, and therefore X has infinite descent. This is a contradiction, so our assumption that $r \geqslant b$ is false, and so it must in fact be the case that $r < b$, as required.

A common application of proof by infinite descent is when proving a statement of the form $\forall n \geqslant n_0, p(n)$ (for some $n_0 \in \mathbb{N}$) by contradiction: if it is false, then there is a counterexample to $p(n)$, and so the set of all counterexamples is inhabited—we prove that the set of all counterexamples has infinite descent, leading to the desired contradiction. This is the strategy we use in the next example.

Example 4.3.21 · *Existence of prime factorisation*

We will prove that every integer $n \geqslant 2$ can be expressed as a product of positive prime numbers.

Towards a contradiction, assume that not every integer $\geqslant 2$ can be expressed as a product of positive prime numbers, and define

$$X = \{n \geqslant 2 \mid n \text{ cannot be expressed as a product of positive prime numbers}\}$$

Note that X is inhabited by assumption.

Let $n \in X$. Then n is not prime, since if it were prime then 'n' would be an expression of n as a product of positive prime numbers (namely, of just one positive prime number).

Since n is not prime, we have $n = ab$ for some $a, b \in \mathbb{Z}$ with $1 < a < n$ and $1 < b < n$. But then a and b cannot both be expressible as products of positive prime numbers, since if $a = p_1 \times \cdots \times p_r$ and $b = q_1 \times \cdots \times q_s$ are two such expressions, then

$$n = ab = p_1 \times \cdots \times p_r \times q_1 \times \cdots \times q_s$$

contradicting the assumption that $n \in X$.

So at least one of a or b cannot be expressed as a product of positive prime numbers. Since $a > 1$ and $b > 1$ we have $a \geqslant 2$ and $b \geqslant 2$, and therefore $a \in X$ or $b \in X$. In either case, we have found an element of X that is less than n. So X has infinite descent, which is a contradiction.

So it must indeed be the case that every integer $\geqslant 2$ can be expressed as a product of primes, as required.

The result of Example 4.3.21 will form part of the *fundamental theorem of arithmetic* (Theorem 7.2.12), but we will give a different proof when we get there. The reason for this is that Example 4.3.21 is non-constructive because it uses the indirect form of proof by contradiction (Strategy 1.3.16), whereas we can give a constructive proof using strong induction instead.

Exercise 4.3.22

Let $n \in \mathbb{N}$. By considering the expression $\sqrt{n} \times \dfrac{\sqrt{n} - \lfloor \sqrt{n} \rfloor}{\sqrt{n} - \lfloor \sqrt{n} \rfloor}$, prove by infinite descent that if $\sqrt{n}$ is not an integer, then $\sqrt{n}$ is irrational.

Section 4.E

Chapter 4 exercises

Recursive definitions

In Questions 4.E.1 to 4.E.5, use the recursive definitions of addition, multiplication and exponentiation directly to prove the desired equation.

4.E.1. $1+3=4$

4.E.2. $0+5=5$

4.E.3. $2\cdot 3=6$

4.E.4. $0\cdot 5=0$

4.E.5. $2^3=8$

4.E.6. Give a recursive definition of new quantifiers $\exists^{=n}$ for $n\in\mathbb{N}$, where given a set X and a predicate $p(x)$, the logical formula $\exists^{=n}x\in X,\ p(x)$ means 'there are exactly n elements $x\in X$ such that $p(x)$ is true'. That is, define $\exists^{=0}$, and then define $\exists^{=n+1}$ in terms of $\exists^{=n}$.

4.E.7. Use the recursive definition of binomial coefficients (Definition 4.1.15) to prove directly that $\binom{4}{2}=6$.

4.E.8. (a) Find the number of trailing 0s in the decimal expansion of 41!.

(b) Find the number of trailing 0s in the binary expansion of 41!.

4.E.9. Let N be a set, let $z\in N$ and let $s:N\to N$. Prove that (N,z,s) is a notion of natural numbers (in the sense of Definition 4.1.1) if and only if, for every set X, every element $a\in X$ and every function $f:X\to X$, there is a unique function $h:N\to X$ such that $h(z)=a$ and $h\circ f=s\circ h$.

Proofs by induction

4.E.10. Find all natural numbers n such that $n^5<5^n$.

4.E.11. Prove that $(1+x)^{123\,456\,789}\geqslant 1+123\,456\,789\,x$ for all real $x\geqslant -1$.

4.E.12. Let $a\in\mathbb{N}$ and assume that the last digit in the decimal expansion of a is 6. Prove that the last digit in the decimal expansion of a^n is 6 for all $n\geqslant 1$.

4.E.13. Let $a,b\in\mathbb{R}$, and let $a_0,a_1,a_2,\ldots$ be a sequence such that $a_0=a$, $a_1=b$ and $a_n=\dfrac{a_{n-1}+a_{n+1}}{2}$ for all $n\geqslant 1$. Prove that $a_n=a+(b-a)n$ for all $n\in\mathbb{N}$.

4.E.14. Let $f:\mathbb{R}\to\mathbb{R}$ be a function such that $f(x+y)=f(x)+f(y)$ for all $x,y\in\mathbb{R}$.

(a) Prove by induction that there is a real number a such that $f(n)=an$ for all $n\in\mathbb{N}$.

(b) Deduce that $f(n)=an$ for all $n\in\mathbb{Z}$.

(c) Deduce further that $f(x) = ax$ for all $x \in \mathbb{Q}$.

4.E.15. Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(0) > 0$ and $f(x+y) = f(x)f(y)$ for all $x, y \in \mathbb{R}$. Prove that there is some positive real number a such that $f(x) = a^x$ for all *rational* numbers x..

4.E.16. Let $a, b \in \mathbb{Z}$. Prove that $b - a$ divides $b^n - a^n$ for all $n \in \mathbb{N}$.

4.E.17. Prove by induction that $7^n - 2 \cdot 4^n + 1$ is divisible by 18 for all $n \in \mathbb{N}$.

Some examples from calculus

Questions 4.E.18 to 4.E.23 assume familiarity with the basic techniques of differential and integral calculus.

4.E.18. Prove that $\dfrac{d^n}{dx^n}(xe^x) = (n+x)e^x$ for all $n \in \mathbb{N}$.

4.E.19. Find and prove an expression for $\dfrac{d^n}{dx^n}(x^2e^x)$ for all $n \in \mathbb{N}$.

4.E.20. Let f and g be differentiable functions. Prove that

$$\frac{d^n}{dx^n}(f(x)g(x)) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x)$$

for all $n \in \mathbb{N}$, where the notation $h^{(r)}$ denotes the r^{th} derivative of h.

4.E.21. Find and prove an expression for the n^{th} derivative of $\log(x)$ for all $n \in \mathbb{N}$.

4.E.22. Prove that $(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$ for all $\theta \in \mathbb{R}$ and all $n \in \mathbb{N}$.

4.E.23. (a) Prove that $\displaystyle\int_0^{\frac{\pi}{2}} \sin^{n+2}(x)\,dx = \frac{n+1}{n+2}\int_0^{\frac{\pi}{2}} \sin^n(x)\,dx$ for all $n \in \mathbb{N}$.

(b) Use induction to prove that $\displaystyle\int_0^{\frac{\pi}{2}} \sin^{2n}(x)\,dx = \frac{\pi}{2^{2n+1}}\binom{2n}{n}$ for all $n \in \mathbb{N}$.

(c) Find and prove an expression for $\displaystyle\int_0^{\frac{\pi}{2}} \sin^{2n+1}(x)\,dx$ for all $n \in \mathbb{N}$.

True–False questions

In Questions 4.E.24 to 4.E.32, determine (with proof) whether the statement is true or false.

4.E.24. Every factorial is positive.

4.E.25. Every binomial coefficient is positive.

4.E.26. Every proof by strong induction can be turned into a proof by weak induction by changing only the induction hypothesis.

4.E.27. Every proof by weak induction can be turned into a proof by strong induction by changing only the induction hypothesis.

4.E.28. Given a set X, a surjection $f : \mathbb{N} \to X$ and a logical formula $p(x)$ with free variable $x \in X$, if $p(f(0))$ is true and $\forall n \in \mathbb{N}$, $(p(f(n)) \Rightarrow p(f(n+1)))$ is true, then $\forall x \in X$, $p(x)$ is true.

4.E.29. Given some logical formula $p(x)$ with free variable $x \in \mathbb{N}$, if for all $x \in \mathbb{N}$ there exists some $y \leqslant x$ such that $p(x)$ is false, then $p(x)$ is false for all $x \in \mathbb{N}$.

4.E.30. Every inhabited subset of the set $\mathbb{Q}^{\geqslant 0}$ of nonnegative rational numbers has a least element.

4.E.31. Every inhabited subset of the set $\mathbb{Z}$ of integers has a least element.

4.E.32. Every inhabited subset of the set $\mathbb{Z}^{-}$ of negative integers has a greatest element.

Always–Sometimes–Never questions

In Questions 4.E.33 to 4.E.35, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

4.E.33. Let $n, k, \ell \in \mathbb{Z}$. Then $\binom{n}{k+\ell} = \binom{n}{k} + \binom{n}{\ell}$.

4.E.34. Let $m, n \geqslant 2$. Then $(m+n)! = m! + n!$.

4.E.35. Let $p(x)$ be a logical formula with free variable $x \in \mathbb{Z}$, and suppose that $p(0)$ is true and, for all $n \in \mathbb{Z}$, if $p(n)$ is true, then $p(n+1)$ and $p(n-2)$ are true. Then $\forall n \in \mathbb{Z}$, $p(n)$ is true.

Chapter 5

Relations

Section 5.1

Relations

Many interesting results or concepts in mathematics arise from observing how the elements of one set interact with the elements of another set, or how elements of a single set interact with each other. We can make this idea of 'interaction' precise using the notion of a *relation*.

Definition 5.1.1

Let X and Y be sets. A (**binary**) **relation** from X to Y is a logical formula $R(x,y)$ with two free variables $x \in X$ and $y \in Y$. We call X the **domain** of R and Y the **codomain** of R.

A relation R is **homogeneous** if it has the same domain and codomain X, in which case we say that R is a relation **on** X.

Given $x \in X$ and $y \in Y$, if $R(x,y)$ is true then we say 'x is **related** to y by R', and write $x\,R\,y$ (LaTeX code: `x \mathrel{R} y`).

Example 5.1.2

We have already seen many examples of relations.

- Divisibility ('x divides y') is a relation on $\mathbb{Z}$.
- The inequality relation $\leqslant$ is a relation on $\mathbb{R}$.
- For any set X, equality $=$ is a relation on X.
- Logical equivalence $\equiv$ is a relation on the set of all logical formulae.
- For any set X, the subset relation $\subseteq$ is a relation on $\mathscr{P}(X)$.

These relations were all homogeneous, but not all relations are:

- For any set X, the elementhood relation $\in$ is a relation from X to $\mathscr{P}(X)$.
- Every function $f : X \to Y$ induces a relation R_f from X to Y, defined by taking $x\,R_f\,y$ to mean $f(x) = y$.

Exercise 5.1.3

Give three more examples of relations, not all of which are homogeneous.

Like with sets and functions, we must determine when to declare that two relations are equal. For example, consider the relation R on $\mathbb{R}$ defined for $a,b \in \mathbb{R}$ by letting $a\,R\,b$ mean $\exists x \in \mathbb{R},\ a + x^2 = b$. It so happens that $a\,R\,b$ if and only if $a \leqslant b$—we'll prove this in Example 5.1.5. So should R be equal to $\leqslant$? On the one hand you might say 'yes', since $\leqslant$ and R relate the same pairs of real numbers. On the other hand you might say 'no', since the fact that $\leqslant$ and R relate the same pairs of real numbers was not immediate and required proof. In fact, if we were to replace $\mathbb{R}$ by $\mathbb{Q}$, it then $\leqslant$ and R would *not* relate the same pairs of elements, since for instance $0 \leqslant 2$ but there is no rational number x such that $0 + x^2 = 2$.

But as with sets and functions, we settle for the *extensional* notion of equality: just as two sets are equal when they have the same elements (Axiom 2.1.14), and two functions are equal when they have the same values (Axiom 3.1.4), we consider two relations to be equal when they relate the same pairs of elements (Axiom 5.1.4).

Axiom 5.1.4 · *Relation extensionality*

Let R and S be relations. Then $R = S$ if and only if R and S have the same domain X and codomain Y, and

$$\forall x \in X, \forall y \in Y, (x\,R\,y \Leftrightarrow x\,S\,y)$$

That is, two relations with the same domain and codomain are equal precisely when they relate the same pairs of elements.

Example 5.1.5

Recall the relation R on $\mathbb{R}$ that we defined above for $a, b \in \mathbb{R}$ by letting $a\,R\,b$ if and only if $a + x^2 = b$ for some $x \in \mathbb{R}$. To see that $R = \leqslant$, note that $a + (b - a) = b$, and that $b - a$ is the square of a real number if and only if $b - a \geqslant 0$, which occurs if and only if $a \leqslant b$.

Exercise 5.1.6

Let R and S be relations on $\mathbb{R}$ defined for $a, b \in \mathbb{R}$ by letting

$$a\,R\,b \;\Leftrightarrow\; b - a \in \mathbb{Q} \quad \text{and} \quad a\,S\,b \;\Leftrightarrow\; \exists n \in \mathbb{Z},\ (n \neq 0) \wedge n(b - a) \in \mathbb{Z}$$

Prove that $R = S$.

The true reason why Axiom 5.1.4 is powerful is that it allows us to reason about relations entirely set theoretically by working with their *graphs*—the sets of pairs of elements that they relate—rather than with the particular formulae defining the relation.

Definition 5.1.7

Let R be a relation from a set X to a set Y. The **graph** of R is the set $\mathrm{Gr}(R)$ (LaTeX code: `\mathrm{Gr}{R}`) of pairs $(x, y) \in X \times Y$ for which $x\,R\,y$. That is

$$\mathrm{Gr}(R) = \{(x, y) \in X \times Y \mid x\,R\,y\} \subseteq X \times Y$$

Example 5.1.8

The graph of the relation $\leqslant$ on $[3]$ is

$$\{(0,0), (0,1), (0,2), (1,1), (1,2), (2,2)\}$$

Likewise, the graph of the relation $\leqslant$ viewed as a relation from $[2]$ to $[4]$ is

$$\{(0,0), (0,1), (0,2), (0,3), (1,1), (1,2), (1,3)\}$$

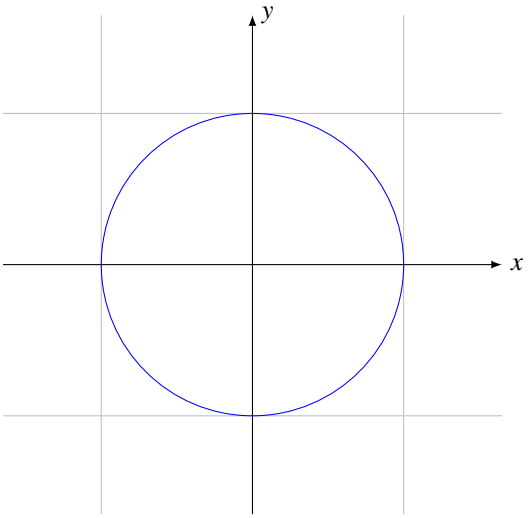
This demonstrates that the graph of a relation is sensitive to the domain (and codomain) of the relation.

Example 5.1.9

Consider the relation C from $\mathbb{R}$ to $\mathbb{R}$ defined by $x\,C\,y \Leftrightarrow x^2+y^2=1$. Then

$$\mathrm{Gr}(C) = \{(x,y) \in \mathbb{R}\times\mathbb{R} \mid x^2+y^2=1\}$$

Plotting $\mathrm{Gr}(C)$ on a standard pair of axes yields a circle with radius 1 centred at the point $(0,0)$, shown below with a unit grid.



Note that $\mathrm{Gr}(C)$ is *not* the graph of a function $f : [0,1] \to \mathbb{R}$, since for example both $(0,1)$ and $(0,-1)$ are elements of $\mathrm{Gr}(C)$, the value $f(0)$ would not be uniquely defined.

Exercise 5.1.10

Let R be the relation on $\mathbb{Z}$ defined for $x,y \in \mathbb{Z}$ by letting $x\,R\,y$ if and only if $x^2 = y^2$. Describe its graph $\mathrm{Gr}(R) \subseteq \mathbb{Z}\times\mathbb{Z}$.

Exercise 5.1.11

Let $f : X \to Y$ be a function, and define the relation R_f from X to Y as in Example 5.1.2. Prove that $\mathrm{Gr}(R_f) = \mathrm{Gr}(f)$—that is, the graph of the *relation* R_f is equal to the graph of the *function* f.

Definition 5.1.12

The **discrete relation** from a set X to a set Y is the relation $D_{X,Y}$ defined by letting $x\,D_{X,Y}\,y$ be true for all $x \in X$ and $y \in Y$.

The **empty relation** from a set X to a set Y is the relation $\varnothing_{X,Y}$ (LaTeX code: `\varnothing`) defined by letting $x\,\varnothing_{X,Y}\,y$ be false for all $x \in X$ and $y \in Y$.

Exercise 5.1.13
Let X and Y be sets. Describe the graphs $\mathrm{Gr}(D_{X,Y})$ and $\mathrm{Gr}(\varnothing_{X,Y})$.

It turns out that, for fixed sets X and Y, relations from X to Y correspond with subsets of $X \times Y$—see Theorem 5.1.14 below. This fact is so convenient that many (if not most) authors actually define 'relation from X to Y' to mean 'subset of $X \times Y$'.

Theorem 5.1.14
Let X and Y be sets. Every subset $G \subseteq X \times Y$ is the graph of a unique relation R from X to Y.

Proof
Fix $G \subseteq X \times Y$. Define a relation R by

$$\forall x \in X, \forall y \in Y, x\,R\,y \Leftrightarrow (x,y) \in G$$

Then certainly $G = \mathrm{Gr}(R)$, since for all $x \in X$ and $y \in Y$ we have

$$(x,y) \in G \quad \Leftrightarrow \quad x\,R\,y \quad \Leftrightarrow \quad (x,y) \in \mathrm{Gr}(R)$$

Moreover, if S is a relation from X to Y such that $G = \mathrm{Gr}(S)$, then, for all $x \in X$ and $y \in Y$

$$x\,S\,y \Leftrightarrow (x,y) \in \mathrm{Gr}(S) \Leftrightarrow (x,y) \in G \Leftrightarrow x\,R\,y$$

so $S = R$.

Hence there is exactly one relation from X to Y whose graph is G. □

Theorem 5.1.14 suggests that, for the purposes of defining relations and proving that relations are equal, we may work entirely set theoretically with the graphs of the relations.

Strategy 5.1.15 · *Relations as graphs*
In order to specify a relation R, it suffices to specify its domain X, its codomain Y, and its graph $\mathrm{Gr}(R) \subseteq X \times Y$. Furthermore, in order to prove that two relations R and S are equal, it suffices to prove that they have the same domain and codomain, and that their graphs are equal.

Example 5.1.16
Consider the set $G = \{(2m+i, 2n+i) \mid m,n \in \mathbb{Z},\ i \in \{0,1\}\}$. Since $G \subseteq \mathbb{Z} \times \mathbb{Z}$, it is the graph of a (unique) relation R on $\mathbb{Z}$, which is necessarily defined for $a,b \in \mathbb{Z}$ by letting $a\,R\,b$ if and only if there are integers m and n, and $i \in \{0,1\}$, such that $a = 2m+i$ and $b = 2n+i$. But this says precisely that a and b both leave the same remainder (namely i) when divided by 2, so that R can be described by saying that, for all $a,b \in \mathbb{Z}$, we have $a\,R\,b$ if and only if a and b are both even or both odd.

Definition 5.1.17
Let X be a set. The **diagonal subset** of $X \times X$ is the set Δ_X (LaTeX code: `\Delta_X`) defined by $\Delta_X = \{(x,x) \mid x \in X\}$.

To see why Δ_X is called the 'diagonal' subset, try plotting $\Delta_{\mathbb{R}} \subseteq \mathbb{R} \times \mathbb{R}$ on a standard pair of axes (like in Example 5.1.9).

Exercise 5.1.18
Let X be a set. Prove that $\Delta_X = \mathrm{Gr}(=)$.

Properties of homogeneous relations

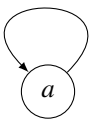
Most of the relations of interest to us in this book are homogeneous—that is, relations *on* a set. In fact, they broadly fall into one of two categories: *equivalence relations*, which are relations that 'behave like $=$'; and *order relations*, which are relations that 'behave like $\leqslant$'. We will study equivalence relations in Section 5.2 and order relations in Section 12.1, but examples of such relations pop up throughout the book. (In fact, we have already seen several!)

Our task for the rest of this section is to isolate the properties that a relation must satisfy in order to be classified as an equivalence relation or an order relation.

To aid with intuition, we will illustrate these properties with diagrams: given a relation R, the fact that $a \mathrel{R} b$ will be represented diagramatically as follows:

A *reflexive* relation is one that relates every element of its domain to itself.

Definition 5.1.19
A relation R on a set X is **reflexive** if $a \mathrel{R} a$ for all $a \in X$.



Example 5.1.20
Given any set X, the equality relation $=$ on X is reflexive, since $a = a$ for all $a \in X$.

Example 5.1.21
Let R be the relation on $\mathbb{R}$ defined for $a, b \in \mathbb{R}$ by $a \mathrel{R} b$ if and only if $b - a \in \mathbb{Q}$. Then R is reflexive, since for all $a \in \mathbb{R}$, we have $a - a = 0 \in \mathbb{Q}$, so that $a \mathrel{R} a$.

Exercise 5.1.22
Let X be a set. Prove that $\subseteq$ is a reflexive relation on $\mathscr{P}(X)$, but $\subsetneq$ is not.

Exercise 5.1.23
Prove that the relation 'x divides y' on $\mathbb{Z}$ is reflexive.

The next exercise demonstrates that when determining if a relation is reflexive, we must be careful to specify its domain.

Exercise 5.1.24
Let $G = \{(0,0),(1,1),(2,2)\}$. Let R be the relation on $[3]$ whose graph is G, and let S be the relation on $[4]$ whose graph is G. Prove that R is reflexive, but S is not.

Symmetric relations are those for which the *direction* of the relation doesn't matter: two elements are either each related to the other, or not related at all.

Definition 5.1.25
A relation R on a set X is **symmetric** if, for all $a,b \in X$, if $a\,R\,b$, then $b\,R\,a$.

Example 5.1.26
Given any set X, the equality relation $=$ on X is symmetric, since for all $a,b \in X$, if $a = b$, then $b = a$.

Example 5.1.27
Let R be the relation on $\mathbb{R}$ defined for $a,b \in \mathbb{R}$ by $a\,R\,b$ if and only if $b-a \in \mathbb{Q}$. Then R is symmetric.

To see this, let $a,b \in \mathbb{R}$ and assume that $a\,R\,b$. Then $b-a \in \mathbb{Q}$, so that $b-a = \frac{p}{q}$ for some $p,q \in \mathbb{Z}$ with $q \neq 0$. But then

$$a-b = -(b-a) = \frac{-p}{q}$$

so that $a-b \in \mathbb{Q}$. Hence $b\,R\,a$, as required.

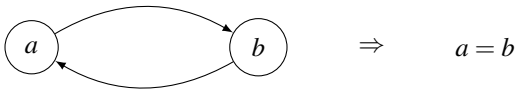
Exercise 5.1.28
Find all subsets $U \subseteq \mathbb{Z}$ such that the relation 'x divides y' on U is symmetric.

We showed in Exercise 5.1.24 that reflexivity of a relation is sensitive to its domain. The next exercise demonstrates that symmetry is *not* sensitive to the domain—that is, it is an *intrinsic* property of the relation.

Exercise 5.1.29
Let R and S be relations such that $\mathrm{Gr}(R) = \mathrm{Gr}(S)$. Note that the domain of R might be different from the domain of S. Prove that R is symmetric if and only if S is symmetric.

A condition related to symmetry, but in a sense opposite to it, is *antisymmetry*. It says that the only way that two elements of a set can each be related to the other is if they are equal.

Definition 5.1.30
Let X be a set. A relation R on X is **antisymmetric** if, for all $a, b \in X$, if $a \mathrel{R} b$ and $b \mathrel{R} a$, then $a = b$.



A word of warning here is that 'antisymmetric' does not mean the same thing as 'not symmetric'—indeed, we we will see, equality is both symmetric and antisymmetric, and many relations are neither symmetric nor antisymmetric. [Even more confusingly, there is a notion of *asymmetric relation*, which also does not mean 'not symmetric'.]

Example 5.1.31
Given any set X, the equality relation $=$ on X is antisymmetric, since for all $a, b \in X$, if $a = b$ and $b = a$, then $a = b$.

Example 5.1.32
The order relation $\leqslant$ on $\mathbb{R}$ is antisymmetric, since for all $a, b \in \mathbb{R}$, if $a \leqslant b$ and $b \leqslant a$, then $a = b$.

Exercise 5.1.33
Prove that the relation 'x divides y' on $\mathbb{N}$ is antisymmetric, but not on $\mathbb{Z}$.

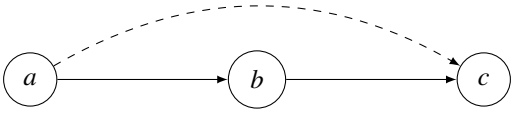
Exercise 5.1.34
Let X be a set. Prove that $\subseteq$ is an antisymmetric relation on $\mathscr{P}(X)$.

Exercise 5.1.35
Let X be a set and let R be a relation on X. Prove that R is both symmetric and antisymmetric if and only if $\mathrm{Gr}(R) \subseteq \Delta_X$, where Δ_X is the diagonal subset of $X \times X$ (see Definition 5.1.17). Deduce that the only reflexive, symmetric and antisymmetric relation on a set X is the equality relation on X.

The last property we will study in some detail is *transitivity*. Transitive relations are those for which we can skip over intermediate related elements—for example, we can deduce $0 < 3$ from the facts that $0 < 1$ and $1 < 2$ and $2 < 3$.

Definition 5.1.36

A relation R on a set X is **transitive** if, for all $a,b,c \in X$, if $a\,R\,b$ and $b\,R\,c$, then $a\,R\,c$.



Example 5.1.37

Given any set X, the equality relation $=$ on X is transitive since, for all $a,b,c \in X$, if $a = b$ and $b = c$, then $a = c$.

Example 5.1.38

Let R be the relation on $\mathbb{R}$ defined for $a,b \in \mathbb{R}$ by $a\,R\,b$ if and only if $b-a \in \mathbb{Q}$. Then R is transitive.

To see this, let $a,b,c \in \mathbb{R}$ and assume that $a\,R\,b$ and $b\,R\,c$. Then $b-a \in \mathbb{Q}$ and $c-b \in \mathbb{Q}$, so there exist $p,q,r,s \in \mathbb{Z}$ with $q,s \neq 0$ such that

$$b-a = \frac{p}{q} \quad \text{and} \quad c-b = \frac{r}{s}$$

It follows that

$$c-a = (c-b)+(b-a) = \frac{p}{q}+\frac{r}{s} = \frac{ps+qr}{qs}$$

so that $c-a \in \mathbb{Q}$. Hence $a\,R\,c$, as required.

Exercise 5.1.39

Let X be a set. Prove that $\subseteq$ is a transitive relation on $\mathscr{P}(X)$.

Exercise 5.1.40

Prove that the relation 'x divides y' on $\mathbb{Z}$ is transitive.

Like symmetry, transitive is an intrinsic property of relations—that is, transitivity is not sensitive to the domain of the relation—as the next exercise demonstrates.

Exercise 5.1.41

Let R and S be relations such that $\mathrm{Gr}(R) = \mathrm{Gr}(S)$. Note that the domain of R might be different from the domain of S. Prove that R is transitive if and only if S is transitive.

A fundamental property of transitive relations is that we can prove two elements a and b are related by finding a chain of related elements starting at a and finishing at b. This is the content of the following proposition.

Proposition 5.1.42

Let R be a relation on a set X. Then R is transitive if and only if, for any finite sequence $x_0, x_1, \ldots, x_n$ of elements of X such that $x_{i-1}\, R\, x_i$ for all $i \in [n]$, we have $x_0\, R\, x_n$.

Proof

For the sake of abbreviation, let $p(n)$ be the assertion that, for any $n \geqslant 1$ and any sequence $x_0, x_1, \ldots, x_n$ of elements of X such that $x_{i-1}\, R\, x_i$ for all $i \in [n]$, we have $x_0\, R\, x_n$.

We prove the two directions of the proposition separately.

- ($\Rightarrow$) Suppose R is transitive. For $n \geqslant 1$. We prove $p(n)$ is true for all $n \geqslant 1$ by induction.
  - ◇ (**Base case**) When $n = 1$ this is immediate, since we assume that $x_0\, R\, x_1$.
  - ◇ (**Induction step**) Fix $n \geqslant 1$ and suppose $p(n)$ is true. Let $x_0, \ldots, x_n, x_{n+1}$ is a sequence such that $x_{i-1}\, R\, x_i$ for all $i \in [n+1]$. We need to prove that $x_0\, R\, x_{n+1}$.

    By the induction hypothesis we know that $x_0\, R\, x_n$. By definition of the sequence we have $x_n\, R\, x_{n+1}$. By transitivity, we have $x_0\, R\, x_{n+1}$.

  So by induction, we have proved the $\Rightarrow$ direction.
- ($\Leftarrow$) Suppose $p(n)$ is true for all $n \geqslant 1$. Then in particular $p(2)$ is true, which is precisely the assertion that R is transitive.

So we're done. □

That is, Proposition 5.1.42 states that for a transitive relation R on a set X, if $x_0, x_1, \ldots, x_n \in X$, then

$$x_0\, R\, x_1\, R \cdots R\, x_n \quad \Rightarrow \quad x_0\, R\, x_n$$

where '$x_0\, R\, x_1\, R \cdots R\, x_n$' abbreviates the assertion that $x_i\, R\, x_{i+1}$ for each $i < n$.

Section 5.2

Equivalence relations and partitions

An equivalence relation on a set X is a relation on X that, to a certain extent, *behaves like equality*. That is, equivalence relations give us a way of saying that two elements of a set are 'similar', without having to be equal. As an example, we might be interested in when the base-10 expansions of two natural numbers end in the *same* digit, or when two finite sets have the *same* number of elements.

Definition 5.2.1

A relation R on a set X is an **equivalence relation** if it is reflexive, symmetric and transitive.

To denote a relation that we know (or suspect) is an equivalence relation, we will usually use a symbol like '$\sim$' (LaTeX code: `\sim`) or '$\equiv$' (LaTeX code: `\equiv`) or '$\approx$' (LaTeX code: `\approx`) instead of a letter like 'R' or 'S'.

Example 5.2.2

Given any set X, it follows from Examples 5.1.20, 5.1.26 and 5.1.37 that the equality relation $=$ is an equivalence relation on X. This is a relief, since we motivated equivalence relations by saying that they are those that behave like equality!

Example 5.2.3

Let R be the relation on $\mathbb{R}$ defined for $a, b \in \mathbb{R}$ by $a\,R\,b$ if and only if $b - a \in \mathbb{Q}$. Piecing together Examples 5.1.21, 5.1.27 and 5.1.38, we see that R is an equivalence relation on $\mathbb{R}$.

Exercise 5.2.4

Given a function $f : X \to Y$, define a relation $\sim_f$ on X by

$$a \sim_f b \quad \Leftrightarrow \quad f(a) = f(b)$$

for all $a, b \in X$. Prove that $\sim_f$ is an equivalence relation on X.

The equivalence relation in the next exercise comes back with a vengeance in Section 10.1, where we will use it to compare the sizes of (finite and) infinite sets.

Exercise 5.2.5

Let $\mathscr{S}$ be some set whose elements are all sets. (For example, we could take $\mathscr{S} = \mathscr{P}(X)$ for some fixed set X.) Define a relation $\cong$ (LaTeX code: `\cong`) on $\mathscr{S}$ by letting $U \cong V$ if and only if there exists a bijection $f : U \to V$, for all $U, V \in \mathscr{S}$. Prove that $\cong$ is an equivalence relation on $\mathscr{S}$.

A first look at modular arithmetic

A particularly useful family of equivalence relations is given by *congruence* of integers, which allows us to do *modular arithmetic*—this is the topic of Section 7.3. For a fixed integer n, this relation identifies two integers when they have the same remainder upon division by n (as in Theorem 0.42).

Definition 5.2.6

Fix $n \in \mathbb{Z}$. Given integers $a, b \in \mathbb{Z}$, we say a **is congruent to** b **modulo** n, and write

$$a \equiv b \pmod{n} \quad (\text{LaTeX code: } \texttt{a \equiv b \pmod\{n\}})$$

if $n \mid b - a$. If a is not congruent to b modulo n, write

$$a \not\equiv b \pmod{n} \quad (\text{LaTeX code: } \texttt{\not\equiv})$$

The number n is called the **modulus** of the congruence.

Before we prove that congruence is modulo n (for fixed $n \in \mathbb{Z}$) is an equivalence relation for all $n \in \mathbb{Z}$, it is worthwhile to get a feel for how it works.

Example 5.2.7

Let $a, b \in \mathbb{Z}$. Then $a \equiv b \pmod{2}$ if and only if a and b are both even or both odd—that is, if and only if they have the same *parity*.

Indeed, by the division theorem, we can write $a = 2k + i$ and $b = 2\ell + j$ for some $k, \ell \in \mathbb{Z}$ and $i, j \in \{0, 1\}$. Then

$$b - a \;=\; (2k + i) - (2\ell + j) \;=\; 2(k - \ell) + (i - j)$$

Note that $i - j \in \{-1, 0, 1\}$, and so $a \equiv b \pmod{2}$ if and only if $i = j$. But this occurs if and only if $i = j = 0$, in which case a and b are both even, or $i = j = 1$, in which case a and b are both odd.

Example 5.2.8

Let $a, b \in \mathbb{N}$. Then $a \equiv b \pmod{10}$ if and only if 10 divides $b - a$, which occurs if and only if the last digit in the decimal expansion of $b - a$ is 0. But this implies that the decimal expansions of a and b have the same last digit. So the relation of congruence modulo 10 on $\mathbb{N}$ is the same as the relation of 'having the same last (decimal) digit'.

Exercise 5.2.9

Let $n \in \mathbb{Z}$. Prove that if $n \neq 0$, then $a \equiv b \pmod{n}$ if and only if a and b have the same remainder when divided by n.

Exercise 5.2.10

Let $a, b \in \mathbb{Z}$. When is it true that $a \equiv b \pmod{0}$? When is it true that $a \equiv b \pmod{1}$?

Having got a better feel for how congruence works, we now prove that, for each $n \in \mathbb{Z}$, congruence modulo n is an equivalence relation on $\mathbb{Z}$.

Theorem 5.2.11

Let $n \in \mathbb{Z}$. Then congruence modulo n is an equivalence relation on $\mathbb{Z}$. That is

(a) $a \equiv a \pmod{n}$ for all $a \in \mathbb{Z}$;

(b) For all $a, b \in \mathbb{Z}$, if $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$;

(c) For all $a, b, c \in \mathbb{Z}$, if $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

Proof

(a) Let $a \in \mathbb{Z}$. Note that $a - a = 0$, which is divisible by n since $0 = 0 \times n$, and hence $a \equiv a \pmod{n}$. So congruence modulo n is reflexive.

(b) Let $a, b \in \mathbb{Z}$ and suppose $a \equiv b \pmod{n}$. Then n divides $a - b$, so that $a - b = kn$ for some $k \in \mathbb{Z}$. Hence $b - a = -kn$, and so n divides $b - a$, so that $b \equiv a \pmod{n}$ as required. So congruence modulo n is symmetric.

(c) Let $a, b, c \in \mathbb{Z}$ and suppose that $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$. Then n divides both $a - b$ and $b - c$, so there exist $k, \ell \in \mathbb{Z}$ such that

$$a - b = kn \quad \text{and} \quad b - c = \ell n$$

Hence $a - c = (a - b) + (b - c) = (k + \ell)n$, so that n divides $a - c$. Hence $a \equiv c \pmod{n}$, as required. So congruence modulo n is transitive.

Since congruence modulo n is reflexive, symmetric and transitive, it is an equivalence relation. □

Equivalence classes and quotients

What makes equivalence relations so useful is they give us a way of ignoring information that is irrelevant to the task at hand.

For example, suppose a and b are two very large natural numbers, each with several trillion (decimal) digits. We want to know what the last digit of ab is. To find this out, it would be silly to compute ab and then look at its last digit. Instead, we can observe that the last digit of a product of two integers depends only on the last digit of each integer—for example, 1527×9502 has the same last digit as $7 \times 2 = 14$. By using the equivalence relation 'has the same last digit as', we are able to ignore the irrelevant information about a and b—that is, all but one of their trillions of digits—and simplify the problem considerably.

To make this precise, we introduce the notion of an *equivalence class*. For a set X with an equivalence relation, the equivalence class of an element $a \in X$ will be the set of elements of X that a is equivalent to. By working with the *equivalence classes* of elements of X, rather than the elements of X themselves, we are able to regard two equivalent elements as being 'the same'.

Definition 5.2.12

Let X be a set and let $\sim$ be an equivalence relation on X. The **$\sim$-equivalence class** of an element $a \in X$ is the set $[a]_\sim$ (LaTeX code: `[x]_{\sim}`) defined by

$$[a]_\sim = \{x \in X \mid a \sim x\}$$

The **quotient** of X by $\sim$ is the set $X/\sim$ (LaTeX code: `X/{\sim}`) of all $\sim$-equivalence classes of elements of X; that is

$$X/\sim = \{[a]_\sim \mid a \in X\}$$

LaTeX tip

Putting {curly brackets} around the command for a symbol like $\sim$ (`\sim`) tells LaTeX to consider the symbol *as a symbol*, rather than as a connective. Compare the following:

TEX code	*Output*
`X/\sim = Y`	$X/\sim = Y$
`X/{\sim} = Y`	$X/{\sim} = Y$

This is because, without braces, LaTeX thinks you're saying 'X-forward-slash is related to is equal to Y', which clearly makes no sense; putting braces around `\sim` signifies to LaTeX that the $\sim$ symbol is being considered as an object in its own right, rather than as a connective.

Example 5.2.13

Let $\sim$ be the relation of congruence modulo 2 on $\mathbb{Z}$. We showed in Example 5.2.7 that, for all $a, b \in \mathbb{Z}$ we have $a \equiv b \pmod 2$ if and only if a and b have the same parity. But this means that, for all $[a]_\sim$ is the set of all integers with the same parity as a—that is:

- If a is even, then $[a]_\sim$ is the set of all even integers; and
- If a is odd, then $[a]_\sim$ is the set of all odd integers.

It follows that $\mathbb{Z}/\sim = \{[0]_\sim, [1]_\sim\} = \{E, O\}$, where E is the set of all even integers and O is the set of all odd integers.

Exercise 5.2.14

Let $\approx$ be the relation of congruence modulo 10 on $\mathbb{N}$. Describe the equivalence classes, and give an explicit expression of the quotient $\mathbb{N}/\approx$ in list notation.

Example 5.2.15

Let $f : X \to Y$ be a function, and let $\sim_f$ be the equivalence relation on X that we defined in Exercise 5.2.4. Given $a \in X$, we have

$$[a]_{\sim_f} = \{x \in X \mid a \sim_f x\} = \{x \in X \mid f(a) = f(x)\}$$

Thus we have $[a]_{\sim_f} = f^{-1}[\{f(a)\}]$.

Exercise 5.2.16

Let $f : X \to Y$ be a function. Prove that f is injective if and only if each $\sim_f$-equivalence class has a unique element, where $\sim_f$ is the equivalence relation defined in Exercise 5.2.4.

The next result demonstrates that an equivalence relation $\sim$ on a set X 'descends' to the equality relation $=$ on the quotient $X/\sim$. This means that if we would rather deal with equality than with the equivalence relation itself, then we may do so by working inside the quotient $X/\sim$ rather than in the set X.

Theorem 5.2.17

Let $\sim$ be an equivalence relation on a set X. Then for all $a, b \in X$, we have $a \sim b$ if and only if $[a]_\sim = [b]_\sim$.

Proof

The proof is an exercise in piecing together the properties of equivalence relations.

Fix $a, b \in X$.

- ($\Rightarrow$) Suppose $a \sim b$. We prove $[a]_\sim = [b]_\sim$ by double containment.
  - ($\subseteq$) Let $x \in [a]_\sim$—then $a \sim x$. We are assuming that $a \sim b$, so that $b \sim a$ by symmetry, and so $b \sim x$ by transitivity. So $x \in [b]_\sim$.
  - ($\supseteq$) Let $x \in [b]_\sim$—then $b \sim x$. We are assuming that $a \sim b$, and so $a \sim x$ by transitivity. So $x \in [a]_\sim$.

  We have shown by double containment that $[a]_\sim = [b]_\sim$.
- ($\Leftarrow$) Assume $[a]_\sim = [b]_\sim$. We have $b \sim b$ by reflexivity, and so $b \in [b]_\sim$. But then $b \in [a]_\sim$, so that $a \sim b$, as required.

So $a \sim b$ if and only if $[a]_\sim = [b]_\sim$. □

Congruence classes

For congruence, special terminology and notation exists for equivalence classes and quotients.

Definition 5.2.18

Let $n \in \mathbb{Z}$. The **congruence class** of an integer a modulo n is defined by

$$[a]_n \;=\; [a]_{\equiv_n} \;=\; \{x \in \mathbb{Z} \mid a \equiv x \pmod{n}\}$$

where $\equiv_n$ denotes the relation of congruence modulo n.

The set of all congruence classes modulo n is denoted by

$$\mathbb{Z}/n\mathbb{Z} \;=\; \mathbb{Z}/{\equiv_n} \;=\; \{[a]_n \mid a \in \mathbb{Z}\}$$

Example 5.2.19

Using the terminology of congruence classes, Example 5.2.13 can be rephrased by saying that $\mathbb{Z}/2\mathbb{Z} = \{[0]_2, [1]_2\}$. Moreover, Theorem 5.2.17 gives us a more succinct proof: for all $a \in \mathbb{Z}$, we have $a \equiv 0 \pmod 2$ if and only if a is even, and $a \equiv 1 \pmod 2$ if and only if a is odd. Therefore for all $a \in \mathbb{Z}$, we have $[a]_2 = [0]_2$ or $[a]_2 = [1]_2$, and so

$$\mathbb{Z}/2\mathbb{Z} \;=\; \{[a]_2 \mid a \in \mathbb{Z}\} \;=\; \{[0]_2, [1]_2\}$$

Additionally, $[0]_2$ is the set of all even integers and $[1]_2$ is the set of all odd integers.

The next exercise generalises the previous one, proving that congruence classes correspond with remainders.

Exercise 5.2.20

Let $n \in \mathbb{Z}$. Prove that the function

$$i : [|n|] \to \mathbb{Z}/n\mathbb{Z}$$

defined by $i(r) = [r]_n$ for all $r \in [|n|]$ is a bijection.

Partitions

A partition of a set X is a way of breaking X up into mutually disjoint subsets. They will be an immensely useful tool for counting how many elements a finite set has in Chapter 8, and will reappear in Section 10.2 for defining arithmetic operations with cardinal numbers.

Definition 5.2.21

A **partition** of a set X is a set $\mathscr{A} \subseteq \mathscr{P}(X)$ (that is, a set of subsets of X) such that the following conditions hold:

(a) Every set $A \in \mathscr{A}$ is inhabited;

(b) $\mathscr{A}$ is a **cover** of X: for all $x \in X$, there exists $A \in \mathscr{A}$ such that $x \in A$; and

(c) $\mathscr{A}$ is **pairwise disjoint**: for all $A, B \in \mathscr{A}$, if $A \cap B$ is inhabited, then $A = B$.

A few remarks about the definition of a partition are in order:

- Later in the textbook, we will encounter contexts in which we want to talk about a partition of a set *without* requiring that the sets in the partition be inhabited. We will address this when we need to, but for the purposes of this section, we *do* need this condition.
- The definition of $\mathscr{A}$ being a cover of X is equivalent to the assertion that $X \subseteq \bigcup \mathscr{A}$. Furthermore, since $\mathscr{A}$ is a subset of X, we automatically have $\bigcup \mathscr{A} \subseteq X$. So a set $\mathscr{A} \subseteq \mathscr{P}(X)$ is a cover of X if and only if $\bigcup \mathscr{A} = X$.
- Pairwise disjointness of $\mathscr{A}$ means, intuitively, that distinct sets in $\mathscr{A}$ do not overlap. If you were asked to write down what this means, you would probably say something like: "for all $A, B \in \mathscr{A}$, if $A \neq B$, then $A \cap B$ is empty". This is good for intuition, but less useful for proofs, since both the hypothesis and the conclusion of the implication are negative statements, which would suggest a proof by contradiction every time we want to show a collection of sets is pairwise disjoint. The definition we gave in Definition 5.2.21 uses the *contrapositive* of this implication, which we know is equivalent by Theorem 1.3.19, but is more useful for proofs of pairwise disjointness because it does not contain any negations.

Example 5.2.22

Let E be the set of all even integers and O be the set of all odd integers. Then $\{E, O\}$ is a partition of $\mathbb{Z}$:

- (**Inhabitation**) E and O are inhabited, since for example $0 \in E$ and $1 \in O$.
- (**Covering**) Every integer n is either even (so $n \in E$) or odd (so $n \in O$), and hence $\{E, O\}$ is a cover of $\mathbb{Z}$.
- (**Pairwise disjointness**) The family $\{E, O\}$ is pairwise disjoint if and only if $E \cap O$ is empty; and it is, since no integer can be both even and odd.

Example 5.2.23

Let $\mathscr{A} = \{\{2n, 2n+1\} \mid n \in \mathbb{N}\}$. Then $\mathscr{A}$ is a partition of $\mathbb{N}$:

- (**Inhabitation**) Let $A \in \mathscr{A}$. Then $A = \{2n, 2n+1\}$ for some $n \in \mathbb{N}$, and so $2n \in A$. Hence A is inhabited.
- (**Covering**) Let $x \in \mathbb{N}$. If x is even, then $x = 2n$ for some $n \in \mathbb{N}$, and if x is odd, then $x = 2n+1$ for some $n \in \mathbb{N}$. In both cases there is a natural number n such that $x \in \{2n, 2n+1\}$, and $\{2n, 2n+1\} \in \mathscr{A}$. So $\mathscr{A}$ is a cover of $\mathbb{N}$.
- (**Pairwise disjointness**) Let $A, B \in \mathscr{A}$ and assume that $A \cap B$ is inhabited. Let $m, n \in \mathbb{N}$ be such that $A = \{2m, 2m+1\}$ and $B = \{2n, 2n+1\}$, and let $x \in A \cap B$, which exists since $A \cap B$ is inhabited. Since A and B each have a unique even element and a unique odd element, we must either have $x = 2m = 2n$ or $x = 2m+1 = 2n+1$ depending on whether x is even or odd. But in both cases we have $m = n$, so that $A = B$ as required. So $\mathscr{A}$ is pairwise disjoint.

Exercise 5.2.24

Let $f : X \to Y$ be a surjection, and define a collection $\mathscr{F}$ of subsets of X by

$$\mathscr{F} = \{f^{-1}[\{b\}] \mid b \in Y\}$$

That is, $\mathscr{F}$ is the set of subsets of X given by the preimages of individual elements of Y under f. Prove that $\mathscr{F}$ is a partition of X. Where in your proof do you use surjectivity of f?

Exercise 5.2.25

Let X be a set and let $\mathscr{A}$ be a set of inhabited subsets of X. Prove that $\mathscr{A}$ is a partition of X if and only if for each $x \in X$, there is a unique set $A \in \mathscr{A}$ with $a \in A$.

Theorem 5.2.26 · *Equivalence relations correspond with partitions*

Let X be a set.

(a) For all equivalence relations $\sim$ on X, the quotient $X/\!\sim$ is a partition of X.

(b) For all partitions $\mathscr{A}$ of X, there is a unique equivalence relation $\sim$ on X such that $\mathscr{A} = X/\!\sim$.

Proof of (b)

Define a relation $\sim$ by

$$x \sim y \quad \Leftrightarrow \quad \exists A \in \mathscr{A}, x \in A \text{ and } y \in A$$

for all $x, y \in X$. That is, $x \sim y$ if and only if x and y are elements of the same set of the partition. We check that $\sim$ is an equivalence relation.

- **Reflexivity.** Let $x \in X$. Then $x \in A$ for some $A \in \mathscr{A}$ since $\mathscr{A}$ is a cover of X.
- **Symmetry.** Let $x, y \in X$ and suppose $x \sim y$. Then there is some $A \in \mathscr{A}$ with $x \in A$ and $y \in A$. But then it is immediate that $y \sim x$.
- **Transitivity.** Let $x, y, z \in X$ and suppose that $x \sim y$ and $y \sim z$. Then there exist $A, B \in \mathscr{A}$ with $x, y \in A$ and $y, z \in B$. Thus $y \in A \cap B$. Since $A \cap B$ is inhabited, it follows by pairwise disjointness that $A = B$. Therefore $x \in A$ and $z \in A$, and so $x \sim z$, as required.

We prove $X/\sim = \mathscr{A}$ by double containment; each inclusion direction will itself require a further double containment proof.

($\subseteq$) Let $E \in X/\sim$. Then $E = [x]_\sim$ for some $x \in X$. Let $A \in \mathscr{A}$ be the unique set in the partition with $x \in A$. We claim that $E = A$.

($\subseteq$) Let $y \in E = [x]_\sim$. Then $x \sim y$, so there exists $B \in \mathscr{A}$ such that $x \in B$ and $y \in B$. But $x \in A$ and $\mathscr{A}$ is pairwise disjoint, so $B = A$, and so $y \in A$ as required.

($\supseteq$) Let $y \in A$. Then since $x \in A$ and $y \in A$, we have $x \sim y$, so that $y \in [x]_\sim = E$, as required.

Therefore $E = A \in \mathscr{A}$, as required.

($\supseteq$) Let $A \in \mathscr{A}$. Then A is inhabited, so there exists an element $x \in A$. We claim that $A = [x]_\sim$.

($\subseteq$) Let $y \in A$. Then $x \in A$ and $y \in A$, so $x \sim y$, and hence $y \in [x]_\sim$.

($\supseteq$) Let $y \in [x]_\sim$. Then $x \sim y$, so there exists $B \in \mathscr{A}$ with $x \in B$ and $y \in B$. But $x \in A$ and $\mathscr{A}$ is pairwise disjoint, so $A = B$, and so $y \in A$ as required.

So $A = [x]_\sim \in X/\sim$, as required.

To prove that $\sim$ is the only such relation, suppose $\approx$ is another equivalence relation on X for which $X/\approx = \mathscr{A}$. Then, given $x, y \in X$, we have:

$$\begin{aligned}
x \sim y &\Leftrightarrow \exists A \in \mathscr{A}, x \in A \wedge y \in A && \text{by definition of } \sim \\
&\Leftrightarrow \exists z \in X, x \in [z]_\approx \wedge y \in [z]_\approx && \text{since } \mathscr{A} = X/\approx \\
&\Leftrightarrow \exists z \in X, x \approx z \wedge y \approx z && \text{by definition of } [z]_\approx \\
&\Leftrightarrow x \approx y && \text{by symmetry and transitivity}
\end{aligned}$$

So $\sim = \approx$. □

Exercise 5.2.27

Prove part (a) of Theorem 5.2.26.

Theorem 5.2.26 implies that equivalence relations and quotients are essentially the same thing: the quotient of a set by an equivalence relation is a partition of the set, and every partition of a set is the quotient by a unique equivalence relation. Another way of stating this is that the function

$$\mathscr{Q} : \left\{ \begin{matrix} \text{equivalence} \\ \text{relations on } X \end{matrix} \right\} \to \left\{ \begin{matrix} \text{partitions} \\ \text{of } X \end{matrix} \right\}$$

defined by $\mathscr{Q}(\sim) = X/\sim$ for all equivalence relations $\sim$ on X, is a bijection. Well-definedness of $\mathscr{Q}$ is ensured by part (a) of Theorem 5.2.26, and the fact that $\mathscr{Q}$ is a bijection is ensured by part (b).

Exercise 5.2.28

For each $i \in [0,1)$ let $A_i = \{i+n \mid n \in \mathbb{Z}\} \subseteq \mathbb{R}$, and define $\mathscr{A} = \{A_i \mid i \in [0,1)\}$.

Prove that $\mathscr{A}$ is a partition of $\mathbb{R}$, and provide an explicit description of the (unique!) equivalence relation $\sim$ on $\mathbb{R}$ such that $\mathbb{R}/\sim = \mathscr{A}$.

The following lemma can be skipped over without grave consequences—it is a technical result with an extremely fiddly proof, but we will use it at a couple of points later in the book. It says that, given two partitioned sets, if we can pair up the sets in the partition, and pair up the elements in each pair of paired-up partitions, then we can pair up the elements of each set.

Lemma 5.2.29

Let X and Y be sets, let $\mathscr{A}$ be a partition of X and let $\mathscr{B}$ be a partition of Y. If there exists:

- A bijection $F : \mathscr{A} \to \mathscr{B}$; and
- For each $A \in \mathscr{A}$, a bijection $g_A : A \to F(A)$;

then there exists a bijection $h : X \to Y$.

Proof

Given $x \in X$, let A_x be the unique set in the partition $\mathscr{A}$ containing x as an element; likewise, given $y \in Y$, let B_y be the unique set in $\mathscr{B}$ containing y as an element. Note that these sets exist and are unique by Exercise 5.2.25.

Define $h : X \to Y$ by $h(x) = g_{A_x}(x)$ for all $x \in X$. To see that h is well-defined, note that for all $x \in X$ we have $g_{A_x} : A_x \to F(A_x)$, so x is in the domain of g_{A_x}; and

$$h(x) = g_{A_x}(x) \in F(A_x) \subseteq Y$$

Furthermore, given $x \in X$ we know that $B_{h(x)}$ is the unique set in $\mathscr{B}$ containing $h(x)$ as an element—since $F(A_x) \in \mathscr{B}$ and $h(x) \in F(A_x)$, it follows that $F(A_x) = B_{h(x)}$, and so $F^{-1}(B_{h(x)}) = A_x$.

Now define $k : Y \to X$ by $k(y) = g^{-1}_{F^{-1}(B_y)}(y)$ for all $y \in Y$. To see that k is well-defined, note that for all $y \in Y$ we have $g_{F^{-1}(B_y)} : F(F^{-1}(B_y)) \to F^{-1}(B_y)$, so that $g_{F^{-1}(B_y)} : B_y \to F^{-1}(B_y)$, meaning that y is in the domain of $g_{F^{-1}(B_y)}$; and

$$k(y) = g_{F^{-1}(B_y)}(y) \in F^{-1}(B_y) \subseteq X$$

Just like with our definition of h, we also see that for all $y \in Y$ we have $F^{-1}(B_y) = A_{k(y)}$.

To see that k is a left inverse for h, let $x \in X$—then

$$\begin{aligned}
k(h(x)) &= g^{-1}_{F^{-1}(B_{h(x)})}(h(x)) && \text{by definition of } k \\
&= g^{-1}_{A_x}(h(x)) && \text{since } F^{-1}(B_{h(x)}) = A_x \\
&= g^{-1}_{A_x}(g_{A_x}(x)) && \text{by definition of } h \\
&= x && \text{since } g^{-1}_{A_x} \circ g_{A_x} = \mathrm{id}_{A_x}
\end{aligned}$$

To see that k is right inverse for h, let $y \in Y$—then

$$\begin{aligned} h(k(y)) &= g_{A_{k(y)}}(k(y)) && \text{by definition of } h \\ &= g_{A_{k(y)}}(g^{-1}_{F^{-1}(B_y)}(y)) && \text{by definition of } k \\ &= g_{A_{k(y)}}(g^{-1}_{A_{k(y)}}(y)) && \text{since } F^{-1}(B_y) = A_{k(y)} \\ &= y && \text{since } g_{A_{k(y)}} \circ g^{-1}_{A_{k(y)}} = \mathrm{id}_{F(A_{k(y)})} \end{aligned}$$

So k is an inverse for h, as required. □

Exercise 5.2.30

Let X and Y be sets, let $\sim$ be an equivalence relation on X and let $\approx$ be an equivalence relation on Y. Assume that there is a bijection $p : X/\sim \to Y/\approx$, and for each equivalence class $E \in X/\sim$ there is a bijection $h_E : E \to p(E)$. Use Lemma 5.2.29 to prove that there is a bijection $h : X \to Y$.

The quotient function

We will now show that equivalence relations on a set X are essentially the same thing as *surjections* from X to another set.

Definition 5.2.31

Let X be a set and let $\sim$ be an equivalence relation on X. The **quotient function** for $\sim$ is the function $q_\sim : X \to X/\sim$ defined by $q(a) = [a]_\sim$ for each $a \in X$. That is, the quotient function sends each element of X to its $\sim$-equivalence class.

Example 5.2.32

Recall that, given $a \in \mathbb{Z}$, we have $[a]_2 = [0]_2$ if a is even, and $[a]_2 = [1]_2$ is a is odd. Thus the quotient function $q_2 : \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ can be viewed as telling us the parity of an integer.

Exercise 5.2.33

Let $n \in \mathbb{Z}$ with $n \neq 0$. Describe the quotient function $q_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ in terms of remainders.

Exercise 5.2.34

Let $\sim$ be an equivalence relation on a set X. Prove that the quotient function $q_\sim : X \to X/\sim$ is surjective.

The theorem we prove next can be viewed as the converse to Exercise 5.2.34. It proves that every surjection 'is' a quotient function, in the sense that given any surjection $p : X \to Z$, we can view Z as a quotient of X by a suitably-defined equivalence relation, and then p 'is' the corresponding quotient function.

Theorem 5.2.35

Let X be a set. Then for every set Z and every surjection $p : X \to Z$, there exist a unique equivalence relation $\sim$ on X and bijection $f : X/\sim \to Z$ such that $f([x]) = p(x)$ for all $x \in X$.

Proof

Let Z be a set and $p : X \to Z$ be a surjection.

- (**Existence**) Define a relation $\sim$ on X for all $x, y \in X$ by letting $x \sim y$ if and only if $p(x) = p(y)$. Then $\sim$ is an equivalence relation by Exercise 5.2.4.

  Moreover, given $x \in X$, we have

  $$[x]_\sim = \{y \in X \mid p(x) = p(y)\} = p^{-1}[\{p(x)\}]$$

  So define $f : X/\sim \to Z$ by letting $f([x]_\sim) = p(x)$. Then f is well-defined, since if $[x]_\sim = [y]_\sim$ then $x \sim y$, so that $p(x) = p(y)$.

  Furthermore, f is a bijection:

  - ◇ (**Injectivity**) Let $[x]_\sim, [y]_\sim \in X/\sim$ and assume $f([x]_\sim) = f([y]_\sim)$. Then $p(x) = p(y)$, so that $x \sim y$, and hence $[x]_\sim = [y]_\sim$.
  - ◇ (**Surjectivity**) Let $z \in Z$. Since p is a surjection, there is some $x \in X$ such that $p(x) = z$. But then $f([x]_\sim) = p(x) = z$.

  So we have established that there exist an equivalence relation $\sim$ on X and a bijection $f : X/\sim \to Z$ such that $f([x]_\sim) = p(x)$ for all $x \in X$.

- (**Uniqueness**) Suppose $\approx$ is another equivalence relation on X and that $g : X/\approx \to Z$ is a bijection such that $g([x]_\approx) = p(x)$ for all $x \in X$. We prove that $\sim = \approx$, and then that $g = f$, so that $\sim$ and f are unique.

  So let $x, y \in X$. Then

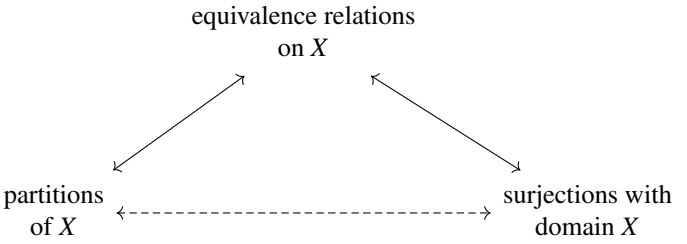
  $$\begin{aligned} x \sim y &\Leftrightarrow p(x) = p(y) && \text{by definition of } \sim \\ &\Leftrightarrow g([x]_\approx) = g([y]_\approx) && \text{by definition of } g \\ &\Leftrightarrow [x]_\approx = [y]_\approx && \text{since } g \text{ is bijective} \\ &\Leftrightarrow x \approx y && \text{by Exercise 5.2.27} \end{aligned}$$

  It follows that $\sim = \approx$, and then for all $x \in X$ we have

  $$f([x]_\sim) = p(x) = g([x]_\approx) = g([x]_\sim)$$

  so that $f = g$, as required. □

In light of Theorem 5.2.35, we have now established the equivalence of three notions for a given set X:



Exercise 5.2.36

Give an explicit description of the dashed arrow in the above diagram. That is, describe the correspondence between partitions of a set X and surjections whose domain is X.

Section 5.E

Chapter 5 exercises

Properties of relations

5.E.1. For each of the eight subsets

$$P \subseteq \{\text{reflexive}, \text{symmetric}, \text{transitive}\}$$

find a relation satisfying (only) the properties in P.

5.E.2. Prove that if R is a symmetric, antisymmetric relation on a set X, then it is a subrelation of the equality relation—that is, $\text{Gr}(R) \subseteq \text{Gr}(=)$.

5.E.3. A relation R on a set X is **left-total** if for all $x \in X$, there exists some $y \in X$ such that $x\,R\,y$. Prove that every left-total, symmetric, transitive relation is reflexive.

Equivalence relations

Definition 5.E.4

Let R be a relation on a set X and let $f : X \to Y$ be a function. The **transport** of R along f is the relation S on Y defined for $c, d \in Y$ by letting $c\,S\,d$ if and only if there exist $a, b \in X$ such that $f(a) = c$, $f(b) = d$ and $a\,R\,b$. That is

$$\text{Gr}(S) = \{(f(a), f(b)) \mid a, b \in X,\ a\,R\,b\}$$

5.E.5. Let X and Y be sets and let $f : X \to Y$. Prove that if $\sim$ is an equivalence relation on X, then the transport of $\sim$ along f is symmetric and transitive. Under what condition(s) on f is the transport of $\sim$ along f an equivalence relation on Y?

Definition 5.E.6

Let R be any relation on a set X. The **equivalence relation generated by** R is the relation $\sim_R$ on X defined as follows. Given $x, y \in X$, say $x \sim_R y$ if and only if for some $k \in \mathbb{N}$ there is a sequence $(a_0, a_1, \ldots, a_k)$ of elements of X such that $a_0 = x$, $a_k = y$ and, for all $i \in [k]$, either $a_i R a_{i+1}$ or $a_{i+1} R a_i$.

5.E.7. Fix $n \in \mathbb{Z}$ and let R be the relation on $\mathbb{Z}$ defined by $x R y$ if and only if $y = x + n$. Prove that $\sim_R$ is the relation of congruence modulo n.

5.E.8. Let X be a set and let R be the subset relation on $\mathscr{P}(X)$. Prove that $U \sim_R V$ for all $U, V \subseteq X$.

5.E.9. Let X be a set, fix two distinct elements $a, b \in X$, and define a relation R on X by declaring $a R b$ only—that is, for all $x, y \in X$, we have $x R y$ if and only if $x = a$ and $y = b$. Prove that the relation $\sim_R$ is defined by $x \sim_R y$ if and only if either $x = y$ or $\{x, y\} = \{a, b\}$.

In Questions 5.E.10 to 5.E.13, let R be a relation on a set X, and let $\sim_R$ be the equivalence relation generated by R (as in Definition 5.E.6). In these questions, you will prove that $\sim_R$ is the 'smallest' equivalence relation extending R.

5.E.10. Prove that $\sim_R$ is an equivalence relation on X.

5.E.11. Prove that $xRy \Rightarrow x \sim_R y$ for all $x, y \in X$.

5.E.12. Prove that if $\approx$ is any equivalence relation on X and $xRy \Rightarrow x \approx y$ for all $x, y \in X$, then $x \sim_R y \Rightarrow x \approx y$ for all $x, y \in X$.

5.E.13. Prove that if R is an equivalence relation, then $\sim_R = R$.

5.E.14. This question requires some familiarity with concepts from calculus. Let $\mathscr{C}^1(\mathbb{R})$ denote the set of all differentiable functions $F : \mathbb{R} \to \mathbb{R}$ whose derivative $F' : \mathbb{R} \to \mathbb{R}$ is continuous, and define a relation $\sim$ on $\mathscr{C}^1(\mathbb{R})$ by letting $F \sim G$ if and only if $F' = G'$.

(a) Prove that $\sim$ is an equivalence relation on $\mathscr{C}^1(\mathbb{R})$.

(b) Let $F \in \mathscr{C}^1(\mathbb{R})$. Prove that

$$[F]_\sim = \{G \in \mathscr{C}^1(\mathbb{R}) \mid \exists c \in \mathbb{R}, \forall x \in \mathbb{R}, G(x) = F(x) + c\}$$

(c) Interpret the notion of an indefinite integral $\int f(x)\,dx$ of a continuous function $f : \mathbb{R} \to \mathbb{R}$ in terms of $\sim$-equivalence classes.

True–False questions

In Questions 5.E.15 to 5.E.21, determine (with proof) whether the statement is true or false.

5.E.15. Every reflexive relation is symmetric.

5.E.16. There is a unique relation on $\mathbb{N}$ that is reflexive, symmetric and antisymmetric.

5.E.17. Every relation that is not symmetric is antisymmetric.

5.E.18. Every symmetric, antisymmetric relation is reflexive.

5.E.19. The empty set is the graph of a relation on $\mathbb{N}$.

5.E.20. For all sets X, the graph of a function $X \to X$ is the graph of a relation on X.

5.E.21. For all sets X, the graph of a relation on X is the graph of a function $X \to X$.

Always–Sometimes–Never questions

In Questions 5.E.22 to 5.E.23, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

5.E.22. Let $\sim$ be an equivalence relation on a set X. Then the relation $\approx$ on $\mathscr{P}(X) \setminus \{\varnothing\}$, defined by $A \approx B$ if and only if $x \sim y$ for all $x \in A$ and $y \in B$, is an equivalence relation.

5.E.23. Let $\sim$ be an equivalence relation on a set X. Then the relation $\approx$ on $\mathscr{P}(X) \setminus \{\varnothing\}$, defined by $A \approx B$ if and only if $x \sim y$ for some $x \in A$ and $y \in B$, is an equivalence relation.

Chapter 6

Finite and infinite sets

Section 6.1

Finite sets

As its title suggests, this section is all about exploring the properties of finite sets, and to do this we must first define what we mean by 'finite'. We certainly know a finite set when we see one—for example:

- The set $\{\text{red}, \text{orange}, \text{yellow}, \text{green}, \text{blue}, \text{purple}\}$ is finite.
- The set $[0,1]$ has infinitely many elements, but it has finite length.
- The set $[0,\infty)$ has infinitely many elements and has infinite length.
- The set $\mathscr{P}(\mathbb{N})$ has infinitely many elements, but has no notion of 'length' to speak of.
- The empty set $\varnothing$ is finite.

If we are to make a definition of 'finite set', we must first figure out what the finite sets above have in common but the infinite sets do not.

It is difficult to define 'finite' without being imprecise. A first attempt at a definition might be something like the following:

A set X is finite if the elements of X don't go on forever.

This is good intuition, but isn't good enough as a mathematical definition, because 'go on' and 'forever' are not precise terms (unless they themselves are defined). So let's try to make this more precise:

A set X is finite if the elements of X can be listed one by one in such a way that the list has both a start and an end.

This is better but is still not entirely precise—it is not entirely clear what is meant by 'listed one by one'. But we can make this precise. Let's label the positions in our list with natural numbers. To list the elements of X means to specify an element of X to go in position 0, then another element of X to go in position 1, and so on, until we eventually run out of elements of X—the fact that we eventually run out of elements is exactly what we want to capture by our definition of 'finite'.

If the listed elements of X occupy positions $0, 1, \ldots, n-1$, this means that we have effectively paired up the elements of $[n]$ (the list positions) with the elements of X (the set being listed). Since 'pairing up' really means 'finding a bijection', we are now ready to define what it means for a set to be finite.

Definition 6.1.1

A set X is **finite** if there exists a bijection $f : [n] \to X$ for some $n \in \mathbb{N}$. The function f is called an **enumeration** of X. If X is not finite we say it is **infinite**.

Example 6.1.2

Let $X = \{\text{red}, \text{orange}, \text{yellow}, \text{green}, \text{blue}, \text{purple}\}$, and define $f : [6] \to X$ by

$$\begin{array}{lll} f(0) = \text{red} & f(1) = \text{orange} & f(2) = \text{yellow} \\ f(3) = \text{green} & f(4) = \text{blue} & f(5) = \text{purple} \end{array}$$

The function f is evidently a bijection, since each element of X can be expressed uniquely as $f(k)$ for some $k \in [6]$. So X is finite.

Exercise 6.1.3

Prove that $[n]$ is finite for each $n \in \mathbb{N}$.

Note that Exercise 6.1.3 implies, in particular, that $\varnothing$ is finite, since $\varnothing = [0]$.

The size of a finite set

Whilst it might sometimes be useful just to know *that* a set is finite, it will be even more useful to know how many elements it has. This quantity is called the *size* of the set. Intuitively, the size of the set should be the length of the list of its elements, but for this to be well-defined, we first need to know that the number of elements in the list is independent of the way we choose to list them.

The 'list of elements' of a finite set X is the bijection $[n] \to X$ given by Definition 6.1.1, and n is the length of the list, this means that we need to prove that if $[m] \to X$ and $[n] \to X$ are bijections, then $m = n$. This will be Theorem 6.1.8.

To be able to prove this, we must first prove some technical results involving functions that we will use in the proof.

Lemma 6.1.4

Let X and Y be sets, let $f : X \to Y$ be an injection and let $a \in X$. Then there is an injection $\tilde{f} : X \setminus \{a\} \to Y \setminus \{f(a)\}$ defined by $\tilde{f}(x) = f(x)$ for all $x \in X \setminus \{a\}$. Moreover if f is a bijection $X \to Y$, then $\tilde{f}$ is a bijection $X \setminus \{a\} \to Y \setminus \{f(a)\}$.

Proof

Note that $\tilde{f}$ is well-defined by injectivity of f: indeed, if we had $f(x) = f(a)$ for some $x \in X \setminus \{a\}$, then that would imply $x = a \in X \setminus \{a\}$, which is a contradiction.

To see that $\tilde{f}$ is injective, let $x, y \in X \setminus \{a\}$ and $\tilde{f}(x) = \tilde{f}(y)$. Then $f(x) = f(y)$, so that $x = y$ by injectivity of f.

Now suppose that f is a bijection. Then f has an inverse $g : Y \to X$, which is itself a bijection (and hence an injection). Repeating the above argument with g yields an injection $\tilde{g} : Y \setminus \{f(a)\} \to X \setminus \{a\}$ defined by $\tilde{g}(y) = g(y)$ for all $y \in Y \setminus \{f(a)\}$. But then for all $x \in X \setminus \{a\}$ and all $y \in Y \setminus \{f(a)\}$ we have

$$\tilde{g}(\tilde{f}(x)) = g(f(x)) = x \quad \text{and} \quad \tilde{f}(\tilde{g}(y)) = f(g(y))$$

so that $\tilde{g}$ is an inverse for $\tilde{f}$. Hence if f is a bijection, then $\tilde{f}$ is a bijection. □

A consequence of Lemma 6.1.4 is the next useful result, which we will use in our proof that every set has a unique 'size'.

Lemma 6.1.5
Let X be an inhabited set. There is a bijection $X \setminus \{a\} \to X \setminus \{b\}$ for all $a, b \in X$.

Proof
Define $f : X \to X$ be the function that swaps a and b and leaves all other elements of X fixed—that is, we define

$$f(x) = \begin{cases} b & \text{if } x = a \\ a & \text{if } x = b \\ x & \text{if } x \notin \{a, b\} \end{cases}$$

Note that $f(f(x)) = x$ for all $x \in X$, so f is a bijection—it is its own inverse! Since $f(a) = b$, it follows from Lemma 6.1.4 that the function $\tilde{f} : X \setminus \{a\} \to X \setminus \{b\}$ defined by $\tilde{f}(x) = f(x)$ for all $x \in X \setminus \{a\}$ is a bijection, as required. □

Theorem 6.1.6
Let $m, n \in \mathbb{N}$.

(a) If there exists an injection $f : [m] \to [n]$, then $m \leqslant n$.

(b) If there exists a surjection $g : [m] \to [n]$, then $m \geqslant n$.

(c) If there exists a bijection $h : [m] \to [n]$, then $m = n$.

Proof of (a)
We will prove by induction on m that, for all $m, n \in \mathbb{N}$, if there exists an injection $[m] \to [n]$, then $m \leqslant n$.

- (**Base case**) We need to prove that, for all $n \in \mathbb{N}$ if there exists an injection $[0] \to [n]$, then $0 \leqslant n$. This is automatically true, since $0 \leqslant n$ for all $n \in \mathbb{N}$.
- (**Induction step**) Fix $m \in \mathbb{N}$ and suppose that, for all $n \in \mathbb{N}$, if there exists an injection $[m] \to [n]$, then $m \leqslant n$.

  Now let $n \in \mathbb{N}$ and suppose that there is an injection $f : [m+1] \to [n]$. We need to prove that $m + 1 \leqslant n$.

  First note that $n > 0$. Indeed, since $m + 1 > 0$, we have $0 \in [m+1]$, and so $f(0) \in [n]$. This means that $[n]$ is inhabited, and so $n > 0$. In particular, $n - 1 \in \mathbb{N}$ and so the set $[n-1]$ is well-defined. It suffices to prove that $m \leqslant n - 1$.

  Let $a = f(m) \in [n]$. Then:

  ◇ By Lemma 6.1.4 there is an injection $\tilde{f} : [m] = [m+1] \setminus \{m\} \to [n] \setminus \{a\}$.

  ◇ By Lemma 6.1.5 there is a bijection $g : [n] \setminus \{a\} \to [n] \setminus \{n-1\} = [n-1]$.

  Composing these two functions gives an injection $g \circ \tilde{f} : [m] \to [n-1]$. But then $m \leqslant n - 1$ by the induction hypothesis, and so $m + 1 \leqslant n$ as required.

The result now follows by induction. □

Exercise 6.1.7
Prove parts (b) and (c) of Theorem 6.1.6.

Phew! That was fun. With these technical results proved, we can now prove the theorem we needed for the size of a finite set to be well-defined.

Theorem 6.1.8 · *Uniqueness of size*
Let X be a finite set and let $f : [m] \to X$ and $g : [n] \to X$ be enumerations of X, where $m, n \in \mathbb{N}$. Then $m = n$.

Proof
Since $f : [m] \to X$ and $g : [n] \to X$ are bijections, the function $g^{-1} \circ f : [m] \to [n]$ is a bijection by Exercises 3.2.20 and 3.2.45. Hence $m = n$ by Theorem 6.1.6(c). □

As we mentioned above, Theorem 6.1.8 tells us that any two ways of listing (enumerating) the elements of a finite set yield the same number of elements. We may now make the following definition.

Definition 6.1.9
Let X be a finite set. The **size** (or **cardinality**) of X, written $|X|$, is the unique natural number n for which there exists a bijection $[n] \to X$.

Example 6.1.10
Example 6.1.2 showed that $|\{\text{red}, \text{orange}, \text{yellow}, \text{green}, \text{blue}, \text{purple}\}| = 6$, and provided the proof was correct, Exercise 6.1.3 showed that $|[n]| = n$ for all $n \in \mathbb{N}$; in particular, $|\varnothing| = 0$.

Example 6.1.11
Fix $n \in \mathbb{N}$ and let $X = \{a \in \mathbb{Z} \mid -n \leqslant a \leqslant n\}$. There is a bijection $f : [2n+1] \to X$ defined by $f(k) = k - n$ for all $k \in [2n+1]$. Indeed:

- f **is well-defined.** We need to prove $f(k) \in X$ for all $k \in [2n+1]$. Well given $k \in [2n+1]$, we have $0 \leqslant k \leqslant 2n$, and so
$$-n = 0 - n \leqslant \underbrace{k - n}_{=f(k)} \leqslant 2n - n = n$$
so that $f(k) \in X$ as claimed.
- f **is injective.** Let $k, \ell \in [2n+1]$ and assume $f(k) = f(\ell)$. Then $k - n = \ell - n$, and so $k = \ell$.
- f **is surjective.** Let $a \in X$ and define $k = a + n$. Then
$$0 = (-n) + n \leqslant \underbrace{a + n}_{=k} \leqslant n + n = 2n < 2n + 1$$
and so $k \in [2n+1]$, and moreover $f(k) = (a+n) - n = a$.

Since f is a bijection, we have $|X| = 2n + 1$ by Definition 6.1.9.

Exercise 6.1.12
Let X be a finite set.

(a) Prove that for all $a \in X$, the set $X \setminus \{a\}$ is finite and $|X \setminus \{a\}| = |X| - 1$; and

(b) Prove that for all $b \notin X$, the set $X \cup \{b\}$ is finite and $|X \cup \{b\}| = |X| + 1$.

What happens in part (a) if we do not require $a \in X$? What happens in part (b) if we do not require $b \notin X$?

Comparing the sizes of finite sets

When we used dots and stars to motivate the definitions of injective and surjective functions at the beginning of Section 3.2, we suggested the following intuition:

- If there is an injection $f : X \to Y$, then X has 'at most as many elements as Y'; and
- If there is a surjection $g : X \to Y$, then X has 'at least as many elements as Y'.

We are now in a position to prove this, at least when X and Y are finite. The following theorem is a generalisation of Theorem 6.1.6.

Theorem 6.1.13
Let X and Y be sets.

(a) If Y is finite and there is an injection $f : X \to Y$, then X is finite and $|X| \leqslant |Y|$.

(b) If X is finite and there is a surjection $f : X \to Y$, then Y is finite and $|X| \geqslant |Y|$.

(c) If one of X or Y is finite and there is a bijection $f : X \to Y$, then X and Y are both finite and $|X| = |Y|$.

Proof of (a)

We prove by induction that, for all $n \in \mathbb{N}$, if Y is a finite set of size n and there is an injection $f : X \to Y$, then X is finite and $|X| \leqslant n$.

- (**Base case**) Let Y be a finite set of size 0—that is, Y is empty. Suppose there is an injection $f : X \to Y$. If X is inhabited, then there exists an element $a \in X$, so that $f(a) \in Y$. This contradicts emptiness of Y, so that X must be empty. Hence $|X| = 0 \leqslant 0$, as required.
- (**Induction step**) Fix $n \in \mathbb{N}$ and assume that, if Z is a finite set of size n and there is an injection $f : X \to Z$, then X is finite and $|X| \leqslant n$.

  Fix a finite set Y of size $n+1$ and an injection $f : X \to Y$. We need to prove that X is finite and $|X| \leqslant n+1$.

  If X is empty, then $|X| = 0 \leqslant n+1$ as required. So assume that X is inhabited, and fix an element $a \in X$.

By Lemma 6.1.4, there is an injection $\tilde{f} : X \setminus \{a\} \to Y \setminus \{f(a)\}$. Moreover, by Exercise 6.1.12, $Y \setminus \{f(a)\}$ is finite and $|Y \setminus \{f(a)\}| = (n+1) - 1 = n$.

Applying the induction hypothesis with $Z = Y \setminus \{f(a)\}$ tells us that, $X \setminus \{a\}$ is finite and $|X \setminus \{a\}| \leqslant n$. But $|X \setminus \{a\}| = |X| - 1$ by Exercise 6.1.12, and so $|X| \leqslant n+1$, as required.

The result now follows by induction. □

Exercise 6.1.14
Prove parts (b) and (c) of Theorem 6.1.13.

Theorem 6.1.13 suggests the following strategies for comparing the sizes of finite sets:

Strategy 6.1.15 · *Comparing the sizes of finite sets*
Let X and Y be finite sets.

(a) In order to prove that $|X| \leqslant |Y|$, it suffices to find an injection $X \to Y$.

(b) In order to prove that $|X| \geqslant |Y|$, it suffices to find a surjection $X \to Y$.

(c) In order to prove that $|X| = |Y|$, it suffices to find a bijection $X \to Y$.

Strategy (c) is commonly known as **bijective proof**.

Closure properties of finite sets

We now use Strategy 6.1.15 to prove some *closure properties* of finite sets—that is, operations we can perform on finite sets to ensure that the result remains finite.

Exercise 6.1.16
Let X be a finite set. Prove that every subset $A \subseteq X$ is finite and $|A| \leqslant |X|$.

Exercise 6.1.17
Let X and Y be finite sets. Prove that $X \cap Y$ is finite.

Proposition 6.1.18
Let X and Y be finite sets. Then $X \cup Y$ is finite, and moreover

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

Proof
We will prove this in the case when X and Y are disjoint. The general case, when X and Y are not assumed to be disjoint, will be Exercise 6.1.19.

Let $m = |X|$ and $n = |Y|$, and let $f : [m] \to X$ and $g : [n] \to Y$ be bijections.

Since X and Y are disjoint, we have $X \cap Y = \varnothing$. Define $h : [m+n] \to X \cup Y$ as follows; given $k \in [m+n]$, let

$$h(k) = \begin{cases} f(k) & \text{if } k < m \\ g(k-m) & \text{if } k \geqslant m \end{cases}$$

Note that h is well-defined: the cases $k < m$ and $k \geqslant m$ are mutually exclusive, they cover all possible cases, and $k - m \in [n]$ for all $m \leqslant k < m+n$ so that $g(k-m)$ is defined. It is then straightforward to check that h has an inverse $h^{-1} : X \cup Y \to [m+n]$ defined by

$$h^{-1}(z) = \begin{cases} f^{-1}(z) & \text{if } z \in X \\ g^{-1}(z) + m & \text{if } z \in Y \end{cases}$$

Well-definedness of h^{-1} relies fundamentally on the assumption that $X \cap Y = \varnothing$, as this is what ensures that the cases $x \in X$ and $x \in Y$ are mutually exclusive.

Hence $|X \cup Y| = m + n = |X| + |Y|$, which is as required since $|X \cap Y| = 0$. □

Exercise 6.1.19

The following steps complete the proof of Proposition 6.1.18:

(a) Given sets A and B, prove that the sets $A \times \{0\}$ and $B \times \{1\}$ are disjoint, and find bijections $A \to A \times \{0\}$ and $B \to B \times \{1\}$. Write $A \sqcup B$ (LaTeX code: `\sqcup`) to denote the set $(A \times \{0\}) \cup (B \times \{1\})$. The set $A \sqcup B$ is called the **disjoint union** of A and B.

(b) Prove that, if A and B are finite then $A \sqcup B$ is finite and

$$|A \sqcup B| = |A| + |B|$$

(c) Let X and Y be sets. Find a bijection

$$(X \cup Y) \sqcup (X \cap Y) \to X \sqcup Y$$

(d) Complete the proof of Proposition 6.1.18—that is, prove that if X and Y are finite sets, not necessarily disjoint, then $X \cup Y$ is finite and

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

Exercise 6.1.20

Let X be a finite set and let $U \subseteq X$. Prove that $X \setminus U$ is finite, and moreover $|X \setminus U| = |X| - |U|$. How does this result change if U is not required to be a subset of X?

Exercise 6.1.21

Let $m, n \in \mathbb{N}$. Prove that $|[m] \times [n]| = mn$.

Proposition 6.1.22
Let X and Y be finite sets. Then $X \times Y$ is finite, and moreover

$$|X \times Y| = |X| \cdot |Y|$$

Proof
Let X and Y be finite sets, let $m = |X|$ and $n = |Y|$, and let $f : [m] \to X$ and $g : [n] \to Y$ be bijections. Define a function $h : [m] \times [n] \to X \times Y$ by

$$h(k, \ell) = (f(k), g(\ell))$$

for each $k \in [m]$ and $\ell \in [n]$. It is easy to see that this is a bijection, with inverse defined by

$$h^{-1}(x, y) = (f^{-1}(x), g^{-1}(y))$$

for all $x \in X$ and $y \in Y$. By Exercise 6.1.21 there is a bijection $p : [mn] \to [m] \times [n]$, and by Exercise 3.2.20 the composite $h \circ p : [mn] \to X \times Y$ is a bijection. Hence $|X \times Y| = mn$. □

In summary, we have proved that the property of finiteness is preserved by taking subsets, pairwise unions, pairwise intersections, pairwise cartesian products, and relative complements.

Infinite sets

We conclude this section by proving that not all sets are finite—specifically, we'll prove that $\mathbb{N}$ is infinite. *Intuitively* this seems extremely easy: of *course* $\mathbb{N}$ is infinite! But in mathematical practice, this isn't good enough: we need to use our definition of 'infinite' to prove that $\mathbb{N}$ is infinite. Namely, we need to prove that there is no bijection $[n] \to \mathbb{N}$ for any $n \in \mathbb{N}$. We will use Lemma 6.1.23 below in our proof.

Lemma 6.1.23
Every inhabited finite set of natural numbers has a greatest element.

Proof
We'll prove by induction on $n \geqslant 1$ that every subset $U \subseteq \mathbb{N}$ of size n has a greatest element.

- (**Base case**) Take $U \subseteq \mathbb{N}$ with $|U| = 1$. then $U = \{m\}$ for some $m \in \mathbb{N}$. Since m is the only element of U, it is certainly the greatest element of U!
- (**Induction step**) Fix $n \geqslant 1$ and suppose that every set of natural numbers of size n has a greatest element.

  Let $U \subseteq \mathbb{N}$ with $|U| = n + 1$. We wish to show that U has a greatest element.

  Since $|U| = n + 1 > 0$, we know that U is inhabited, so there is an element $a \in U$. By Exercise 6.1.12(a) we have $|U \setminus \{a\}| = |U| - 1 = n$, and so by the induction hypothesis, $U \setminus \{a\}$ has a greatest element $m \in U \setminus \{a\}$. Now either $a < m$ or $a > m$.

- ◇ If $a < m$, then m is the greatest element of U; and
- ◇ If $a > m$, then a is the greatest element of U.

In any case, we see that U has a greatest element. □

Lemma 6.1.23 is equivalent to the assertion that every inhabited subset of $\mathbb{N}$ with no greatest element is infinite. We can therefore deduce that $\mathbb{N}$ is infinite very easily, as we do in the next theorem.

Theorem 6.1.24
The set $\mathbb{N}$ is infinite.

Proof
We proceed by contradiction. Suppose $\mathbb{N}$ is finite. Then $|\mathbb{N}| = n$ for some $n \in \mathbb{N}$, and hence $\mathbb{N}$ is either empty (nonsense, since $0 \in \mathbb{N}$) or, by Lemma 6.1.23, it has a greatest element g. But $g+1 \in \mathbb{N}$ since every natural number has a successor, and $g+1 > g$, so this contradicts maximality of g. Hence $\mathbb{N}$ is infinite. □

Section 6.2

Countable and uncountable sets

In Section 6.1, we defined what it means for a set to be 'finite' in order to capture the idea that its elements can be listed in such a way that the list has a start and an end. We did so by observing that a list of the elements of a finite set is essentially the same thing as a bijection $f : [n] \to X$ for some $n \in \mathbb{N}$, with the element $f(k) \in X$ playing the role of the element in position k in the list.

We are now interested in *infinite* sets. We certainly can't expect a list of all the elements of an infinite set to end, so the question now is: can its elements be listed if we allow the list to be infinite? We will call such sets *countable sets*.

It is perhaps surprising that not every set is countable: some sets are 'too big' for their elements to be listed! We will prove that such *uncountable* sets exist later in this section.

The precise definition of what it means for a set X to be countable (Definition 6.2.1) captures the idea that we can list the elements of X one-by-one such that, even if the list goes on forever, each element of X appears *at some finite stage* in the list. The list might or might not be finite; it has a start, but it might not have an end.

To illustrate, consider the following list of the elements of $\mathbb{N}$:

$$0,\ 1,\ 2,\ 3,\ 4,\ \ldots,\ n,\ n+1,\ \ldots$$

The list does not end, since $\mathbb{N}$ is infinite (Theorem 6.1.24), but nevertheless every natural number appears at some finite stage along the list.

As another example, consider the set $\mathbb{Z}$ of all integers. We might wish to list the elements of $\mathbb{Z}$ in the usual way:

$$\ldots,\ -(n+1),\ -n,\ \ldots,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ \ldots,\ n,\ n+1,\ \ldots$$

This does not fulfil our criterion that the list must have a start: it is infinite in both directions. However, it is still possible to list them, by slotting the negative integers in-between the non-negative integers:

$$0,\ -1,\ 1,\ -2,\ 2,\ -3,\ 3,\ \ldots,\ -n,\ n,\ -(n+1),\ n+1,\ \ldots$$

This is not how we would usually think about listing the integers, but we have nonetheless found a way of doing it so that every integer appears at some finite stage on the list.

But specifying a list of the elements of an infinite set X, such that every element of X appears at some finite stage on the list, is equivalent to specifying a bijection $f : \mathbb{N} \to X$, where the element $f(k) \in X$ plays the role of the element in position k of the list—that is, the elements of X can be listed as

$$f(0),\ f(1),\ f(2),\ \ldots,\ f(n),\ f(n+1),\ \ldots$$

This motivates the following definition.

Definition 6.2.1

A set X is **countably infinite** if there exists a bijection $f : \mathbb{N} \to X$. The bijection f is called an **enumeration** of X. We say X is **countable** if it is finite or countably infinite.

Some authors prefer to use 'countable' to mean 'countably infinite', in which case they would say 'finite or countable' to mean 'countable'.

Example 6.2.2

The set $\mathbb{N}$ is countably infinite, since by Exercise 3.2.19, the identity function $\mathrm{id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}$ is a bijection. This enumeration yields the usual list of natural numbers

$$0,\ 1,\ 2,\ 3,\ \ldots,\ n,\ n+1,\ \ldots$$

Example 6.2.3

The function $f : \mathbb{Z} \to \mathbb{N}$ defined for $x \in \mathbb{Z}$ by

$$f(x) = \begin{cases} 2x & \text{if } x \geqslant 0 \\ -(2x+1) & \text{if } x < 0 \end{cases}$$

is a bijection. Indeed, it has an inverse is given by

$$f^{-1}(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ -\frac{x+1}{2} & \text{if } x \text{ is odd} \end{cases}$$

Hence the set of integers $\mathbb{Z}$ is countably infinite. The corresponding list of integers is given by

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

which is exactly the list we presented earlier! The fact that f is a bijection ensures that each integer appears on this list exactly once.

Exercise 6.2.4

Prove that the set of all positive integers is countably infinite.

Exercise 6.2.5

Prove that the set of all even natural numbers is countably infinite, and that the set of all odd natural numbers is countably infinite.

Since the inverse of a bijection is a bijection, we may also prove that a set X is countable by finding a bijection $X \to \mathbb{N}$.

Exercise 6.2.6

Prove that the function $p : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $p(x,y) = 2^x(2y+1) - 1$ is a bijection. Deduce that $\mathbb{N} \times \mathbb{N}$ is countable.

Closure properties of countable sets

Proving that a set is infinite by finding an explicit bijection with the set of natural numbers can be overly burdensome, so we will now develop some *closure properties* of countable sets. This will allow us to prove that a X set is countable by relating it to sets that we already know are countable, without having to find a bijection $f : \mathbb{N} \to X$ every time.

Proposition 6.2.7
Let $f : X \to Y$ be a bijection. Then X is countably infinite if and only if Y is countably infinite.

Proof
Suppose X is countably infinite. Then there is a bijection $g : \mathbb{N} \to X$. But then $f \circ g : \mathbb{N} \to Y$ is a composite of bijections, so is bijective, meaning that Y is countably infinite.

Conversely, suppose Y is countably infinite. Then there is a bijection $h : \mathbb{N} \to Y$. But then $f^{-1} \circ h : \mathbb{N} \to X$ is a composite of bijections, so is bijective, meaning that X is countably infinite. □

Exercise 6.2.8
Prove if X and Y are countably infinite sets, then $X \times Y$ is countably infinite.

Exercise 6.2.8 allows us to prove that the product of finitely many countably infinite sets are countably infinite.

Proposition 6.2.9 · *The product of finitely many countable sets is countable*
Let $n \geqslant 1$ and let $X_1, \ldots, X_n$ be countably infinite sets. Then the product $\prod_{i=1}^{n} X_i$ is countably infinite.

Proof
We proceed by induction on $n \geqslant 1$.

- (**Base case**) When $n = 1$ the assertion is trivial: if X_1 is countably infinite then $\prod_{i=1}^{1} X_i$ is equal to X_1, which is countably infinite.
- (**Induction step**) Fix $n \geqslant 1$ and suppose that for any sets $X_1, \ldots, X_n$, the product $\prod_{i=1}^{n} X_i$ is countably infinite. Fix sets $X_1, \ldots, X_{n+1}$. Then $\prod_{i=1}^{n} X_i$ is countably infinite by the induction hypothesis, and X_{n+1} is countably infinite by assumption, so by Exercise 6.2.8, the set

$$\left(\prod_{i=1}^{n} X_i\right) \times X_{n+1}$$

is countably infinite. But by Exercise 4.2.12 there is a bijection

$$\prod_{i=1}^{n+1} X_i \to \left(\prod_{i=1}^{n} X_i\right) \times X_{n+1}$$

and so by Exercise 6.2.8 we have that $\prod_{i=1}^{n+1} X_i$ is countably infinite, as required.

By induction, we're done. □

We often just want to know whether a set is *countable*, rather than countably infinite. This might be because we only seek an upper bound on how large the set is; or it might because we already know that the set is infinite, so proving that it is countable suffices.

The following theorem allows us to prove that a set X is countable—that is, finite or countably infinite—by either surjecting $\mathbb{N}$ onto X, or injecting X into $\mathbb{N}$. Using the intuition of Section 6.1, where we compared the sizes of finite sets using injections and surjections, this says that a set is countable if and only if X is at most as large as $\mathbb{N}$.

Theorem 6.2.10

Let X be an inhabited set. The following are equivalent:

(i) X is countable;

(ii) There exists an injection $g : X \to \mathbb{N}$;

(iii) There exists a surjection $h : \mathbb{N} \to X$.

Proof

We'll prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i).

- (i) ⇒ (ii). Suppose X is countable.
  - ◇ Suppose X is countably infinite, then there exists a bijection $f : \mathbb{N} \to X$; but then $f^{-1} : X \to \mathbb{N}$ is bijective (and therefore injective) by Exercise 3.2.45.
  - ◇ Suppose X is finite. Then there exists a bijection $f : [n] \to X$ for some $n \in \mathbb{N}$.
    Define $g : X \to \mathbb{N}$ by $g(x) = f^{-1}(x)$ for all $x \in X$. Note that g is well-defined since for all $x \in X$ we have $f^{-1}(x) \in [n] \subseteq \mathbb{N}$. Moreover for all $x_1, x_2 \in X$, if $g(x_1) = g(x_2)$, then $f^{-1}(x_1) = f^{-1}(x_2)$, and so applying f to both sides of the equation gives $x_1 = x_2$. Therefore g is injective, as required.
- (ii) ⇒ (iii). Suppose there exists an injection $g : X \to \mathbb{N}$. Since X is inhabited, g has a left inverse $h : \mathbb{N} \to X$ by Proposition 3.2.27, which is surjective by Exercise 3.2.28.
- (iii) ⇒ (i). Suppose that there exists a surjection $h : \mathbb{N} \to X$. If X is finite, then X is countable by definition of countability, so assume that X is infinite.

  Define a sequence of natural numbers $k_0, k_1, k_2, \ldots$ recursively by

  $$k_0 = 0 \quad \text{and} \quad k_{n+1} = \min\{k \in \mathbb{N} \mid h(k) \neq h(k_i) \text{ for all } i \leqslant n\}$$

  Intuitively, this is the sequence we obtain by starting from 0 and proceeding through the natural numbers, listing only those inputs to h that produce *new* values. Note that for all $n \in \mathbb{N}$, k_{n+1} is always defined since $h[\mathbb{N}] = X$, which is is infinite.

  Now define $f : \mathbb{N} \to X$ by $f(n) = h(k_n)$ for all $n \in \mathbb{N}$. Then:

- f **is injective.** Let $m, n \in \mathbb{N}$ and assume that $f(m) = f(n)$, so that $h(k_m) = h(k_n)$.

  If $m = 0$, then $n = 0$: indeed, if $n > 0$, then we'd have $h(k_n) \neq h(k_i)$ for all $i \leqslant n-1$ by the recursive definition of the sequence, and so $h(k_n) \neq h(k_0) = h(k_m)$, contradicting the assumption that $h(k_m) = h(k_n)$. Likewise, if $n = 0$, then $m = 0$.

  Suppose $m > 0$ and $n > 0$. Without loss of generality, assume $m \leqslant n$—otherwise, swap the roles of m and n in what follow. If $m > n$, then by the recursive definition of the sequence we have $h(k_m) \neq h(k_i)$ for all $i \leqslant m-1$, and so in particular $h(k_m) \neq h(k_n)$, contradicting our assumption that $h(k_m) = h(k_n)$. Therefore $m = n$.

  In all cases, we see that $m = n$, as required.

- f **is surjective.** Before we prove this, note first that the recursive definition of the sequence ensures that $k_{n+1} > k_n$ for all $n \in \mathbb{N}$, and therefore that $k_n \geqslant n$ for all $n \in \mathbb{N}$.

  Now, towards a contradiction, assume f is not surjective. Since h *is* surjective, every element of X is a value of h, so by the well-ordering principle there exists $n \in \mathbb{N}$ such that $h(n) \neq f(m)$ for all $m \in \mathbb{N}$, and therefore $h(n) \neq h(k_m)$ for all $m \in \mathbb{N}$.

  But then $h(n) \neq h(k_n)$, and since $n \leqslant k_n < k_{n+1}$, this contradicts minimality of k_{n+1}—indeed, $k = k_{n+1}$ is supposed to be least natural number such that $h(k) \neq h(k_i)$ for all $i \leqslant n$, but setting $k = n$ gives an even smaller such natural number.

So f is a bijection, as required. □

Exercise 6.2.11

Let X be a countably infinite set. In the proof of (iii) ⇒ (i) in Theorem 6.2.10, we gave an explicit construction of a bijection $f : \mathbb{N} \to X$ using a surjection $h : \mathbb{N} \to X$. Prove (ii) ⇒ (i) the same way—that is, explain how to construct a bijection $f : \mathbb{N} \to X$ using an injection $g : X \to \mathbb{N}$.

Theorem 6.2.10 allows us to deduce that sets are countable by finding injections with codomain $\mathbb{N}$ or surjections with domain $\mathbb{N}$, and without needing to find a bijection explicitly. It turns out we can do even better: we can replace $\mathbb{N}$ in Theorem 6.2.10 with any set that is already known to be countable.

Exercise 6.2.12

Let X be an inhabited set. The following are equivalent:

(i) X is countable;

(ii) There exists an injection $g : X \to C$ for some countable set C;

(iii) There exists a surjection $h : C \to X$ for some countable set C.

Exercise 6.2.12 is useful for proving the countability of many other sets: as we build up our repertoire of countable sets, all we need to do in order to prove a set X is countable is find a surjection from a set we already know is countable to X, or an injection from X into a set we already know is countable.

This proof technique yields an incredibly short proof of the following counterintuitive result, which can be interpreted to mean that there are exactly as many rational numbers as there are natural numbers.

Theorem 6.2.13
The set $\mathbb{Q}$ of rational numbers is countable.

Proof
Define a function $q : \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \to \mathbb{Q}$ by letting $q(a,b) = \frac{a}{b}$ for all $a, b \in \mathbb{Z}$ with $b \neq 0$.

By Example 6.2.3 and Exercise 6.2.8, the set $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is countable.

The function q is surjective by definition of $\mathbb{Q}$—indeed, to say $x \in \mathbb{Q}$ is precisely to say that $x = \dfrac{a}{b} = q(a,b)$ for some $(a,b) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$.

By Exercise 6.2.12, it follows that $\mathbb{Q}$ is countable. □

Exercise 6.2.14
Let X be a countable set. Prove that $\binom{X}{k}$ is countable for each $k \in \mathbb{N}$.

Theorem 6.2.15 · *The union of countably many countable sets is countable*
Let $\{X_n \mid n \in \mathbb{N}\}$ be a family of countable sets. Then the set X defined by

$$X = \bigcup_{n \in \mathbb{N}} X_n$$

is countable.

Proof
We may assume that the sets X_n are all inhabited, since the empty set does not contribute to the union.

For each $n \in \mathbb{N}$ there is a surjection $f_n : \mathbb{N} \to X_n$. Define $f : \mathbb{N} \times \mathbb{N} \to X$ by $f(m,n) = f_m(n)$ for all $m, n \in \mathbb{N}$. Then f is surjective: if $x \in X$ then $x \in X_m$ for some $m \in \mathbb{N}$. Since f_m is surjective, it follows that $x = f_m(n)$ for some $n \in \mathbb{N}$. But then $x = f(m,n)$. Since $\mathbb{N} \times \mathbb{N}$ is countable, it follows from Exercise 6.2.12 that X is countable. □

Example 6.2.16
Let X be a countable set. The set of all finite subsets of X is countable. Indeed, the set of all finite subsets of X is equal to $\bigcup_{k \in \mathbb{N}} \binom{X}{k}$, which is a union of countably many countable sets by Exercise 6.2.14, so is countable by Theorem 6.2.15.

Uncountable sets

We have now seen plenty of examples of countable sets. It is not immediately obvious that not every set is countable. How do we *know* that there are sets out there whose elements can't be listed?

We prove in Theorem 6.2.17 that there exists an uncountable set, namely the power set of the natural numbers. The proof is deceptively simple, but the implications are important.

Theorem 6.2.17
$\mathscr{P}(\mathbb{N})$ is uncountable.

Proof
We proved in Exercise 3.2.16 that no function $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$ is surjective. Specifically, given $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$, we can show that the element

$$B = \{k \in \mathbb{N} \mid k \notin f(k)\} \in \mathscr{P}(\mathbb{N})$$

is not equal to $f(n)$ for any $n \in \mathbb{N}$. It follows from Exercise 6.2.12(iii) that $\mathscr{P}(\mathbb{N})$ is uncountable.□

The argument used in Theorem 6.2.17 is an example of *Cantor's diagonal argument*, and is typical .

Theorem 6.2.18
Let X be a set, and assume that for every function $f : \mathbb{N} \to X$, there is:

(i) For each $n \in \mathbb{N}$, a logical formula $p_n(x)$ with free variable $x \in X$; and

(ii) An element $b \in X$;

such that $\forall n \in \mathbb{N}$, $[p_n(b) \Leftrightarrow \neg p_n(f(n))]$. Then X is uncountable.

Proof
We prove that no function $f : \mathbb{N} \to X$ is surjective. So let $f : \mathbb{N} \to X$ be an arbitrary function and let $p_n(x)$ and $b \in X$ be as in the statement of the theorem.

To see that f is not surjective, assume towards a contradiction that $b = f(k)$ for some $k \in \mathbb{N}$.

- If $p_k(b)$ is true, then $p_k(f(k))$ is true since $b = f(k)$, and $\neg p_k(f(k))$ is true since $p_k(b) \Rightarrow \neg p_k(f(k))$. This is a contradiction.
- If $\neg p_k(b)$ is true, then $\neg p_k(f(k))$ is false since $b = f(k)$, and so $p_k(b)$ is true since $[\neg p_k(f(k))] \Rightarrow p_k(b)$. This is a contradiction.

In both cases we arrive at a contradiction. So $b \neq f(k)$ for any $k \in \mathbb{N}$, and so f is not surjective.

Hence X is uncountable by Exercise 6.2.12(iii). □

Strategy 6.2.19 · *Cantor's diagonal argument*
In order to prove that a set X is uncountable, it suffices to prove that no function $f : \mathbb{N} \to X$ is surjective using the following argument: given a function $f : \mathbb{N} \to X$, find an element $b \in X$ that 'disagrees' with all the values of f, in the sense that for each $n \in \mathbb{N}$, b and $f(n)$ disagree about something to do with n.

Example 6.2.20

In the proof of Theorem 6.2.17, we proved that $\mathscr{P}(\mathbb{N})$ is uncountable as follows. Given $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$, we needed to find an element $B \in \mathscr{P}(\mathbb{N})$—that is, a subset $B \subseteq \mathbb{N}$—that 'disagreed' with each value $f(n)$ about something involving n.

Keep in mind that, for each $n \in \mathbb{N}$, the element $f(n) \in \mathscr{P}(\mathbb{N})$ is a subset of $\mathbb{N}$, so it makes sense to ask whether n is an element of $f(n)$ or not.

With this in mind, for each $n \in \mathbb{N}$, we forced B to disagree with $f(n)$ about whether n is an element. That is

$$n \in B \quad \Leftrightarrow \quad n \notin f(n)$$

In the language of Theorem 6.2.18, the statement $p_n(U)$ (for $n \in \mathbb{N}$ and $U \in \mathscr{P}(\mathbb{N})$ is the statement '$n \in U$'; for each $n \in \mathbb{N}$, this statement is true with $U = B$ if and only if it is false with $U = f(n)$.

The definition $B = \{k \in \mathbb{N} \mid k \notin f(k)\}$ forces $n \in B \Leftrightarrow n \notin f(n)$ to be true for all $n \in \mathbb{N}$, and so Cantor's diagonal argument applies.

Exercise 6.2.21

Use Cantor's diagonal argument to prove that the set $\mathbb{N}^{\mathbb{N}}$ of all functions $\mathbb{N} \to \mathbb{N}$ is uncountable.

Using Cantor's diagonal argument every time we want to prove that a set is uncountable can be cumbersome. However, just like with, Theorem 6.2.10, we can prove that sets are uncountable by relating them to other sets that we already know are uncountable.

Theorem 6.2.22

Let X be a set. The following are equivalent:

(i) X is uncountable;

(ii) There exists a surjection $f : X \to U$ for some uncountable set U;

(iii) There exists an injection $f : U \to X$ for some uncountable set U.

Proof

For (i)⇒(ii) and (i)⇒(iii), take $U = X$ and $f = \mathrm{id}_X$.

For (ii)⇒(i), suppose that there is a surjection $f : X \to U$ for some uncountable set U. If X were countable, then there would be a surjection $g : \mathbb{N} \to X$ by Theorem 6.2.10(ii); but then $f \circ g : \mathbb{N} \to U$ would be a surjection, contradicting the fact that U is uncountable.

For (iii)⇒(i), suppose there is an injection $f : U \to X$ for some uncountable set U. If X were countable, then there would be an injection $g : X \to \mathbb{N}$ by Theorem 6.2.10(iii); but then $g \circ f : U \to \mathbb{N}$ would be an injection, contradicting the fact that U is uncountable. □

Proposition 6.2.23

The set $[2]^{\mathbb{N}}$ of all functions $\mathbb{N} \to [2]$ is uncountable.

Proof
Define $F : \mathscr{P}(\mathbb{N}) \to [2]^{\mathbb{N}}$ for $U \subseteq \mathbb{N}$ by letting $F(U) = \chi_U : \mathbb{N} \to [2]$ be the characteristic function of U (Definition 3.1.23).

Given $U, V \subseteq \mathbb{N}$, we have

$$F(U) = F(V) \quad \Rightarrow \quad \chi_U = \chi_V \quad \Rightarrow \quad U = V$$

by Theorem 3.1.25. But then F is injective, and $\mathscr{P}(\mathbb{N})$ is uncountable by Cantor's theorem, so that $[2]^{\mathbb{N}}$ is uncountable by Theorem 6.2.22. □

It so happens that the function F defined in the proof of Proposition 6.2.23 is *bijective*, not just injective; however, we did not prove this, since it was not needed in our proof.

Exercise 6.2.24
Prove that if a set X has an uncountable subset, then X is uncountable.

Exercise 6.2.25
Let X be a set. Prove that $\mathscr{P}(X)$ is either finite or uncountable.

Detecting countability

We have now seen several examples of countable and uncountable sets. But being able to prove that a set is countable, or that a set is uncountable, only gets us so far—it will be beneficial to gain some intuition for how to detect whether a set is countable or uncountable.

We now work towards proving the following heuristic:

A set is countable if each of its elements has a finite description.

Before we make this precise, let's see this heuristic in action. Consider some of the sets that we have already seen are countable:

- The set $\mathbb{N}$ is countable; every natural number can be expressed as a finite string of digits, so has a finite description—for example, 7 or 15213.
- The set $\mathbb{Z}$ is countable; every integer has a finite description as a finite string of digits, possibly modified by a minus sign—for example, 7 or -15213.
- The set $\mathbb{Q}$ is countable; every rational number has a finite description as a pair of integers (which themselves have finite descriptions), one atop the other, separated by a horizontal line—for example, $\frac{7}{1}$ or $\frac{-15213}{21128}$.
- The set of all *finite* subsets of $\mathbb{N}$ is countable; every finite subset of $\mathbb{N}$ has a finite expression in list notation by writing an open curly bracket '{', followed by a finite number of natural numbers (which themselves have finite expressions) separated by commas, and finally writing a closed curly bracket '}'—for example $\{15, 0, 281\}$ or $\{\}$.

Now consider some of the sets that we know (or have been told) are uncountable:

- The set $\mathscr{P}(\mathbb{N})$ is uncountable; intuitively, the infinite subsets of $\mathbb{N}$ cannot be expressed finitely in a uniform way.
- The set $\mathbb{R}$ is uncountable; intuitively, there is no uniform way of finitely describing real numbers—for example, decimal expansions are no good since they are infinite.

Definition 6.2.26

Let Σ be a set. A **word** over Σ is an expression of the form $a_0 a_1 \dots a_{n-1}$, where $n \in \mathbb{N}$ and $a_i \in \Sigma$ for all $i \in [n]$. The natural number n is called the **length** of the word. The unique word of length 0 is called the **empty word**, and is denoted by ε (LaTeX code: `\varepsilon`).

The set Σ is called the **alphabet**. The set of all words over Σ is denoted by Σ^* (LaTeX code: `\Sigma^*`); the operation $(-)^*$ is called the **Kleene star**.

Formally, Σ^* is defined to be the union $\bigcup_{n\in\mathbb{N}} \Sigma^n$, where the sets Σ^n are recursively defined by

$$\Sigma^0 = \{\varepsilon\} \quad \text{and} \quad \Sigma^{n+1} = \{wa \mid w \in \Sigma^n,\ a \in \Sigma\} \text{ for all } n \geqslant 0$$

That is, given $n \in \mathbb{N}$, the elements of Σ^n are precisely the words over Σ of length n.

Example 6.2.27

The elements of $\{0,1\}^*$ are precisely the finite binary strings:

$$\{0,1\}^* = \{\varepsilon,\ 0,\ 1,\ 00,\ 01,\ 10,\ 11,\ 000,\ 001,\ 010,\ 011,\ \dots\}$$

Likewise, the elements of $\{0,1,2,3,4,5,6,7,8,9\}^*$ are the finite strings of the base-10 digits. Keep in mind, though, that formally the elements of this string are *words* and not *numbers*, and that two words may correspond with the same natural number—for example 123 and 00123 are distinct elements of $\{0,1,2,3,4,5,6,7,8,9\}^*$, despite the fact that they both *represent* the same natural number.

Exercise 6.2.28

Let $\Sigma = \{0,1,2,3,4,5,6,7,8,9,\div,-\}$. Describe which words over Σ immediately represent rational numbers (in the sense of treating the digit symbols as numbers and the symbols $\div$ and $-$ as arithmetic operations). Find some examples of elements of Σ^* that do not represent rational numbers.

Exercise 6.2.29

Let Σ be an alphabet. Prove that if Σ is countable, then Σ^* is countable.

As hinted by Example 6.2.27, we can use words over sets to make precise what we mean by 'finite description'.

The idea is that finitely describing the elements of a set X amounts to defining a function $D : X \to \Sigma^*$ for some set Σ; given $x \in X$, the word $D(x) \in \Sigma^*$ can be thought of as a finite description of the element $x \in X$.

But D cannot be any old function—if two elements of X have the same description, then they should be equal. That is

$$\forall x, y \in X,\ D(x) = D(y) \Rightarrow x = y$$

But this says exactly that D must be injective!

Definition 6.2.30

Let X be a set and let Σ be an alphabet. A **finite description** of the elements of X over Σ is an injection $D : X \to \Sigma^*$.

Example 6.2.31

Let $\mathscr{F}(\mathbb{N})$ be the set of all finite subsets of $\mathbb{N}$. We already know that this set is countable by Example 6.2.16; now we exhibit a finite description of its elements.

Let $\Sigma = \left\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \boxed{\{}, \boxed{\}}, \boxed{,}\right\}$—that is, the elements of the alphabet Σ are the digits 0 through 9, an open curly bracket, a closed curly bracket and a comma. (The curly brackets and comma are boxed to distinguish them from the same symbols used in the list notation describing Σ.)

Define $D : \mathscr{F}(\mathbb{N}) \to \Sigma^*$ by letting $D(U)$ be the expression of the finite set $U \subseteq \mathbb{N}$ in list notation, with its elements listed in increasing order. This ensures well-definedness of D—for example

$$\{1,2,3\} = \{2,3,1\} \quad \text{but} \quad D(\{1,2,3\}) = D(\{2,3,1\}) = \{1,2,3\}$$

But then D is injective: given finite subsets $U, V \subseteq \mathbb{N}$, if $D(U) = D(V)$, then U and V have identical expressions in list notation—this means that they have the same elements, so they are equal!

Example 6.2.32

Every set has a trivial finite description, taking the set itself to be the alphabet. Indeed, the function $D : X \to X^*$ defined by $D(x) = x$ for all $x \in X$ is evidently an injection!

Example 6.2.32 raises the point that if we want to prove that a set is countable by showing that its elements have finite descriptions, then we must place some restrictions on the alphabets that we may use. For example, the alphabet that we used in Example 6.2.31 is finite—it has 13 elements.

Exercise 6.2.33

Give an explicit finite description of the rational numbers over a finite alphabet.

Exercise 6.2.34

Let $\mathscr{C}(\mathbb{N})$ be the set of all *cofinite* subsets of $\mathbb{N}$; that is

$$\mathscr{C}(\mathbb{N}) = \{U \subseteq \mathbb{N} \mid \mathbb{N} \setminus U \text{ is finite}\}$$

Give an explicit finite description of the elements of $\mathscr{C}(\mathbb{N})$ over a finite alphabet.

Theorem 6.2.35
Let X be a set. Then X is countable if and only if there is a finite description of the elements of X over a countable alphabet.

Proof
If X is countable, then we may take X itself to be our alphabet! The function $D : X \to X^*$ defined by $D(x) = x$ for all $x \in X$ is evidently injective.

Conversely, if there is an injection $D : X \to \Sigma^*$ for some countable alphabet Σ, then Σ^* is countable by Exercise 6.2.29, and so X is countable by Exercise 6.2.12(ii). □

Example 6.2.36
In light of Theorem 6.2.35, we may deduce from Example 6.2.31 and Exercises 6.2.33 and 6.2.34 that $\mathscr{F}(\mathbb{N})$, $\mathbb{Q}$ and $\mathscr{C}(\mathbb{N})$ are all countable sets.

Exercise 6.2.37
Prove that the set of all polynomials with rational coefficients is countable.

Exercise 6.2.38
Consider the following erroneous argument that claims to be a proof that $\mathscr{P}(\mathbb{N})$ is countable.

> Let $\Sigma = \left\{\mathbb{N}, x, U, V, \boxed{|}, \boxed{\in}, \boxed{\{}, \boxed{\}}\right\}$—again, we have boxed some symbols to emphasise that they are elements of Σ.
>
> Define $D : \mathscr{P}(\mathbb{N}) \to \Sigma^*$ by
>
> $$D(U) = \{x \in \mathbb{N} \mid x \in U\} \quad \text{and} \quad D(V) = \{x \in \mathbb{N} \mid x \in V\}$$
>
> for all $U, V \subseteq \mathbb{N}$. Then Σ is finite, and D is injective since for all $U, V \subseteq \mathbb{N}$ we have
>
> $$D(U) = D(V) \;\Rightarrow\; \{x \in \mathbb{N} \mid x \in U\} = \{x \in \mathbb{N} \mid x \in V\} \;\Rightarrow\; U = V$$
>
> Since D is a finite description of the subsets of $\mathbb{N}$ over a finite alphabet, it follows from Theorem 6.2.35 that $\mathscr{P}(\mathbb{N})$ is countable. □

This proof cannot be valid, since we already proved in Theorem 6.2.17 that $\mathscr{P}(\mathbb{N})$ is uncountable. Where in the proof did we go wrong?

Exercise 6.2.39
Suppose we were to attempt to prove that $\mathscr{P}(\mathbb{N})$ is countable using the same argument as Example 6.2.31. Where does the proof fail?

Exercise 6.2.40
Prove that, for every countable set X, there is a finite description of the elements of X over a *finite* alphabet.

Section 6.E
Chapter 6 exercises

Part II

Topics in pure mathematics

Chapter 7

Number theory

Section 7.1

Division

This section introduces the notion of *divisibility*. As we have already mentioned, it is not always the case that one integer can divide another. As you read through this section, note that we never use fractions; everything we do is *internal* to $\mathbb{Z}$, and does not require that we 'spill over' to $\mathbb{Q}$ at any point. This will help you when you study ring theory in the future, and is a good practice to mimic in your own work.

The following theorem, called the division theorem, is the crux of everything that is to follow.

Theorem 7.1.1 · *Division theorem*
Let $a, b \in \mathbb{Z}$ with $b \neq 0$. There exists a unique pair $(q, r) \in \mathbb{Z}^2$ such that

$$a = qb + r \quad \text{and} \quad 0 \leqslant r < |b|$$

Strategy 7.1.2
Let's first consider the case when $a \geqslant 0$ and $b > 0$ (so that $|b| = b$). Dropping the requirement that $0 \leqslant r < b$, it is easy to find $(q, r) \in \mathbb{Z}^2$ such that $a = qb + r$—indeed, $a = 0 \cdot b + a$, so we could take $(q, r) = (0, a)$. However, this isn't good enough, since if $a \geqslant b$ then setting $r = a$ would not satisfy the requirement that $0 \leqslant r < b$.

However, if $r \geqslant b$, then (intuitively) we can make r smaller by making q larger—to see this, note that

$$qb + r = (q+1)b + (r-b)$$

so whenever we have a pair $(q, r) \in \mathbb{Z}^2$ such that $a = qb + r$, we can replace (q, r) by $(q+1, r-b)$ to obtain another such pair with a smaller value of r.

We want to make r as small as possible without it becoming negative—in other words, we want r to be the least *natural number* such that $a = qb + r$ for some $q \in \mathbb{Z}$. This suggests that we should prove the division theorem using the well-ordering principle (Theorem 4.3.11).

Proof
We may assume that $b > 0$: if not, replace b by $-b$ and q by $-q$ in the proof below.

- **Existence.** Define $X = \{r \in \mathbb{N} \mid a = qb + r \text{ for some } q \in \mathbb{Z}\} \subseteq \mathbb{N}$.

  Note first that X is inhabited. Indeed:

  - Suppose $a \geqslant 0$. Then $a \in X$, since $a = 0 \cdot b + a$.
  - Suppose $a < 0$. Since $b > 0$ we have $1 - b \leqslant 0$, so that $(1-b)a \geqslant 0$. But then $a = ab + (1-b)a$, and so $(1-b)a \in X$.

  By the well-ordering principle, X has a least element. Let r be the least element of X, and let $q \in \mathbb{Z}$ be such that $a = qb + r$—note that q exists by definition of X.

  We have $r \geqslant 0$ since $r \in \mathbb{N}$, so it remains to prove that $r < b$. Well, if $r \geqslant b$, then $r - b \in \mathbb{N}$ and $a = qb + r = (q+1)b + (r-b)$, which implies that $r - b \in X$; but $r - b < r$ since $b > 0$, so this contradicts minimality of r. Therefore $r < b$, as required.

- **Uniqueness.** Let $(q_0, r_0), (q_1, r_1) \in \mathbb{Z}^2$ and assume that $a = q_0 b + r_0 = q_1 b + r_1$ and that $0 \leqslant r_0, r_1 < b$.

  We may assume that $r_0 \leqslant r_1$—otherwise, swap the roles of (q_0, r_0) and (q_1, r_1) in what follows.

  Note first that

  $$(q_1 - q_0)b + (r_1 - r_0) = (q_1 b + r_1) - (q_0 b + r_0) = a - a = 0$$

  Since $r_0 \leqslant r_1$ we have $r_1 - r_0 \geqslant 0$, and since $r_1 < b$ and $r_0 \geqslant 0$ we have $r_1 - r_0 < b$.

  If $q_1 - q_0 < 0$, then $q_1 - q_0 \leqslant -1$ since $q_1 - q_0 \in \mathbb{Z}$, and so

  $$0 = (q_1 - q_0)b + (r_1 - r_0) \leqslant -b + (r_1 - r_0) < b - b = 0$$

  If $q_1 - q_0 > 0$, then $q_1 - q_0 \geqslant 1$ since $q_1 - q_0 \in \mathbb{Z}$, and so

  $$0 = (q_1 - q_0)b + (r_1 - r_0) \geqslant b + (r_1 - r_0) \geqslant b > 0$$

  Both of these are contradictions, so it must be the case that $q_1 - q_0 = 0$, and so $q_0 = q_1$.

  Therefore we have $0 = (q_1 - q_0)b + (r_1 - r_0) = r_1 - r_0$, and so $r_0 = r_1$.

  So $(q_0, r_0) = (q_1, r_1)$, as required. □

Definition 7.1.3

Let $a, b \in \mathbb{Z}$ with $b \neq 0$, and let q, r be the unique integers such that

$$a = qb + r \quad \text{and} \quad 0 \leqslant r < |b|$$

We say q is the **quotient** and r is the **remainder** of a divided by b.

Example 7.1.4

Some examples of division include:

$$14 = 2 \times 5 + 4, \qquad -14 = -3 \times 5 + 1, \qquad 15 = 3 \times 5 + 0$$

Note in particular that, even though the remainder of 14 when divided by 5 is 4, the remainder of -14 when divided by 5 is 1, not -4. This is because remainders are non-negative by definition.

Definition 7.1.5

Let $a, b \in \mathbb{Z}$. We say b **divides** a, or that b is a **divisor** (or **factor**) of a, if there exists $q \in \mathbb{Z}$ such that $a = qb$. To denote the fact that b divides a we write $b \mid a$ (LaTeX code: `\mid`). For the negation $\neg(b \mid a)$ write $b \nmid a$ (LaTeX code: `\nmid`).

Thus, when $b \neq 0$, saying $b \mid a$ is equivalent to saying that the remainder of a divided by b is 0.

Example 7.1.6

5 divides 15 since $15 = 3 \times 5$. However, 5 does not divide 14: we know that the remainder of 14 divided by 5 is 4, not 0—and it can't be both since we proved in the division theorem that remainders are unique!

Exercise 7.1.7
Show that if $a \in \mathbb{Z}$ then $1 \mid a$, $-1 \mid a$ and $a \mid 0$. For which integers a does $a \mid 1$? For which integers a does $0 \mid a$?

We now introduce the very basic notion of a *unit*. This notion is introduced to rule out trivialities. Units become interesting when talking about general rings, but in $\mathbb{Z}$, the units are very familiar.

Definition 7.1.8
Let $u \in \mathbb{Z}$. We say u is a **unit** if $u \mid 1$; that is, u is a unit if there exists $v \in \mathbb{Z}$ such that $uv = 1$.

Proposition 7.1.9
The only units in $\mathbb{Z}$ are 1 and -1.

Proof
First note that 1 and -1 are units, since $1 \cdot 1 = 1$ and $(-1) \cdot (-1) = 1$. Now suppose that $u \in \mathbb{Z}$ is a unit, and let $v \in \mathbb{Z}$ be such that $uv = 1$. Certainly $u \neq 0$, since $0v = 0 \neq 1$. If $u > 1$ or $u < -1$ then $v = \frac{1}{u} \notin \mathbb{Z}$. So we must have $u \in \{-1, 1\}$. □

Exercise 7.1.7 shows that -1, 0 and 1 are fairly trivial from the point of view of divisibility, so we often want to remove them from consideration when defining concepts or proving facts about integers. For this reason, we have a term for integers that are *not* trivial in this sense.

Definition 7.1.10
A **nonzero nonunit** is an integer that is neither zero nor a unit.

By Proposition 7.1.9, the nonzero nonunits are precisely the integers other than -1, 0 and 1. In other words, an integer n is a nonzero nonunit if and only if $|n| > 1$.

Greatest common divisors

Definition 7.1.11
Let $a, b \in \mathbb{Z}$.

- A **common divisor** of a and b is an integer c such that $c \mid a$ and $c \mid b$;
- A **greatest common divisor** of a and b is an integer d such that (i) d is a common divisor of a and b, and (ii) for all $c \in \mathbb{Z}$, if c is a common divisor of a and b, then $c \mid d$.

Condition (ii) in the definition of 'greatest common divisor' in Definition 7.1.11 is a little counter-intuitive: the word 'greatest' suggests that we might want $d \geqslant c$ for all common divisors c of a and b, but the definition has '$c \mid d$' instead of '$d \geqslant c$'. That is, instead of d being *greater than or equal to* all common divisors, it is *divisible by* all common divisors.

There are two good reasons for defining it this way. First, we will use greatest common divisors primarily for reasoning about divisibility—the order of the integers as they appear on the number line turns out to be much less relevant in proving facts in number theory. Second, it allows us to generalise the definition of 'greatest common divisor' to other settings where we *can* make sense of divisibility but don't have a natural notion of order, such as polynomials.

Example 7.1.12

2 is a greatest common divisor of 4 and 6; indeed:

(i) $4 = 2 \times 2$, and $6 = 3 \times 2$, so $2 \mid 4$ and $2 \mid 6$;

(ii) Suppose $q \mid 4$ and $q \mid 6$. The divisors of 4 are $\pm 1, \pm 2, \pm 4$ and the divisors of 6 are ± 1, ± 2, ± 3, ± 6. Since q divides both, it must be the case that $q \in \{-2, -1, 1, 2\}$; in any case, $q \mid 2$.

Likewise, -2 is a greatest common divisor of 4 and 6.

Exercise 7.1.13

There are two greatest common divisors of 6 and 15; find both.

We will now prove that greatest common divisors *exist*—that is, any two integers have a greatest common divisor—and that they are *unique up to sign*.

Theorem 7.1.14

Every pair of integers a, b has a greatest common divisor.

Proof

First note that the greatest common divisor of 0 and 0 is 0—the conditions in Definition 7.1.11 follow immediately from the fact that every integer divides 0 (as you hopefully found in Exercise 7.1.7).

So now let's assume that a and b are not both zero.

Define a subset $X \subseteq \mathbb{N}$ by

$$X = \{n \in \mathbb{N} \mid n > 0 \text{ and } n = au + bv \text{ for some } u, v \in \mathbb{Z}\}$$

X is inhabited. To see this, note that $a^2 > 0$ or $b^2 > 0$ since a and b are not both zero, so letting $u = a$ and $v = b$ in the expression $au + bv$, we see that $au + bv = a^2 + b^2 > 0$, and so $a^2 + b^2 \in X$.

By the well-ordering principle, X has a least element d, and by definition of X there exist $u, v \in \mathbb{Z}$ such that $d = au + bv$.

We will prove that d is a greatest common divisor for a and b.

- $d \mid a$. To see this, let r be the remainder of a when divided by d, so that $0 \leqslant r < d$ and $a = qd + r$ for some $q \in \mathbb{Z}$. But then

$$r = a - qd = a - q(au + bv) = a(1 - qu) + b(-qv)$$

If $r > 0$ then this implies that $r \in X$; but this would contradict minimality of $d \in X$, since $r < d$. So we must have $r = 0$. Therefore $a = qd$, and so $d \mid a$, as required.

- $d \mid b$. The proof of this is identical to the proof that $d \mid a$, but with the roles of a and b swapped.
- Suppose c is a common divisor of a and b. Then $c \mid au + bv$ by Exercise 0.40, and so $c \mid d$ since $au + bv = d$.

So we have verified that d is a greatest common divisor of a and b. □

Exercise 7.1.15

Let $a, b \in \mathbb{Z}$. If d_1 and d_2 are two greatest common divisors of a and b, then either $d_1 = d_2$ or $d_1 = -d_2$.

Aside

A consequence of Theorem 7.1.14 and Exercise 7.1.15 is that every pair of integers has a unique non-negative greatest common divisor! Written symbolically, we can say

$$\forall (a,b) \in \mathbb{Z} \times \mathbb{Z}, \exists! d \in \mathbb{Z}, \left(\begin{array}{c} d \geqslant 0 \text{ and } d \text{ is a greatest} \\ \text{common divisor for } a \text{ and } b \end{array} \right)$$

As discussed in Section 3.1, since this is a formula of the form 'for all … there exists a unique …', this defines a function $\gcd : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$. We won't explicitly refer to the fact that gcd is a function; rather, we'll just concern ourselves with its values, as in Notation 7.1.16.

Exercise 7.1.15 justifies our use of the following notation to refer to greatest common divisors.

Notation 7.1.16

Let $a, b \in \mathbb{Z}$. Denote by $\gcd(a,b)$ (LaTeX code: `\mathrm{gcd}`) the (unique!) non-negative greatest common divisor of a and b.

Example 7.1.17

In Example 7.1.12, we saw that both 2 and -2 are greatest common divisors of 4 and 6. Using Notation 7.1.16, we can now write $\gcd(4,6) = 2$.

Exercise 7.1.18

For each $n \in \mathbb{Z}$, let $D_n \subseteq \mathbb{Z}$ be the set of divisors of n. Prove that $D_a \cap D_b = D_{\gcd(a,b)}$ for all $a, b \in \mathbb{Z}$.

Our goal for the rest of this subsection is to investigate the behaviour of greatest common divisors, find out how to compute them, and look into the implications they have for solutions to certain kinds of equations.

Theorem 7.1.19

Let $a, b, m, n \in \mathbb{Z}$, and suppose that $a = mb + n$. Then

$$\gcd(a, b) = \gcd(b, n)$$

Proof

Let $d = \gcd(a, b)$. We check that d satisfies the conditions required to be a greatest common divisor of b and r.

Note that $d \mid a$ and $d \mid b$, so let $s, t \in \mathbb{Z}$ be such that $a = sd$ and $b = td$.

- We already know that $d \mid b$; and $d \mid n$ since

$$n = a - mb = sd - mtd = (s - mt)d$$

so d is a common divisor of b and n.

- Let $c \in \mathbb{Z}$ and suppose that $c \mid b$ and $c \mid n$; say $b = uc$ and $n = vc$ with $u, v \in \mathbb{Z}$. Then $c \mid a$, since

$$a = mb + n = muc + vc = (mu + v)c$$

so $c \mid d$ since $d = \gcd(a, b)$.

So d is a greatest common divisor of b and n. Since $d \geqslant 0$, we have $d = \gcd(b, n)$ by uniqueness (Exercise 7.1.15), as required. □

A consequence of Theorem 7.1.19 is that, for all $a, b \in \mathbb{Z}$ with $b \neq 0$, if r is the remainder of a when divided by b, then $\gcd(a, b) = \gcd(b, r)$. To see this, note that $a = qb + r$ for some $q \in \mathbb{Z}$, and so we can set $m = q$ and $n = r$ in the result of Theorem 7.1.19.

Exercise 7.1.20

Prove the following facts about greatest common divisors:

(a) $\gcd(a, b) = \gcd(|a|, |b|)$ for all $a, b \in \mathbb{Z}$.

(b) $\gcd(a, b) = \gcd(b, a)$ for all $a, b \in \mathbb{Z}$.

(c) $\gcd(a, 0) = |a|$ for all $a \in \mathbb{Z}$.

Combined with the division theorem (Theorem 7.1.1), the results of Theorem 7.1.19 and Exercise 7.1.20 give rise to a relatively fast algorithm for computing the greatest common divisor of two integers, known as the **Euclidean algorithm**.

Strategy 7.1.21 · *Euclidean algorithm*

Let $a, b \in \mathbb{Z}$. To compute $\gcd(a, b)$, proceed as follows.

(1) Replace a by $|a|$ and b by $|b|$, and then if $a < b$, swap a and b. This ensures that $a \geqslant b \geqslant 0$, and it doesn't affect the value of the greatest common divisor by parts (a) and (b) of Exercise 7.1.20.

(2) If $b = 0$, return a. This is valid since we know $\gcd(a,0) = a$ by Exercise 7.1.20(c).

(3) If $b > 0$, let r be the remainder of a when divided by b, and return $\gcd(b,r)$. This doesn't change the value of the greatest common divisor by Theorem 7.1.19.

Before we work through any examples of the Euclidean algorithm, note that it always terminates: in (3) note that $0 \leqslant r < b$ since r is the remainder of a when divided by b, so the second argument to the algorithm decreases—but remains non-negative—each time it is called. Therefore the second argument must eventually reach 0, at which point the algorithm terminates in (2).

Example 7.1.22

We will compute $\gcd(148,28)$ using the Euclidean algorithm. Note that $148 \geqslant 28 \geqslant 0$, so we don't need to do anything in step (1). Iterating step (3), we obtain:

- $148 = 5 \times 28 + 8$, so $\gcd(148,28) = \gcd(28,8)$.
- $28 = 3 \times 8 + 4$, so $\gcd(28,8) = \gcd(8,4)$.
- $8 = 2 \times 4 + 0$, so $\gcd(8,4) = \gcd(4,0)$.

Finally, step (2) gives $\gcd(4,0) = 4$. So $\gcd(148,48) = 4$.

Example 7.1.23

The Euclidean algorithm works surprisingly quickly, even for relatively large numbers. Consider the problem of computing $\gcd(1311,5757)$ for example:

$$\begin{aligned} 5757 &= 4 \times 1311 + 513 & \rightsquigarrow \gcd(5757,1311) &= \gcd(1311,513) \\ 1311 &= 2 \times 513 + 285 & &= \gcd(513,285) \\ 513 &= 1 \times 285 + 228 & &= \gcd(285,228) \\ 285 &= 1 \times 228 + 57 & &= \gcd(228,57) \\ 228 &= 4 \times 57 + 0 & &= \gcd(57,0) = 57 \end{aligned}$$

So $\gcd(5757,1311) = 57$.

Example 7.1.24

Here's an example where one of the numbers is negative: we compute the value of $\gcd(76,-492)$. Using step (1), we instead compute $\gcd(492,76)$.

$$\begin{aligned} 492 &= 6 \times 76 + 36 & \rightsquigarrow \gcd(492,76) &= \gcd(76,36) \\ 76 &= 2 \times 36 + 4 & &= \gcd(36,4) \\ 36 &= 9 \times 4 + 0 & &= \gcd(4,0) = 4 \end{aligned}$$

Hence $\gcd(76,-492) = \gcd(492,76) = 4$.

Example 7.1.25

Use the Euclidean algorithm to compute the greatest common divisors of the following pairs of integers

$$(12,9), \quad (100,35), \quad (7125,1300), \quad (1010,101010), \quad (-4,14)$$

The following theorem will be useful when we study modular arithmetic in Section 7.3; it is called a 'lemma' for historical reasons, and is really an important result in its own right.

Theorem 7.1.26 · *Bézout's lemma*
Let $a, b, c \in \mathbb{Z}$, and let $d = \gcd(a, b)$. The equation

$$ax + by = c$$

has a solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ if and only if $d \mid c$.

Proof
($\Rightarrow$) Write $a = a'd$ and $b = b'd$, for $a', b' \in \mathbb{Z}$. If there exist $x, y \in \mathbb{Z}$ such that $ax + by = c$, then

$$c = ax + by = a'dx + b'dy = (a'x + b'y)d$$

and so $d \mid c$.

($\Leftarrow$) Suppose $d \mid c$, and let $c = kd$ for some $k \in \mathbb{Z}$.

If $c = 0$, then a solution is $x = y = 0$. If $c < 0$, then $ax + by = c$ if and only if $a(-x) + b(-y) = -c$; so we may assume that $c > 0$.

We proved in Theorem 7.1.14 that a greatest common divisor of a and b is a least element of the set

$$X = \{au + bv \mid u, v \in \mathbb{Z},\ au + bv > 0\}$$

So let $u, v \in \mathbb{Z}$ be such that $au + bv = d$. Then

$$a(ku) + b(kv) = k(au + bv) = kd = c$$

and so letting $x = ku$ and $y = kv$, we see that the equation $ax + by = c$ has a solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.□

Bézout's lemma completely characterises when the equation $ax + by = c$ has a solution. An easy generalisation of Bézout's lemma provides a complete characterisation of when solutions to **linear Diophantine equations** exist, that is equations of the form

$$ax + by = c$$

where $a, b, c \in \mathbb{Z}$. We will soon develop an algorithm for computing *all* solutions to these equations.

Example 7.1.27
Here are some examples of applications of Bézout's lemma.

- Consider the equation $1311x + 5757y = 12963$. We computed in Example 7.1.23 that $\gcd(1311, 5757) = 57$. But $57 \nmid 12963$ since $12963 = 227 \times 57 + 24$. By Bézout's lemma, the equation $1311x + 5757y = 12963$ has no integer solutions.
- For fixed z, the equation $4u + 6v = z$ has solutions exactly when z is even, since $\gcd(4, 6) = 2$.

- For fixed a, b, the equation $au + bv = 0$ always has solution. Indeed, setting $u = b$ and $v = -a$ gives a solution; but we knew one had to exist since by Exercise 7.1.7 we know that $d \mid 0$ for all $d \in \mathbb{Z}$.

Exercise 7.1.28
Which of the following equations have solutions?

(a) $12u + 9v = -18$

(b) $12u + 9v = 1$

(c) $100u + 35v = 125$

(d) $7125u + 1300v = 0$

(e) $1010u + 101010v = 1010101010101010$

(f) $14u - 4v = 12$

Coprimality

Definition 7.1.29
Let $a, b \in \mathbb{Z}$. We say a and b are **coprime** (or **relatively prime**), and write $a \perp b$ (LaTeX code: `\perp`) (read 'a is coprime to b'), if $\gcd(a, b) = 1$.

Example 7.1.30
$4 \perp 9$. To see this, note that if $d \mid 4$ then $d \in \{-4, -2, -1, 1, 2, 4\}$, and if $d \mid 9$ then $d \in \{-9, -3, -1, 1, 3, 9\}$. Hence if $d \mid 4$ and $d \mid 9$, then $d = 1$ or $d = -1$. It follows that $\gcd(4, 9) = 1$.

Exercise 7.1.31
Which integers in the set $[15]$ are coprime to 15?

Proposition 7.1.32
Let $a, b \in \mathbb{Z}$. The following are equivalent:

(1) a and b are coprime;

(2) If $d \in \mathbb{Z}$ with $d \mid a$ and $d \mid b$, then d is a unit.

Proof
We prove that condition (1) implies condition (2), and vice versa.

- (1)⇒(2). Suppose a and b are coprime, and fix $d \in \mathbb{Z}$ with $d \mid a$ and $d \mid b$. Then $d \mid \gcd(a, b) = 1$, so d is a unit.

- (2)⇒(1). Suppose condition (2) above holds. We prove that 1 satisfies the conditions required to be a greatest common divisor of a and b. The fact that $1 \mid a$ and $1 \mid b$ is automatic; and the fact that if $d \mid a$ and $d \mid b$ implies $d \mid 1$ is precisely the condition (2) that we are assuming.

Hence the two conditions are equivalent. □

Exercise 7.1.33
Let a and b be integers, not both zero, and let $d = \gcd(a,b)$. The integers $\frac{a}{d}$ and $\frac{b}{d}$ are coprime.

The following corollary is a specialisation of Bézout's lemma to the case when a and b are coprime.

Corollary 7.1.34
Let $a, b \in \mathbb{Z}$. The equation $au + bv = 1$ has a solution if and only if a and b are coprime. Moreover, if a and b are coprime, then the equation $au + bv = z$ has a solution for all $z \in \mathbb{Z}$.

Proof
By Bézout's lemma (Theorem 7.1.26), the equation $au + bv = 1$ has a solution if and only if $\gcd(a,b) \mid 1$. But the only positive divisor of 1 is 1, so a solution exists if and only if $\gcd(a,b) = 1$, which is precisely the assertion that a and b are coprime.

If a and b are coprime, then $1 = \gcd(a,b) \mid z$ for all $z \in \mathbb{Z}$. So by Bézout's lemma again, the equation $au + bv = z$ has a solution for all $z \in \mathbb{Z}$. □

A useful consequence of Bézout's lemma is the following result:

Proposition 7.1.35
Let $a, b, c \in \mathbb{Z}$. If a and b are coprime and $a \mid bc$, then $a \mid c$.

Proof
By Bézout's lemma (Theorem 7.1.26) there exist integers u and v such that $au + bv = 1$. Multiplying by c gives $acu + bcv = c$. Since $a \mid bc$, we can write $bc = ka$ for some $k \in \mathbb{Z}$, and so $acu + kav = c$. But then

$$(cu + kv)a = c$$

which proves that $a \mid c$. □

Linear Diophantine equations

We have now seen two important results:

- The **Euclidean algorithm**, which was a procedure for computing the greatest common divisor of two integers.
- **Bézout's lemma**, which provides a necessary and sufficient condition for equations of the form $ax + by = c$ to have an integer solution.

We will now develop the **extended Euclidean algorithm**, which provides a method for computing a solutions to (bivariate) linear Diophantine equations, when such a solution exists. Then we will prove a theorem that characterises *all* integer solutions in terms of a given solution.

Example 7.1.36

Suppose we want to find integers x and y such that $327x + 114y = 18$. Running the Euclidean algorithm yields that $\gcd(327, 114) = 3$ — see below. For reasons soon to become apparent, we rearrange each equation to express the remainder on its own.

$$\begin{aligned}
327 &= 2\times 114 + 99 &\Rightarrow\quad 99 &= 327 - 2\times 114 &(1)\\
114 &= 1\times 99 + 15 &\Rightarrow\quad 15 &= 114 - 1\times 99 &(2)\\
99 &= 6\times 15 + 9 &\Rightarrow\quad 9 &= 99 - 6\times 15 &(3)\\
15 &= 1\times 9 + 6 &\Rightarrow\quad 6 &= 15 - 1\times 9 &(4)\\
9 &= 1\times 6 + 3 &\Rightarrow\quad 3 &= 9 - 1\times 6 &(5)\\
6 &= 2\times 3 + 0
\end{aligned}$$

We can then express 3 in the form $327u + 114v$ by successively substituting the equations into each other:

- Equation (5) expresses 3 as a linear combination of 6 and 9. Substituting equation (4) yields:

$$3 = 9 - 1\times(15 - 1\times 9) \quad\Rightarrow\quad 3 = 2\times 9 - 1\times 15$$

- This now expresses 3 as a linear combination of 9 and 15. Substituting equation (3) yields:

$$3 = 2\times(99 - 6\times 15) - 1\times 15 \quad\Rightarrow\quad 3 = (-13)\times 15 + 2\times 99$$

- This now expresses 3 as a linear combination of 15 and 99. Substituting equation (2) yields:

$$3 = (-13)\times(114 - 1\times 99) + 2\times 99 \quad\Rightarrow\quad 3 = 15\times 99 - 13\times 114$$

- This now expresses 3 as a linear combination of 99 and 114. Substituting equation (1) yields:

$$3 = 15\times(327 - 2\times 114) - 13\times 114 \quad\Rightarrow\quad 3 = (-43)\times 114 + 15\times 327$$

Now that we've expressed 3 as a linear combination of 114 and 327, we're nearly done: we know that $18 = 6\times 3$, so multiplying through by 6 gives

$$18 = (-258)\times 114 + 90\times 327$$

Hence $(x, y) = (90, -258)$ is a solution to the equation $327x + 114y = 18$.

Proof tip

Let $a, b \in \mathbb{Z}$ and let $d = \gcd(a, b)$. To find integers x, y such that $ax + by = d$:

(i) Run the Euclidean algorithm on the pair (a, b), keeping track of all quotients and remainders; note that in the penultimate equation, the remainder is $\gcd(a, b)$.

(ii) Rearrange each equation to isolate the remainder.

(iii) Starting from the (rearranged) penultimate equation, substitute for the remainders in reverse order until $\gcd(a,b)$ is expressed in the form $ax+by$ for some $x,y\in\mathbb{Z}$.

This process is called the **extended Euclidean algorithm**.

Exercise 7.1.37
Find a solution $(x,y)\in\mathbb{Z}\times\mathbb{Z}$ to the equation $630x+385y=4340$.

Now that we have a procedure for computing *one* solution to the equation $ax+by=c$, we need to come up with a procedure for computing *all* solutions. This can be done by proving the following theorem.

Theorem 7.1.38
Let $a,b,c\in\mathbb{Z}$, where a and b are not both zero. Suppose that x_0 and y_0 are integers such that $ax_0+by_0=c$. Then, $(x,y)\in\mathbb{Z}\times\mathbb{Z}$ is another solution to the equation $ax+by=c$ if and only if

$$x=x_0+k\cdot\frac{b}{\gcd(a,b)} \qquad \text{and} \qquad y=y_0-k\cdot\frac{a}{\gcd(a,b)}$$

for some $k\in\mathbb{Z}$.

Thus, as soon as we've found one solution $(x,y)=(x_0,y_0)$ to the equation $ax+by=c$, this theorem tells us what all other solutions must look like.

Proof of Theorem 7.1.38
We prove the two directions separately.

($\Rightarrow$). First suppose that (x_0,y_0) is an integer solution to the equation $ax+by=c$. Let $k\in\mathbb{Z}$ and let

$$x=x_0+k\cdot\frac{b}{\gcd(a,b)} \qquad \text{and} \qquad y=y_0-k\cdot\frac{a}{\gcd(a,b)}$$

Then

$$\begin{aligned}
&ax+by\\
&=a\left(x_0+k\cdot\frac{b}{\gcd(a,b)}\right)+b\left(y_0-k\cdot\frac{a}{\gcd(a,b)}\right) && \text{by definition of } x \text{ and } y\\
&=(ax_0+by_0)+ak\cdot\frac{b}{\gcd(a,b)}-kb\cdot\frac{a}{\gcd(a,b)} && \text{rearranging}\\
&=(ax_0+by_0)+\frac{kab-kab}{\gcd(a,b)} && \text{combining the fractions}\\
&=ax_0+by_0 && \text{since } kab-kab=0\\
&=c && \text{since } (x_0,y_0) \text{ is a solution}
\end{aligned}$$

so (x,y) is indeed a solution to the equation.

($\Leftarrow$). First suppose that $a\perp b$. Fix a solution (x_0,y_0) to the equation $ax+by=c$, and let (x,y) be another solution. Then

$$a(x-x_0)+b(y-y_0)=(ax_0+by_0)-(ax+by)=c-c=0$$

so that

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

Now a and b are coprime, so by Proposition 7.1.35, we have $a \mid y_0 - y$ and $b \mid x - x_0$. Let $k, \ell \in \mathbb{Z}$ be such that $x - x_0 = kb$ and $y_0 - y = \ell a$. Then substituting into the above equation yields

$$a \cdot kb = b \cdot \ell a$$

and hence $(k-\ell)ab = 0$. Since $ab \neq 0$, we have $k = \ell$, so that

$$x = x_0 + kb \quad \text{and} \quad y = y_0 - ka$$

Now we drop the assumption that $a \perp b$. Let $\gcd(a,b) = d \geqslant 1$. We know that $d \mid c$, by Bézout's lemma (Theorem 7.1.26), and so

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

is another linear Diophantine equations, and moreover $\frac{a}{d} \perp \frac{b}{d}$ by Exercise 7.1.33. By what we proved above, we have

$$x = x_0 + k \cdot \frac{b}{d} \quad \text{and} \quad y = y_0 - k \cdot \frac{a}{d}$$

for some $k \in \mathbb{Z}$. But this is exactly what we sought to prove! □

Example 7.1.39

We know that $(x,y) = (90,-258)$ is a solution to the equation $327x + 114y = 18$, and

$$\frac{327}{\gcd(327,114)} = \frac{327}{3} = 109 \qquad \text{and} \qquad \frac{114}{\gcd(327,114)} = \frac{114}{3} = 38$$

so this theorem tells us that $(x,y) \in \mathbb{Z} \times \mathbb{Z}$ is a solution to the equation $327x + 114y = 18$ if and only if

$$x = 90 + 38k \qquad \text{and} \qquad y = -258 - 109k$$

for some $k \in \mathbb{Z}$.

Exercise 7.1.40

Find all integers x, y such that

$$630x + 385y = 4340$$

Least common multiples

You would be forgiven for wondering why so much of the foregoing section was devoted to greatest common divisors, with no mention of least common multiples. We will now give the latter some attention.

Definition 7.1.41

Let $a, b \in \mathbb{Z}$. An integer m is a **least common multiple** of a and b if:

(a) $a \mid m$ and $b \mid m$;

(b) If c is another integer such that $a \mid c$ and $b \mid c$, then $m \mid c$.

The definition of least common multiple is *dual* to that of greatest common divisor (Definition 7.1.11). This means that many properties of greatest common divisors have corresponding 'dual' properties, which hold of least common multiples. As such, we will not say much here about least common multiples, and that which we *do* say is in the form of exercises.

Exercise 7.1.42

Let $a, b \in \mathbb{Z}$. Prove that a and b have a least common multiple. Furthermore, prove that least common multiples are unique up to sign, in the sense that if m, m' are two least common multiples of a and b, then $m = m'$ or $m = -m'$.

As with greatest common divisors, Exercise 7.1.42 justifies the following definition.

Definition 7.1.43

Given $a, b \in \mathbb{Z}$, denote by $\operatorname{lcm}(a,b)$ (LaTeX code: `\mathrm{lcm}`) the non-negative least common multiple of a and b.

Exercise 7.1.44

Let $a, b \in \mathbb{Z}$. Prove that $\gcd(a,b) \cdot \operatorname{lcm}(a,b) = |ab|$.

Section 7.2

Prime numbers

Thinking of divisibility as a way of *breaking down* an integer, for example $12 = 2 \times 2 \times 3$, our goal now is to show that there are particular integers that are *atomic*—they are the building blocks of the integers, in the sense that:

- Every integer can be broken into a product of these atomic integers...
- ...and these atomic integers cannot themselves be broken down any further...
- ...*and* there is an essentially unique way to write an integer as a product of these atomic integers.

There are a couple of fairly vague terms used here: 'atomic' and 'essentially unique'. But as always, we will make these terms precise when we need to.

Primes and irreducibles

There are two ways that we might want to characterise the so-called *atomic* integer that we just mentioned.

- One way that an integer might be atomic is if it allows us to break down products of integers—this leads to the notion of *prime* (Definition 7.2.1).
- Another way is that an integer might be atomic is if it cannot be split up as a product of more than one integer (in a nontrivial way)—this leads to the notion of *irreducible* (Definition 7.2.6).

Conveniently, as we will show in Theorem 7.2.11, these two notions coincide. But the fact that they coincide is not obvious, and uses essential properties of the integers that do not hold in more general structures.

The definition of *prime* that we are about to give comes from abstract algebra (specifically, from ring theory). It might seem strange, but we will soon be able to show that the more familiar definition—that is, having exactly two positive divisors—is equivalent to this one.

Definition 7.2.1

An integer p is (**ring theoretically**) **prime** if p is a nonzero nonunit and, for all $a, b \in \mathbb{Z}$, if $p \mid ab$ then $p \mid a$ or $p \mid b$.

Example 7.2.2

2 is prime. To see this, suppose it isn't. Then there exist $a, b \in \mathbb{Z}$ such that $2 \mid ab$ but 2 divides neither a nor b. Thus a and b are both odd, meaning that ab is odd... but this contradicts the assumption that $2 \mid ab$.

However, 18 is not prime. Indeed, $18 \mid 12 \times 15$ but 18 divides neither 12 nor 15.

Exercise 7.2.3
Using Definition 7.2.1, prove that 3 and 5 are prime, and that 4 is not prime.

Example 7.2.4
Let $k \in \mathbb{Z}$ with $0 < k < 5$. We'll show that $5 \mid \binom{5}{k}$.

Well, by Theorem 4.2.14 we know that

$$5! = \binom{5}{k} k!(5-k)!$$

By Definition 4.1.14, we have

$$\underbrace{5 \times 4!}_{=5!} = \binom{5}{k} \times \underbrace{1 \times \cdots \times k}_{=k!} \times \underbrace{1 \times \cdots \times (5-k)}_{=(5-k)!}$$

Since 5 is prime, it must divide one of the factors on the right-hand side of this equation. Thus, either 5 divides $\binom{5}{k}$, or it divides ℓ for some $1 \leqslant \ell \leqslant k$ or $1 \leqslant \ell \leqslant 5-k$. But $k < 5$ and $5-k < 5$, so it cannot divide any of these values of ℓ—if it did, it would imply $5 \leqslant \ell \leqslant k$ or $5 \leqslant \ell \leqslant 5-k$, which is nonsense. Hence 5 must divide $\binom{5}{k}$.

Exercise 7.2.5
Let $p \in \mathbb{Z}$ be a positive prime and let $0 < k < p$. Show that $p \mid \binom{p}{k}$.

We now arrive at our second notion of *atomic*, capturing the idea that it should not be possible to break an atomic integer into smaller parts.

Definition 7.2.6
An integer a is **irreducible** if a is a nonzero nonunit and, for all $m, n \in \mathbb{Z}$, if $a = mn$, then either m or n is a unit. Otherwise, a is **reducible**.

The notion of *irreducible* captures more closely the more familiar notion of 'prime', as the next result shows.

Proposition 7.2.7
Let $p \in \mathbb{Z}$ be a nonzero nonunit. Then p is irreducible if and only if the only divisors of p are p, $-p$, 1 and -1.

Proof
Suppose p is irreducible and that $a \mid p$. Then $p = ab$ for some $b \in \mathbb{Z}$. Since p is irreducible, either a or b is a unit. If a is a unit then $b = \pm p$, and if b is a unit then $a = \pm p$. So the only divisors of p are ± 1 and $\pm p$.

Conversely, suppose that the only divisors of p are ± 1 and $\pm p$, and let $a, b \in \mathbb{Z}$ with $p = ab$. We want to prove that a or b is a unit. Since $a \mid p$, we have $a \in \{1, -1, p, -p\}$. If $a = \pm 1$, then a is a unit; if $a = \pm p$, then $b = \pm 1$, so that b is a unit. In any case, either a or b is a unit, and hence p is irreducible. □

Example 7.2.8

A couple of examples of reducible and irreducible numbers are:

- 2 is irreducible: if $2 = mn$ then either m or n is even, otherwise we'd be expressing an even number as the product of two odd numbers. We may assume m is even, say $m = 2k$; then $2 = 2kn$, so $kn = 1$ and hence n is a unit.
- 20 is reducible since $20 = 4 \times 5$ and neither 4 nor 5 is a unit.

Exercise 7.2.9

Let $p \in \mathbb{Z}$. Prove that if p is ring theoretically prime, then p is irreducible.

Lemma 7.2.10

Let $a \in \mathbb{Z}$ be a nonzero nonunit. Then there are irreducibles $p_1, \ldots, p_n$ such that $a = p_1 \times \cdots \times p_n$.

Proof

We may assume $a > 0$, since if $a < 0$ we can just multiply by -1.

We proceed by strong induction on $a \geqslant 2$. The base case has $a = 2$ since we consider only nonunits.

- (**Base case**) We have shown that 2 is irreducible, so setting $p_1 = 2$ yields a product of primes.
- (**Induction step**) Let $a \geqslant 2$ and suppose that each integer k with $2 \leqslant k \leqslant a$ has an expression as a product of irreducibles. If $a+1$ is irreducible then we're done; otherwise we can write $a+1 = st$, where $s, t \in \mathbb{Z}$ are nonzero nonunits. We may assume further that s and t are positive. Moreover, $s < a+1$ and $t < a+1$ since $s, t \geqslant 2$.

  By the induction hypothesis, s and t have expressions as products of irreducibles. Write

  $$s = p_1 \times \cdots \times p_m \quad \text{and} \quad t = q_1 \times \cdots \times q_n$$

  This gives rise to an expression of a as a product of irreducibles:

  $$a = st = \underbrace{p_1 \times \cdots \times p_m}_{=s} \times \underbrace{q_1 \times \cdots \times q_n}_{=t}$$

The result follows by induction. □

Theorem 7.2.11

Let $p \in \mathbb{Z}$. Then p is ring theoretically prime if and only if p is irreducible.

Proof

We prove the two directions separately.

- **Prime ⇒ irreducible.** This was Exercise 7.2.9.
- **Irreducible ⇒ prime.** Suppose p is irreducible. Let $a, b \in \mathbb{Z}$ and suppose $p \mid ab$. We need to show that $p \mid a$ or $p \mid b$. It suffices to show that if $p \nmid a$ then $p \mid b$.

So suppose $p \nmid a$. Let $d = \gcd(p,a)$. Since $d \mid p$ and p is irreducible, we must have $d = 1$ or $d = p$ by Proposition 7.2.7. Since $p \nmid a$ and $d \mid a$, we must therefore have $d = 1$.

By Bézout's lemma (Theorem 7.1.26), there exist $u, v \in \mathbb{Z}$ such that $au + pv = 1$. Multiplying by b gives $abu + pbv = b$. Since $p \mid ab$, there exists $k \in \mathbb{Z}$ such that $pk = ab$. Define $q = ku + bv$; then

$$b = abu + pbv = pku + pbv = p(ku + bv) = qp$$

so $p \mid b$, as required.

So we're done. □

Since primes and irreducibles are the same thing in $\mathbb{Z}$, we will refer to them as 'primes', unless we need to emphasise a particular aspect of them.

Prime factorisation

Having described prime numbers in two ways, each of which emphasises their nature of being 'unbreakable' by multiplication, we will extend Lemma 7.2.10 to prove that every integer can be expressed as a product of primes in an essentially unique way.

Theorem 7.2.12 · *Fundamental theorem of arithmetic*

Let $a \in \mathbb{Z}$ be a nonzero nonunit. There exist primes $p_1, \ldots, p_k \in \mathbb{Z}$ such that

$$a = p_1 \times \cdots \times p_k$$

Moreover, this expression is essentially unique: if $a = q_1 \times \cdots \times q_\ell$ is another expression of a as a product of primes, then $k = \ell$ and, re-ordering the q_i if necessary, for each i there is a unit u_i such that $q_i = u_i p_i$.

Proof

We showed that such a factorisation exists in Lemma 7.2.10, with the word 'prime' replaced by the word 'irreducible'. It remains to prove (essential) uniqueness.

Let k be least such that there is an expression of a as a product of k primes, namely $a = p_1 \times \cdots \times p_k$. Let $a = q_1 \times \cdots \times q_\ell$ be any other such expression. We prove by induction on k that $\ell = k$ and, after re-ordering if necessary, for each i there is a unit u_i such that $q_i = u_i p_i$.

- (**Base case**) If $k = 1$ then $a = p_1$ is itself prime. Then we have $p_1 = q_1 \times \cdots \times q_\ell$. Since p_1 is prime, $p_1 \mid q_j$ for some j; by relabelling q_1 and q_j we may assume that $j = 1$, so that $p_1 \mid q_1$. By irreducibility of q_1 we have $q_1 = u_1 p_1$ for some unit u_1.
- (**Induction step**) Let $k \geqslant 1$ and suppose that any integer which can be expressed as a product of k primes is (essentially) uniquely expressible in such a way. Suppose a has an expression as a product of $k+1$ primes, and that $k+1$ is the least such number. Suppose also that

$$a = p_1 \times \cdots \times p_k \times p_{k+1} = q_1 \times \cdots \times q_\ell$$

Note that $\ell \geqslant k+1$. Since p_{k+1} is prime we must have $p_{k+1} \mid q_j$ for some j; by relabelling q_j and q_ℓ if necessary, we may assume that $j = \ell$, so that $p_{k+1} \mid q_\ell$. As before, $q_\ell = u_{k+1}p_{k+1}$ for some unit u_{k+1}. Dividing through by p_{k+1} gives

$$p_1 \times \cdots \times p_k = q_1 \times \cdots \times q_{\ell-1} \times u_{k+1}$$

Replacing $q_{\ell-1}$ by $q_{\ell-1}u_{k+1}$, which is still prime, we can apply the induction hypothesis to obtain $k = \ell - 1$, so $k+1 = \ell$, and, after reordering if necessary $q_i = u_i p_i$ for all $i \leqslant k$. Since this also holds for $i = k+1$, the induction step is complete.

The result follows by induction. □

Example 7.2.13

Here are some examples of numbers written as products of primes:

- $12 = 2 \times 2 \times 3$. We could also write this as $2 \times 3 \times 2$ or $(-2) \times (-3) \times 2$, and so on.
- $53 = 53$ is an expression of 53 as a product of primes.
- $-1000 = 2 \times 5 \times (-2) \times 5 \times 2 \times 5$.
- We can view any unit as a product of *no* primes. (Don't dwell on this point for too long as it will not arise very often!)

Exercise 7.2.14

Express the following numbers as products of primes:

$$16 \qquad -240 \qquad 5050 \qquad 111111 \qquad -123456789$$

To make things slightly more concise, we introduce a standard way of expressing a number as a product of primes:

Definition 7.2.15

The **canonical prime factorisation** of a nonzero integer a is the expression in the form

$$a = up_1^{j_1} \cdots p_r^{j_r}$$

where $r \geqslant 0$ and:

- $u = 1$ if $a > 0$, and $u = -1$ if $a < 0$;
- The numbers p_i are all positive primes;
- $p_1 < p_2 < \cdots < p_r$;
- $j_i \geqslant 1$ for all i.

We call j_i the **multiplicity** of p_i in a, and we call u the **sign** of a.

Typically we omit u if $u = 1$ (unless $a = 1$), and just write a minus sign $(-)$ if $u = -1$.

Example 7.2.16

The canonical prime factorisations of the integers given in Example 7.2.13 are:

- $12 = 2^2 \cdot 3$.
- $53 = 53$.
- $-1000 = -2^3 \cdot 5^3$.
- $1 = 1$.

Exercise 7.2.17

Write out the canonical prime factorisations of the numbers from Exercise 7.2.14, which were:

$$16 \qquad -240 \qquad 5050 \qquad 111111 \qquad -123456789$$

The following exercise provides another tool for computing greatest common divisors of pairs of integers by looking at their prime factorisations.

Exercise 7.2.18

Let $p_1, p_2, \ldots, p_r$ be distinct primes, and let $k_i, \ell_i \in \mathbb{N}$ for all $1 \leqslant i \leqslant r$. Define

$$m = p_1^{k_1} \times p_2^{k_2} \times \cdots \times p_r^{k_r} \quad \text{and} \quad n = p_1^{\ell_1} \times p_2^{\ell_2} \times \cdots \times p_r^{\ell_r}$$

Prove that

$$\gcd(m,n) = p_1^{u_1} \times p_2^{u_2} \times \cdots \times p_r^{u_r}$$

where $u_i = \min\{k_i, \ell_i\}$ for all $1 \leqslant i \leqslant r$.

Example 7.2.19

We use Exercise 7.2.18 to compute the greatest common divisor of 17640 and 6468.

First we compute the prime factorisations of 17640 and 6468:

$$17640 = 2^3 \cdot 3^2 \cdot 5 \cdot 7^2 \quad \text{and} \quad 6468 = 2^2 \cdot 3 \cdot 7^2 \cdot 11$$

It now follows from Exercise 7.2.18 that

$$\begin{aligned} \gcd(17640, 6468) &= 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^2 \cdot 11^0 \\ &= 4 \cdot 3 \cdot 1 \cdot 49 \cdot 1 \\ &= 588 \end{aligned}$$

Exercise 7.2.18 allows us to provide a concise proof of the following result.

Corollary 7.2.20
Let $p \in \mathbb{Z}$ be prime, let $a \in \mathbb{Z}$ be nonzero, and let $k \geqslant 1$. Then $a \perp p^k$ if and only if $p \nmid a$.

Proof
By the fundamental theorem of arithmetic, we can write

$$a = p^j \times p_1^{j_1} \times \cdots \times p_r^{j_r}$$

where $p_1, \ldots, p_r$ are the primes other than p appearing in the prime factorisation of a, and $j, j_i \in \mathbb{N}$ for all $1 \leqslant i \leqslant r$. Note that $p \mid a$ if and only if $j \geqslant 1$.

Furthermore we have

$$p^k = p^k \times p_1^0 \times \cdots \times p_r^0$$

By Exercise 7.2.18 it follows that

$$\gcd(a, p^k) = p^{\min\{j,k\}} \times p_1^0 \times \cdots \times p_r^0 = p^{\min\{j,k\}}$$

Now:

- If $\min\{j,k\} = 0$ then $j = 0$, in which case $p \nmid a$, and $\gcd(a, p^k) = p^0 = 1$;
- If $\min\{j,k\} > 0$ then $j \geqslant 1$, in which case $p \mid a$, and $p \mid \gcd(a, p^k)$, so $\gcd(a, p^k) \neq 1$.

In particular, $p \nmid a$ if and only if $a \perp p^k$. □

Distribution of primes

So far we have seen several examples of prime numbers; to name a few, we've seen 2, 3, 5 and 53. It might seem like the prime numbers go on forever, but proving this is less than obvious.

Exercise 7.2.21
Let P be an inhabited finite set of positive prime numbers and let m be the product of all the elements of P. That is, for some $n \geqslant 1$ let

$$P = \{p_1, \ldots, p_n\} \quad \text{and} \quad m = p_1 \times \cdots \times p_n$$

where each $p_k \in P$ is a positive prime. Using the fundamental theorem of arithmetic, show that $m + 1$ has a positive prime divisor which is not an element of P.

Theorem 7.2.22
There are infinitely many primes.

Proof
We prove that there are infinitely many *positive* prime numbers—the result then follows immediately. Let P be the set of all positive prime numbers. Then P is inhabited, since $2 \in P$, for example.

If P were finite, then by Exercise 7.2.21, there would be a positive prime which is not an element of P—but P contains all positive primes, so that is impossible. Hence there are infinitely many positive primes. □

This is one proof of many and is attributed to Euclid, who lived around 2300 years ago. We might hope that a proof of the infinitude of primes gives some insight into how the primes are *distributed*. That is, we might ask questions like: how frequently do primes occur? How fast does the sequence of primes grow? How likely is there to be a prime number in a given set of integers?

As a starting point, Euclid's proof gives an algorithm for writing an infinite list of primes:

- Let $p_1 = 2$; we know that 2 is prime;
- Given $p_1, \ldots, p_n$, let p_{n+1} be the smallest positive prime factor of $p_1 \times \cdots \times p_n + 1$.

The first few terms produced would be:

- $p_1 = 2$ by definition;
- $2+1 = 3$, which is prime, so $p_2 = 3$;
- $2 \times 3 + 1 = 7$, which is prime, so $p_3 = 7$;
- $2 \times 3 \times 7 + 1 = 43$, which is prime, so $p_4 = 43$;
- $2 \times 3 \times 7 \times 43 + 1 = 1807 = 13 \times 139$, so $p_5 = 13$;
- $2 \times 3 \times 7 \times 43 \times 13 + 1 = 23479 = 53 \times 443$, so $p_6 = 53$;
- ... and so on.

The sequence obtained, called the *Euclid–Mullin sequence*, is a bit bizarre:

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, 5471, \ldots$$

Big primes like 38709183810571 often appear before small primes like 11. It remains unknown whether or not every positive prime number appears in this list!

The chaotic nature of this sequence makes it difficult to extract information about how the primes are distributed: the numbers $p_1 \times \cdots \times p_n + 1$ grow very quickly—indeed, it must be the case that $p_1 \times \cdots \times p_n + 1 > 2^n$ for all n—so the upper bounds for the sequence grow at least exponentially.

Another proof of the infinitude of primes that gives a (slightly) tighter bound can be obtained using the following exercise.

Exercise 7.2.23

Let $n \in \mathbb{Z}$ with $n > 2$. Prove that the set $\{k \in \mathbb{Z} \mid n < k < n!\}$ contains a prime number.

Section 7.3

Modular arithmetic

Recall the definition of *congruence* modulo an integer from Section 5.2.

Definition 5.2.6

Fix $n \in \mathbb{Z}$. Given integers $a, b \in \mathbb{Z}$, we say a **is congruent to** b **modulo** n, and write

$$a \equiv b \pmod{n}$$ (LaTeX code: `a \equiv b \pmod{n}`)

if $n \mid b - a$. If a is not congruent to b modulo n, write

$$a \not\equiv b \pmod{n}$$ (LaTeX code: `\not\equiv`)

The number n is called the **modulus** of the congruence.

In Section 5.2, we proved that congruence is an equivalence relation:

Theorem 5.2.11

Let $n \in \mathbb{Z}$. Then congruence modulo n is an equivalence relation on $\mathbb{Z}$. That is

(a) $a \equiv a \pmod{n}$ for all $a \in \mathbb{Z}$;

(b) For all $a, b \in \mathbb{Z}$, if $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$;

(c) For all $a, b, c \in \mathbb{Z}$, if $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

In this section, we turn our attention to addition, subtraction, multiplication and division: our goal is to find out how much arithmetic can be done with *equality* replaced by *congruence*. For example:

(i) Can we add a number to both sides of a congruence? That is, given $a, b, c, n \in \mathbb{Z}$, is it the case that $a \equiv b \pmod{n}$ implies $a + c \equiv b + c \pmod{n}$?

(ii) Can we multiply both sides of a congruence by a number? That is, given $a, b, c, n \in \mathbb{Z}$, is it the case that $a \equiv b \pmod{n}$ implies $ac \equiv bc \pmod{n}$?

(iii) Can we divide both sides of a congruence by a nonzero common factor? That is, given $a, b, c, n \in \mathbb{Z}$ with $c \not\equiv 0 \pmod{n}$, is it the case that if $ac \equiv bc \pmod{n}$ implies $a \equiv b \pmod{n}$?

The answers to (i) and (ii) are 'yes', as we will prove; but surprisingly, the answer to (iii) is 'no' (except under certain circumstances). For example, $2 \times 3 \equiv 4 \times 3 \pmod{6}$, but $2 \not\equiv 4 \pmod{6}$, even though $3 \not\equiv 0 \pmod{6}$.

In light of this, it is important from the outset to point out that, although congruence is written with a symbol that looks like that of equality ('$\equiv$' vs. '$=$'), and although it is an equivalence relation, we can only treat congruence like equality inasmuch as we prove that we can. Specifically:

- In Theorem 5.2.11 we proved that congruence is an equivalence relation. This allows us to make some basic inferences about congruences—for example, transitivity means that the following im-

plication is valid:

$$-5 \equiv 18 \equiv 41 \equiv 64 \pmod{23} \quad \Rightarrow \quad -5 \equiv 64 \pmod{23}$$

- Theorem 7.3.3, which we will prove soon, tells us that we can treat congruence like equality for the purposes of addition, multiplication and subtraction. Thus it will be valid to write things like

$$x \equiv 7 \pmod{12} \quad \Rightarrow \quad 2x+5 \equiv 19 \pmod{12}$$

and we'll be able to replace values by congruent values in congruences, provided they're only being added, subtracted or multiplied. For example, from the knowledge that $2^{60} \equiv 1 \pmod{61}$ and $60! \equiv -1 \pmod{61}$, we will be able to deduce

$$2^{60} \cdot 3 \equiv 60! \cdot x \pmod{61} \quad \Rightarrow \quad 3 \equiv -x \pmod{61}$$

After we have worked out what arithmetic properties carry over to congruence, we will be able to prove some interesting theorems involving congruences and discuss their applications.

The first result we prove gives us a few equivalent ways of talking about congruence.

Proposition 7.3.1

Fix a modulus n and let $a, b \in \mathbb{Z}$. The following are equivalent:

(i) a and b leave the same remainder when divided by n;

(ii) $a = b + kn$ for some $k \in \mathbb{Z}$;

(iii) $a \equiv b \pmod{n}$.

Proof

We prove (i) $\Leftrightarrow$ (iii) and (ii) $\Leftrightarrow$ (iii).

- (i) $\Rightarrow$ (iii). Suppose a and b leave the same remainder when divided by n, and let $q_1, q_2, r \in \mathbb{Z}$ be such that

$$a = q_1 n + r, \quad b = q_2 n + r \quad \text{and} \quad 0 \leqslant r < n$$

Then $a - b = (q_1 - q_2)n$, which proves that $n \mid a - b$, and so $a \equiv b \pmod{n}$.

- (iii) $\Rightarrow$ (i). Suppose that $a \equiv b \pmod{n}$, so that $b - a = qn$ for some $q \in \mathbb{Z}$. Write

$$a = q_1 n + r_1, \quad b = q_2 n + r_2 \quad \text{and} \quad 0 \leqslant r_1, r_2 < n$$

We may further assume that $r_1 \leqslant r_2$. (If not, swap the roles of a and b—this is fine, since $n \mid b - a$ if and only if $n \mid a - b$.) Now we have

$$\begin{aligned} b - a = qn &\Rightarrow (q_2 n + r_2) - (q_1 n + r_1) = qn \\ &\Rightarrow (q_2 - q_1 - q)n + (r_2 - r_1) = 0 \qquad \text{rearranging} \end{aligned}$$

since $0 \leqslant r_1 \leqslant r_2 < n$ we have $0 \leqslant r_2 - r_1 < n$, so that $r_2 - r_1$ is the remainder of 0 when divided by n. That is, $r_2 - r_1 = 0$, so $r_1 = r_2$. Hence a and b have the same remainder when divided by n.

- (ii) $\Leftrightarrow$ (iii). We unpack the definitions of (ii) and (iii) to see that they are equivalent. Indeed

$$\begin{aligned}
&\text{(ii)} \Leftrightarrow a = b + kn \text{ for some } k \in \mathbb{Z} && \\
&\quad \Leftrightarrow a - b = kn \text{ for some } k \in \mathbb{Z} && \text{rearranging} \\
&\quad \Leftrightarrow n \mid a - b && \text{by definition of divisibility} \\
&\quad \Leftrightarrow a \equiv b \pmod{n} && \text{by definition of congruence} \\
&\quad \Leftrightarrow \text{(iii)} &&
\end{aligned}$$

Discussion 7.3.2

Where in the proof of Proposition 7.3.1 did we rely on the convention that the modulus n is positive? Is the result still true if n is negative?

We now prove that we can treat congruence like equality for the purposes of adding, subtracting and multiplying (but not dividing!) integers.

Theorem 7.3.3 · *Modular arithmetic*

Fix a modulus n, and let $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ be such that

$$a_1 \equiv a_2 \pmod{n} \qquad \text{and} \qquad b_1 \equiv b_2 \pmod{n}$$

Then the following congruences hold:

(a) $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$;

(b) $a_1 b_1 \equiv a_2 b_2 \pmod{n}$;

(c) $a_1 - b_1 \equiv a_2 - b_2 \pmod{n}$.

Proof

By Definition 5.2.6 that $n \mid a_1 - a_2$ and $n \mid b_1 - b_2$, so there exist $q_1, q_2 \in \mathbb{Z}$ such that

$$a_1 - a_2 = qn \qquad \text{and} \qquad b_1 - b_2 = rn$$

This implies that

$$(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) = qn + rn = (q + r)n$$

so $n \mid (a_1 + b_1) - (a_2 + b_2)$. This proves (a).

The algebra for (b) is slightly more involved:

$$\begin{aligned}
a_1 b_1 - a_2 b_2 &= (qn + a_2)(rn + b_2) - a_2 b_2 \\
&= qrn^2 + a_2 rn + b_2 qn + a_2 b_2 - a_2 b_2 \\
&= qrn^2 + a_2 rn + b_2 qn \\
&= (qrn + a_2 r + b_2 q)n
\end{aligned}$$

This shows that $n \mid a_1 b_1 - a_2 b_2$, thus proving (b).

Now (a) and (b) together imply (c). Indeed, we know that $-1 \equiv -1 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$, so by (b) we have $-b_1 \equiv -b_2 \pmod{n}$. We also know that $a_1 \equiv a_2 \pmod{n}$, and hence $a_1 - b_1 \equiv a_2 - b_2 \pmod{n}$ by (a). □

Theorem 7.3.3 allows us to perform algebraic manipulations with congruences as if they were equations, provided all we're doing is adding, multiplying and subtracting.

Example 7.3.4

We will solve the congruence $3x - 5 \equiv 2x + 3 \pmod{7}$ for x:

$$\begin{aligned}
& 3x - 5 \equiv 2x + 3 \pmod{7} && && \\
\Leftrightarrow\ & x - 5 \equiv 3 \pmod{7} && (\Rightarrow) \text{ subtract } 2x && (\Leftarrow) \text{ add } 2x \\
\Leftrightarrow\ & x \equiv 8 \pmod{7} && (\Rightarrow) \text{ add } 5 && (\Leftarrow) \text{ subtract } 5 \\
\Leftrightarrow\ & x \equiv 1 \pmod{7} && \text{since } 8 \equiv 1 \pmod{7} &&
\end{aligned}$$

So the integers x for which $3x - 5$ and $2x + 3$ leave the same remainder when divided by 7, are precisely the integers x which leave a remainder of 1 when divided by 7:

$$3x - 5 \equiv 2x + 3 \pmod{7} \quad \Leftrightarrow \quad x = 7q + 1 \text{ for some } q \in \mathbb{Z}$$

Exercise 7.3.5

For which integers x does the congruence $5x + 1 \equiv x + 8 \pmod{3}$ hold? Characterise such integers x in terms of their remainder when divided by 3.

So far this all feels like we haven't done very much: we've just introduced a new symbol $\equiv$ which behaves just like equality... but does it really? The following exercises should expose some more ways in which congruence *does* behave like equality, and some in which it *doesn't*.

Exercise 7.3.6

Fix a modulus n. Is it true that

$$a \equiv b \pmod{n} \quad \Rightarrow \quad a^k \equiv b^k \pmod{n}$$

for all $a, b \in \mathbb{Z}$ and $k \in \mathbb{N}$? If so, prove it; if not, provide a counterexample.

Exercise 7.3.7

Fix a modulus n. Is it true that

$$k \equiv \ell \pmod{n} \quad \Rightarrow \quad a^k \equiv a^\ell \pmod{n}$$

for all $k, \ell \in \mathbb{N}$ and $a \in \mathbb{Z}$? If so, prove it; if not, provide a counterexample.

Exercise 7.3.8

Fix a modulus n. Is it true that

$$qa \equiv qb \pmod{n} \quad \Rightarrow \quad a \equiv b \pmod{n}$$

for all $a, b, q \in \mathbb{Z}$ with $q \not\equiv 0 \pmod{n}$? If so, prove it; if not, provide a counterexample.

Example 7.3.9

Now that we have seen several things that we *can* do with modular arithmetic, let's look at some things that we *cannot* do:

- We cannot talk about fractions in modular arithmetic; for instance, it is invalid to say $2x \equiv 1 \pmod{5}$ implies $x \equiv \frac{1}{2} \pmod{5}$.
- We cannot take square roots in modular arithmetic; for instance, it is invalid to say $x^2 \equiv 3 \pmod{4}$ implies $x \equiv \pm\sqrt{3} \pmod{4}$. In fact, it is invalid to say $x^2 \equiv 1 \pmod{8}$ implies $x \equiv \pm 1 \pmod{8}$, since for example $3^2 \equiv 1 \pmod{8}$ but $3 \not\equiv \pm 1 \pmod{8}$.
- We cannot replace numbers in exponents by other numbers they are congruent to; for instance, it is invalid to say $x^3 \equiv 2^3 \pmod{4}$ implies $x \equiv 2 \pmod{4}$.

Multiplicative inverses

We made a big deal about the fact that fractions don't make sense in modular arithmetic. That is, it is invalid to say

$$2x \equiv 1 \pmod{5} \quad \Rightarrow \quad x \equiv \frac{1}{2} \pmod{5}$$

Despite this, we can still make sense of 'division', provided we change what we mean when we say 'division'. Indeed, the congruence $2x \equiv 1 \pmod{5}$ has a solution:

$$\begin{aligned} & 2x \equiv 1 \pmod{5} && && \\ \Leftrightarrow\ & 6x \equiv 3 \pmod{5} && (\Rightarrow) \text{ multiply by } 3 && (\Leftarrow) \text{ subtract } 3 \\ \Leftrightarrow\ & x \equiv 3 \pmod{5} && \text{since } 6 \equiv 1 \pmod{5} && \end{aligned}$$

Here we didn't divide by 2, but we still managed to cancel the 2 by instead multiplying through by 3. For the purposes of solving the equation this had the same effect as division by 2 would have had if we were allowed to divide. The key here was that $2 \times 3 \equiv 1 \pmod{5}$.

Definition 7.3.10

Fix a modulus n. Given $a \in \mathbb{Z}$, a **multiplicative inverse** for a modulo n is an integer u such that $au \equiv 1 \pmod{n}$.

Example 7.3.11

Some examples of multiplicative inverses are as follows:

- 2 is a multiplicative inverse of itself modulo 3, since $2 \times 2 \equiv 4 \equiv 1 \pmod{3}$.

- 2 is a multiplicative inverse of 3 modulo 5, since $2 \times 3 \equiv 6 \equiv 1 \pmod 5$.
- 7 is also a multiplicative inverse of 3 modulo 5, since $3 \times 7 \equiv 21 \equiv 1 \pmod 5$.
- 3 has no multiplicative inverse modulo 6. Indeed, suppose $u \in \mathbb{Z}$ with $3u \equiv 1 \pmod 6$. Then $6 \mid 3u - 1$, so $3u - 1 = 6q$ for some $q \in \mathbb{Z}$. But then

  $$1 = 3u - 6q = 3(u - 2q)$$

  which implies that $3 \mid 1$, which is nonsense.

Knowing when multiplicative inverses exist is very important for solving congruences: if u is a multiplicative inverse for a modulo n, then we can solve equations of the form $ax \equiv b \pmod n$ extremely easily:

$$ax \equiv b \pmod n \quad \Rightarrow \quad x \equiv ub \pmod n$$

Exercise 7.3.12

For $n = 7, 8, 9, 10, 11, 12$, either find a multiplicative inverse for 6 modulo n, or show that no multiplicative inverse exists. Can you spot a pattern?

Some authors write a^{-1} to denote multiplicative inverses. We refrain from this, since it suggests that multiplicative inverses are unique—but they're not, as you'll see in the following exercise.

Exercise 7.3.13

Let n be a modulus and let $a \in \mathbb{Z}$. Suppose that u is a multiplicative inverse for a modulo n. Prove that, for all $k \in \mathbb{Z}$, $u + kn$ is a multiplicative inverse for a modulo n.

Proposition 7.3.14

Let $a \in \mathbb{Z}$ and let n be a modulus. Then a has a multiplicative inverse modulo n if and only if $a \perp n$.

Proof

Note that a has a multiplicative inverse u modulo n if and only if there is a solution (u, v) to the equation $au + nv = 1$. Indeed, $au \equiv 1 \pmod n$ if and only if $n \mid au - 1$, which occurs if and only if there is some $q \in \mathbb{Z}$ such that $au - 1 = nq$. Setting $q = -v$ and rearranging yields the desired equivalence.

By Bézout's lemma (Theorem 7.1.26), such a solution (u, v) exists if and only if $\gcd(a, n) \mid 1$. This occurs if and only if $\gcd(a, n) = 1$, i.e. if and only if $a \perp n$. □

Proof tip

To solve a congruence of the form $ax \equiv b \pmod n$ when $a \perp n$, first find a multiplicative inverse u for a modulo n, and then simply multiply through by u to obtain $x \equiv ub \pmod n$.

Corollary 7.3.15
Let $a, p \in \mathbb{Z}$, where p is a positive prime. If $p \nmid a$ then a has a multiplicative inverse modulo p.

Proof
Suppose $p \nmid a$, and let $d = \gcd(a, p)$. Since $d \mid p$ and p is prime we have $d = 1$ or $d = p$. Since $d \mid a$ and $p \nmid a$ we can't have $d = p$; therefore $d = 1$. By Proposition 7.3.14, a has a multiplicative inverse modulo p. □

Example 7.3.16
11 is prime, so each of the integers a with $1 \leqslant a \leqslant 10$ should have a multiplicative inverse modulo 11. And indeed, the following are all congruent to 1 modulo 11:

$$\begin{array}{lllll} 1 \times 1 = 1 & 2 \times 6 = 12 & 3 \times 4 = 12 & 4 \times 3 = 12 & 5 \times 9 = 45 \\ 6 \times 2 = 12 & 7 \times 8 = 56 & 8 \times 7 = 56 & 9 \times 5 = 45 & 10 \times 10 = 100 \end{array}$$

Exercise 7.3.17
Find all integers x such that $25x - 4 \equiv 4x + 3 \pmod{13}$.

Orders and totients

For any modulus n, there are only finitely many possible remainders modulo n. A nice consequence of this finiteness is that, when $a \perp n$, we can choose some power of a to be its multiplicative inverse, as proved in the following exercise.

Exercise 7.3.18
Let n be a modulus and let $a \in \mathbb{Z}$ with $a \perp n$. Prove that there exists $k \geqslant 1$ such that $a^k \equiv 1 \pmod{n}$.

Exercise 7.3.18, together with the well-ordering principle, justify the following definition.

Definition 7.3.19
Let n be a modulus and let $a \in \mathbb{Z}$ with $a \perp n$. The **order** of a modulo n is the least $k \geqslant 1$ such that $a^k \equiv 1 \pmod{n}$.

Note that this definition makes sense by Exercise 7.3.18 and the well-ordering principle.

Example 7.3.20
The powers of 7 modulo 100 are:

- $7^1 = 7$, so $7^1 \equiv 7 \pmod{100}$;
- $7^2 = 49$, so $7^2 \equiv 49 \pmod{100}$;

- $7^3 = 343$, so $7^3 \equiv 43 \pmod{100}$;
- $7^4 = 2401$, so $7^4 \equiv 1 \pmod{100}$.

Hence the order of 7 modulo 100 is 4, and 7^3 and 43 are multiplicative inverses of 7 modulo 100.

Our focus turns to computing specific values of k such that $a^k \equiv 1 \pmod{n}$, whenever $a \in \mathbb{Z}$ and $a \perp n$. We first focus on the case when n is prime; then we develop the machinery of *totients* to study the case when n is not prime.

Lemma 7.3.21
Let $a, b \in \mathbb{Z}$ and let $p \in \mathbb{Z}$ be a positive prime. Then $(a+b)^p \equiv a^p + b^p \pmod{p}$.

Proof
By the binomial theorem (Theorem 4.2.17), we have

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^k b^{p-k}$$

By Exercise 7.2.5, $p \mid \binom{p}{k}$ for all $0 < k < p$, and hence $\binom{p}{k} a^k b^{p-k} \equiv 0 \pmod{p}$ for all $0 < k < p$. Thus

$$(a+b)^p \equiv \binom{p}{0} a^0 b^{p-0} + \binom{p}{p} a^p b^{p-p} \equiv a^p + b^p \pmod{p}$$

as desired. □

Theorem 7.3.22 · *Fermat's little theorem*
Let $a, p \in \mathbb{Z}$ with p a positive prime. Then $a^p \equiv a \pmod{p}$.

Proof
We may assume that $a \geqslant 0$, otherwise replace a by its remainder modulo p.

We will prove that $a^p \equiv a \pmod{p}$ by induction on a.

- (**Base case**) Since $p > 0$ we have $0^p = 0$, hence $0^p \equiv 0 \pmod{p}$.
- (**Induction step**) Fix $a \geqslant 0$ and suppose $a^p \equiv a \pmod{p}$. Then $(a+1)^p \equiv a^p + 1^p \pmod{p}$ by Lemma 7.3.21. Now $a^p \equiv a \pmod{p}$ by the induction hypothesis, and $1^p = 1$, so we have $(a+1)^p \equiv a+1 \pmod{p}$.

By induction, we're done. □

The following consequence of Theorem 7.3.22 is often also referred to as 'Fermat's little theorem', but is slightly less general since it requires an additional hypothesis. In keeping with the wider mathematical community, we will refer to both Theorem 7.3.22 and Corollary 7.3.23 as 'Fermat's little theorem'.

Corollary 7.3.23 · *Fermat's little theorem — alternative version*
Let $a, p \in \mathbb{Z}$ with p a positive prime and $p \nmid a$. Then $a^{p-1} \equiv 1 \pmod{p}$.

Proof
Since $p \nmid a$, it follows that $a \perp p$. Theorem 7.3.22 tells us that $a^p \equiv a \pmod{p}$. By Proposition 7.3.14, a has a multiplicative inverse b modulo p. Hence

$$a^p b \equiv ab \pmod{p}$$

But $a^p b \equiv a^{p-1} ab \pmod{p}$, and $ab \equiv 1 \pmod{p}$, so we get

$$a^{p-1} \equiv 1 \pmod{p}$$

as required. □

Corollary 7.3.23 can be useful for computing remainders of humongous numbers when divided by smaller primes.

Example 7.3.24
We compute the remainder of 2^{1000} when divided by 7. Since $7 \nmid 2$, it follows from Fermat's little theorem (Corollary 7.3.23) that $2^6 \equiv 1 \pmod{7}$. Now $1000 = 166 \times 6 + 4$, so

$$2^{1000} \equiv 2^{166\times 6+4} \equiv (2^6)^{166} \cdot 2^4 \equiv 2^4 \equiv 16 \equiv 2 \pmod{7}$$

so the remainder of 2^{1000} when divided by 7 is 2.

Exercise 7.3.25
Find the remainder of 3^{244886} when divided by 13.

Unfortunately, the hypothesis that p is prime in Fermat's little theorem cannot be disposed of. For example, 6 is not prime, and $5^{6-1} = 5^5 = 3125 = 520 \times 6 + 5$, so $5^5 \equiv 5 \pmod{6}$. Our next order of business is to generalise Corollary 7.3.23 by removing the requirement that the modulus p be prime, and replacing $p - 1$ by the *totient* of the modulus.

Definition 7.3.26
Let $n \in \mathbb{Z}$. The **totient** of n is the natural number $\varphi(n)$ (LaTeX code: `\varphi(n)`) defined by

$$\varphi(n) = |\{k \in [|n|] \mid k \perp n\}|$$

That is, $\varphi(n)$ is the number of natural numbers less than $|n|$ which are coprime to n. The function $\varphi : \mathbb{Z} \to \mathbb{N}$ is called **Euler's totient function**.

Example 7.3.27
Here are some examples of totients:

- The elements of $[6]$ which are coprime to 6 are 1 and 5, so $\varphi(6) = 2$.

- If p is a positive prime, then by Corollary 7.2.20, every element of $[p]$ is coprime to p except for p itself. Hence if p is a positive prime then $\varphi(p) = p - 1$. More generally, if p is prime then $\varphi(p) = |p| - 1$.
- $[1] = \{0\}$ and $\gcd(0,1) = 1$, so that $0 \perp 1$, and hence $\varphi(1) = 1$.

Exercise 7.3.28
Let $n \in \mathbb{Z}$ and let $p > 0$ be prime. Prove that if $p \mid n$, then $\varphi(pn) = p \cdot \varphi(n)$. Deduce that $\varphi(p^k) = p^k - p^{k-1}$ for all prime $p > 0$ and all $k \geqslant 1$.

Exercise 7.3.29
Let $n \in \mathbb{Z}$ and let $p > 0$ be prime. Prove that if $p \nmid n$, then $\varphi(pn) = (p-1)\varphi(n)$.

Together, Exercises 7.3.28 and 7.3.29 allow us to compute the totient of any integer with up to two primes in its prime factorisation.

Example 7.3.30
We compute $\varphi(100)$. The prime factorisation of 100 is $2^2 \times 5^2$. Applying Exercise 7.3.28 twice

$$\varphi(2^2 \times 5^2) = 2 \times 5 \times \varphi(2 \times 5) = 10\varphi(10)$$

Finally, Exercise 7.3.29 tells us that

$$\varphi(10) = \varphi(2 \times 5) = 1 \times \varphi(5) = 1 \times 4 = 4$$

Hence $\varphi(100) = 40$.

Exercise 7.3.31
Prove that $\varphi(100) = 40$, this time using the inclusion–exclusion principle.

Euler's theorem uses totients to generalise Fermat's little theorem (Theorem 7.3.22) to arbitrary moduli, not just prime ones.

Theorem 7.3.32 · *Euler's theorem*
Let n be a modulus and let $a \in \mathbb{Z}$ with $a \perp n$. Then

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

Proof
By definition of totient, the set X defined by

$$X = \{k \in [n] \mid k \perp n\}$$

has $\varphi(n)$ elements. List the elements as

$$X = \{x_1, x_2, \ldots, x_{\varphi(n)}\}$$

Note that $ax_i \perp n$ for all i, so let y_i be the (unique) element of X such that $ax_i \equiv y_i \pmod{n}$.

Note that if $i \neq j$ then $y_i \neq y_j$. We prove this by contraposition; indeed, since $a \perp n$, by Proposition 7.3.14, a has a multiplicative inverse, say b. Then

$$y_i \equiv y_j \pmod{n} \Rightarrow ax_i \equiv ax_j \pmod{n} \Rightarrow bax_i \equiv bax_j \pmod{n} \Rightarrow x_i \equiv x_j \pmod{n}$$

and $x_i \equiv x_j \pmod{n}$ if and only if $i = j$. Thus

$$X = \{x_1, x_2, \ldots, x_{\varphi(n)}\} = \{y_1, y_2, \ldots, y_{\varphi(n)}\}$$

This means that the product of the 'x_i's is equal to the product of the 'y_i's, and hence

$$\begin{aligned}
&x_1 \cdot \ldots \cdot x_{\varphi(n)} \\
&\equiv y_1 \cdot \ldots \cdot y_{\varphi(n)} \pmod{n} && \text{since } \{x_1, \ldots\} = \{y_1, \ldots\} \\
&\equiv (ax_1) \cdot \ldots \cdot (ax_{\varphi(n)}) \pmod{n} && \text{since } y_i \equiv ax_i \pmod{n} \\
&\equiv a^{\varphi(n)} \cdot x_1 \cdot \ldots \cdot x_{\varphi(n)} \pmod{n} && \text{rearranging}
\end{aligned}$$

Since each x_i is coprime to n, we can cancel the x_i terms (by multiplying by their multiplicative inverses) to obtain

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

as required. □

Example 7.3.33

Some examples of Euler's theorem in action are as follows:

- We have seen that $\varphi(6) = 2$, and we know that $5 \perp 6$. And, indeed,

  $$5^{\varphi(6)} = 5^2 = 25 = 4 \times 6 + 1$$

  so $5^{\varphi(6)} \equiv 1 \pmod{6}$.

- By Exercise 7.3.28, we have

  $$\varphi(121) = \varphi(11^2) = 11^2 - 11^1 = 121 - 11 = 110$$

  Moreover, given $a \in \mathbb{Z}$, $a \perp 121$ if and only if $11 \nmid a$ by Corollary 7.2.20. Hence $a^{110} \equiv 1 \pmod{121}$ whenever $11 \nmid a$.

Exercise 7.3.34

Use Euler's theorem to prove that the last two digits of 3^{79} are '67'.

Example 7.3.35

Let n be a modulus and let $a \in \mathbb{Z}$ with $a \perp n$. Prove that the order of a modulo n divides $\varphi(n)$.

A formula for the totient of an arbitrary nonzero integer is proved in Theorem 7.3.59—its proof is an application of the Chinese remainder theorem Theorem 7.3.46, and uses the techniques for counting finite sets discussed in Section 8.1.

Wilson's theorem

We conclude this chapter on number theory with *Wilson's theorem*, which is a nice result that completely characterises prime numbers in the sense that we can tell when a number is prime by computing the remainder of $(n-1)!$ when divided by n.

Let's test a few numbers first:

n	$(n-1)!$	remainder
2	1	1
3	2	2
4	6	2
5	24	4
6	120	0
7	720	6
8	5040	0

n	$(n-1)!$	remainder
9	40320	0
10	362880	0
11	3628800	10
12	39916800	0
13	479001600	12
14	6227020800	0
15	87178291200	0

It's tempting to say that an integer $n > 1$ is prime if and only if $n \nmid (n-1)!$, but this isn't true since it fails when $n = 4$. But it's extremely close to being true.

Theorem 7.3.36 · *Wilson's theorem*
Let $n > 1$ be a modulus. Then n is prime if and only if $(n-1)! \equiv -1 \pmod{n}$.

The following sequence of exercises will piece together into a proof of Wilson's theorem.

Exercise 7.3.37
Let $n \in \mathbb{Z}$ be composite. Prove that if $n > 4$, then $n \mid (n-1)!$.

Exercise 7.3.38
Let p be a positive prime and let $a \in \mathbb{Z}$. Prove that, if $a^2 \equiv 1 \pmod{p}$, then $a \equiv 1 \pmod{p}$ or $a \equiv -1 \pmod{p}$.

Exercise 7.3.38 implies that the only elements of $[p-1]$ that are their own multiplicative inverses are 1 and $p-1$; this morsel of information allows us to deduce result in the following exercise.

Exercise 7.3.39
Let p be a positive prime. Prove that $(p-1)! \equiv -1 \pmod{p}$.

Proof of Wilson's theorem (Theorem 7.3.36)
Let $n > 1$ be a modulus.

- If n is prime, then $(n-1)! \equiv -1 \pmod{n}$ by Exercise 7.3.39.
- If n is composite, then either $n = 4$ or $n > 4$. If $n = 4$ then

$$(n-1)! = 3! = 6 \equiv 2 \pmod{4}$$

and so $(n-1)! \not\equiv -1 \pmod{n}$. If $n > 4$, then

$$(n-1)! \equiv 0 \pmod{n}$$

by Exercise 7.3.37.

Hence $(n-1)! \equiv -1 \pmod{n}$ if and only if n is prime, as desired. □

Since Wilson's theorem completely characterises the positive prime numbers, we could have defined 'n is prime', for $n > 1$, to mean that $(n-1)! \equiv -1 \pmod{n}$. We don't do this because, although this is an interesting result, it is not particularly useful in applications. We might even hope that Wilson's theorem gives us an easy way to test whether a number is prime, but unfortunately even this is a bust: computing the remainder $(n-1)!$ on division by n is not particularly efficient.

However, there are some nice applications of Wilson's theorem, which we will explore now.

Example 7.3.40
We'll compute the remainder of $3^{45} \cdot 44!$ when divided by 47. Note that $3^{45} \cdot 44!$ is equal to a monstrous number with 76 digits; I don't recommend doing the long division! Anyway...

- 47 is prime, so we can apply both Fermat's little theorem (Theorem 7.3.22) and Wilson's theorem (Theorem 7.3.36).
- By Fermat's little theorem, we know that $3^{46} \equiv 1 \pmod{47}$. Since $3 \cdot 16 = 48 \equiv 1 \pmod{47}$, we have

$$3^{45} \equiv 3^{45} \cdot (3 \cdot 16) \equiv 3^{46} \cdot 16 \equiv 16 \pmod{47}$$

- By Wilson's theorem, we have $46! \equiv -1 \pmod{47}$. Now
  - $46 \equiv -1 \pmod{47}$, so 46 is its own multiplicative inverse modulo 47.
  - The extended Euclidean algorithm yields $45 \cdot 23 \equiv 1 \pmod{47}$.

  So we have

$$44! = 44! \cdot (45 \cdot 23) \cdot (46 \cdot 46) \equiv 46! \cdot 23 \cdot 46 \equiv (-1) \cdot 23 \cdot (-1) \equiv 23 \pmod{47}$$

Putting this information together yields

$$3^{45} \cdot 44! \equiv 16 \cdot 23 = 368 \equiv 39 \pmod{47}$$

So the remainder left when $3^{45} \cdot 44!$ is divided by 47 is 39.

Exercise 7.3.41

Let p be an odd positive prime. Prove that

$$\left[\left(\frac{p-1}{2}\right)!\right]^2 \equiv (-1)^{\frac{p+1}{2}} \pmod{p}$$

Chinese remainder theorem

We introduce the Chinese remainder theorem with an example.

Example 7.3.42

We find all integer solutions x to the system of congruences

$$x \equiv 2 \pmod{5} \quad \text{and} \quad x \equiv 4 \pmod{8}$$

Note that $x \equiv 4 \pmod{8}$ if and only if $x = 4 + 8k$ for some $k \in \mathbb{Z}$. Now, for all $k \in \mathbb{Z}$ we have

$x \equiv 2 \pmod{5}$	
$\Leftrightarrow 4 + 8k \equiv 2 \pmod{5}$	since $x = 4 + 8k$
$\Leftrightarrow 8k \equiv -2 \pmod{5}$	subtracting 4
$\Leftrightarrow 3k \equiv 3 \pmod{5}$	since $8 \equiv -2 \equiv 3 \pmod{5}$
$\Leftrightarrow k \equiv 1 \pmod{5}$	multiplying by a multiplicative inverse for 3 modulo 5

So $4 + 8k \equiv 2 \pmod{5}$ if and only if $k = 1 + 5\ell$ for some $\ell \in \mathbb{Z}$.

Combining this, we see that x satisfies both congruences if and only if

$$x = 4 + 8(1 + 5\ell) = 12 + 40\ell$$

for some $\ell \in \mathbb{Z}$.

Hence the integers x for which both congruences are satisfied are precisely those integers x such that $x \equiv 12 \pmod{40}$.

Exercise 7.3.43

Find all integer solutions x to the system of congruences:

$$\begin{cases} x \equiv -1 \pmod{4} \\ x \equiv 1 \pmod{9} \\ x \equiv 5 \pmod{11} \end{cases}$$

Express your solution in the form $x \equiv a \pmod{n}$ for suitable $n > 0$ and $0 \leqslant a < n$.

Exercise 7.3.44

Let m, n be coprime moduli and let $a, b \in \mathbb{Z}$. Let $u, v \in \mathbb{Z}$ be such that

$$mu \equiv 1 \pmod{n} \quad \text{and} \quad nv \equiv 1 \pmod{m}$$

In terms of a, b, m, n, u, v, find an integer x such that

$$x \equiv a \pmod{m} \quad \text{and} \quad x \equiv b \pmod{n}$$

Exercise 7.3.45

Let m, n be coprime moduli and let $x, y \in \mathbb{Z}$. Prove that if $x \equiv y \pmod{m}$ and $x \equiv y \pmod{n}$, then $x \equiv y \pmod{mn}$.

Theorem 7.3.46 · *Chinese remainder theorem*

Let m, n be moduli and let $a, b \in \mathbb{Z}$. If m and n are coprime, then there exists an integer solution x to the simultaneous congruences

$$x \equiv a \pmod{m} \quad \text{and} \quad x \equiv b \pmod{n}$$

Moreover, if $x, y \in \mathbb{Z}$ are two such solutions, then $x \equiv y \pmod{mn}$.

Proof

Existence of a solution x is precisely the content of Exercise 7.3.44.

Now let $x, y \in \mathbb{Z}$ be two solutions to the two congruences. Then

$$\begin{cases} x \equiv a \pmod{m} \\ y \equiv a \pmod{m} \end{cases} \quad \Rightarrow \quad x \equiv y \pmod{m}$$

$$\begin{cases} x \equiv b \pmod{n} \\ y \equiv b \pmod{n} \end{cases} \quad \Rightarrow \quad x \equiv y \pmod{n}$$

so by Exercise 7.3.45, we have $x \equiv y \pmod{mn}$, as required. □

We now generalise the Chinese remainder theorem to the case when the moduli m, n are not assumed to be coprime. There are two ways we could make this generalisation: either we could reduce the more general version of the theorem to the version we proved in Theorem 7.3.46, or we could prove the more general version from scratch. We opt for the latter approach, but you might want to consider what a 'reductive' proof would look like.

Theorem 7.3.47

Let m, n be moduli and let $a, b \in \mathbb{Z}$. There exists an integer solution x to the system of congruences

$$x \equiv a \pmod{m} \quad \text{and} \quad x \equiv b \pmod{n}$$

if and only if $a \equiv b \pmod{\gcd(m,n)}$.

Moreover, if $x, y \in \mathbb{Z}$ are two such solutions, then $x \equiv y \pmod{\operatorname{lcm}(m,n)}$

Proof
Let $d = \gcd(m,n)$, and write $m = m'd$ and $n = n'd$ for some $m', n' \in \mathbb{Z}$.

We prove that an integer solution x to the system of congruences exists if and only if $a \equiv b \pmod{d}$.

- ($\Rightarrow$) Suppose an integer solution x to the system of congruences exists. Then there exist integers k, ℓ such that
$$x = a + mk = b + n\ell$$
But $m = m'd$ and $n = n'd$, so we have $a + m'dk = b + n'd\ell$, and so
$$a - b = (n'\ell - m'k)d$$
so that $a \equiv b \pmod{d}$, as required.
- ($\Leftarrow$) Suppose $a \equiv b \pmod{d}$, and let $t \in \mathbb{Z}$ be such that $a - b = td$. Let $u, v \in \mathbb{Z}$ be such that $mu + nv = d$—these exist by Bézout's lemma (Theorem 7.1.26). Note also that, since $m = m'd$ and $n = n'd$, dividing through by d yields $m'u + n'v = 1$.
Define
$$x = an'v + bm'u$$
Now we have
$$\begin{aligned} x &= an'v + bm'u && \text{by definition of } x \\ &= an'v + (a - td)m'u && \text{since } a - b = td \\ &= a(m'u + n'v) - tdm'u && \text{rearranging} \\ &= a - tdm'u && \text{since } m'u + n'v = 1 \\ &= a - tum && \text{since } m = m'd \end{aligned}$$
so $x \equiv a \pmod{m}$. Likewise
$$\begin{aligned} x &= an'v + bm'u && \text{by definition of } x \\ &= (b + td)n'v + bm'u && \text{since } a - b = td \\ &= b(m'u + n'v) + tdn'v && \text{rearranging} \\ &= b + tdn'v && \text{since } m'u + n'v = 1 \\ &= b + tvn && \text{since } n = n'd \end{aligned}$$
so $x \equiv b \pmod{n}$.
Hence $x = an'v + bm'u$ is a solution to the system of congruences.

We now prove that if x, y are two integer solutions to the system of congruences, then they are congruent modulo $\operatorname{lcm}(a,b)$. First note that we must have
$$x \equiv y \pmod{m} \quad \text{and} \quad x \equiv y \pmod{n}$$

so that $x = y + km$ and $x = y + \ell n$ for some $k, \ell \in \mathbb{Z}$. But then

$$x - y = km = \ell n$$

Writing $m = m'd$ and $n = n'd$, we see that $km'd = \ell n'd$, so that $km' = \ell n'$. But m', n' are coprime by Exercise 7.1.33, and hence $m' \mid \ell$ by Proposition 7.1.35. Write $\ell = \ell' m'$ for some $\ell' \in \mathbb{Z}$. Then we have

$$x - y = \ell n = \ell' m' n$$

and hence $x \equiv y \pmod{m'n}$. But $m'n = \operatorname{lcm}(m, n)$ by Exercise 7.1.44. □

This theorem is in fact *constructive*, in that it provides an algorithm for finding all integer solutions x to a system of congruences

$$x \equiv a \pmod{m} \quad \text{and} \quad x \equiv b \pmod{n}$$

as follows:

- Use the Euclidean algorithm to compute $d = \gcd(m, n)$.
- If $d \nmid a - b$ then there are no solutions, so stop. If $d \mid a - b$, then proceed to the next step.
- Use the extended Euclidean algorithm to compute $u, v \in \mathbb{Z}$ such that $mu + nv = d$.
- The integer solutions x to the system of congruences are precisely those of the form

$$x = \frac{anv + bmu + kmn}{d} \quad \text{for some } k \in \mathbb{Z}$$

✎ **Exercise 7.3.48**

Verify that the algorithm outlined above is correct. Use it to compute the solutions to the system of congruences

$$x \equiv 3 \pmod{12} \quad \text{and} \quad x \equiv 15 \pmod{20}$$

✎ ⋆ **Exercise 7.3.49**

Generalise the Chinese remainder theorem to systems of arbitrarily (finitely) many congruences. That is, given $r \in \mathbb{N}$, find precisely the conditions on moduli $n_1, n_2, \ldots, n_r$ and integers $a_1, a_2, \ldots, a_r$ such that an integer solution exists to the congruences

$$x \equiv a_1 \pmod{n_1}, \quad x \equiv a_2 \pmod{n_2}, \quad \cdots \quad x_r \equiv a_r \pmod{n_r}$$

Find an explicit formula for such a value of x, and find a suitable modulus n in terms of $n_1, n_2, \ldots, n_r$ such that any two solutions to the system of congruences are congruent modulo n.

✎ **Exercise 7.3.50**

Prove that gaps between consecutive primes can be made arbitrarily large. That is, prove that for all $n \in \mathbb{N}$, there exists an integer a such that the numbers

$$a,\ a+1,\ a+2,\ \ldots,\ a+n$$

are all composite.

Application: tests for divisibility

The language of modular arithmetic provides a practical setting for proving tests for divisibility using number bases. Number bases were introduced in Chapter 0, and we gave a preliminary definition in Definition 0.31 of what a number base is. Our first job will be to justify why this definition makes sense at all—that is, we need to prove that every natural number *has* a base-b expansion, and moreover, that it only has one of them. Theorem 7.3.51 says exactly this.

Theorem 7.3.51

Let $n \in \mathbb{N}$ and let $b \in \mathbb{N}$ with $b \geqslant 2$. Then there exist unique $r \in \mathbb{N}$ and $d_0, d_1, \ldots, d_r \in \{0, 1, \ldots, b-1\}$ such that

$$n = \sum_{i=0}^{r} d_i b^i$$

and such that $d_r \neq 0$, except $n = 0$, in which case $r = 0$ and $d_0 = 0$.

Proof

We proceed by strong induction on n.

- (**BC**) We imposed the requirement that if $n = 0$ then $r = 0$ and $d_0 = 0$; and this evidently satisfies the requirement that $n = \sum_{i=0}^{r} d_i b^i$.
- (**IS**) Fix $n \geqslant 0$ and suppose that the requirements of the theorem are satisfied for all the natural numbers up to and including n.

  By the division theorem (Theorem 7.1.1), there exist unique $u, v \in \mathbb{N}$ such that

  $$n + 1 = ub + v \quad \text{and} \quad v \in \{0, 1, \ldots, b-1\}$$

  Since $b \geqslant 2$, we have $u < n + 1$, and so $u \leqslant n$. It follows from the induction hypothesis that there exist unique $r \in \mathbb{N}$ and $d_1, \ldots, d_r \in \{0, 1, \ldots, b-1\}$ such that

  $$u = \sum_{i=0}^{r} d_{i+1} b^i$$

  and $d_r \neq 0$. Writing $d_0 = v$ yields

  $$n = ub + v = \sum_{i=0}^{r} d_{i+1} b^{i+1} + d_0 = \sum_{i=0}^{r} d_i b^i$$

  Since $d_r \neq 0$, this proves existence.

  For uniqueness, suppose that there exists $s \in \mathbb{N}$ and $e_0, \ldots, e_s \in \{0, 1, \ldots, b-1\}$ such that

  $$n + 1 = \sum_{j=0}^{s} e_j b^j$$

  and $e_s \neq 0$. Then

  $$n + 1 = \left(\sum_{j=1}^{s} e_j b^{j-1} \right) b + e_0$$

so by the division theorem we have $e_0 = d_0 = v$. Hence

$$u = \frac{n+1-v}{b} = \sum_{j=1}^{s} e_j b^{j-1} = \sum_{i=1}^{r} d_i b^{j-1}$$

so by the induction hypothesis, it follows that $r = s$ and $d_i = e_i$ for all $1 \leqslant i \leqslant r$. This proves uniqueness.

By induction, we're done. □

We now re-state the definition of base-b expansion, confident in the knowledge that this definition makes sense.

Definition 7.3.52

Let $n \in \mathbb{N}$. The **base-b expansion** of n is the unique string $d_r d_{r-1} \dots d_0$ such that the conditions in Theorem 7.3.51 are satisfied. The base-2 expansion is also known as the **binary expansion**, and the base-10 expansion is called the **decimal expansion**.

Example 7.3.53

Let $n \in \mathbb{N}$. Then n is divisible by 3 if and only if the sum of the digits in the decimal expansion of n is divisible by 3. Likewise, n is divisible by 9 if and only if the sum of the digits in the decimal expansion n is divisible by 9.

We prove this for divisibility by 3. Let

$$n = d_r d_{r-1} \cdots d_1 d_0$$

be the decimal expansion of n, and let $s = \sum_{i=0}^{r} d_i$ be the sum of the digits of n.

Then we have

$$\begin{aligned}
n &\equiv \sum_{i=0}^{r} d_i 10^i \pmod 3 && \text{since } n = \sum_i d_i 10^i \\
&\equiv \sum_{i=0}^{r} d_i 1^i \pmod 3 && \text{since } 10 \equiv 1 \pmod 3 \\
&\equiv \sum_{i=0}^{r} d_i && \text{since } 1^i = 1 \text{ for all } i \\
&\equiv s && \text{by definition of } s
\end{aligned}$$

Since $n \equiv s \pmod 3$, it follows that n is divisible by 3 if and only if s is divisible by 3.

Exercise 7.3.54

Let $n \in \mathbb{N}$. Prove that n is divisible by 5 if and only if the final digit in the decimal expansion of n is 5 or 0.

More generally, fix $k \geqslant 1$ and let m be the number whose decimal expansion is given by the last k digits of that of n. Prove that n is divisible by 5^k if and only if m is divisible by 5^k. For example, we have

$$125 \mid 9\,550\,828\,230\,495\,875 \quad \Leftrightarrow \quad 125 \mid 875$$

✎ **Exercise 7.3.55**

Let $n \in \mathbb{N}$. Prove that n is divisible by 11 if and only if the *alternating sum* of the digits of n is divisible by 11. That is, prove that if the decimal expansion of n is $d_r d_{r-2} \cdots d_0$, then

$$11 \mid n \quad \Leftrightarrow \quad 11 \mid d_0 - d_1 + d_2 - \cdots + (-1)^r d_r$$

✎ **Exercise 7.3.56**

Let $n \in \mathbb{N}$. Find a method for testing if n is divisible by 7 based on the decimal expansion of n.

Application: public-key cryptography

Public-key cryptography is a method of encryption and decryption that works according to the following principles:

- Encryption is done using a *public key* that is available to anyone.
- Decryption is done using a *private key* that is only known to the recipient.
- Knowledge of the private key should be extremely difficult to derive from knowledge of the public key.

Specifically, suppose that Alice wants to securely send Bob a message. As the recipient of the message, Bob has a public key and a private key. So:

- Bob sends the *public key* to Alice.
- Alice uses the public key to encrypt the message.
- Alice sends the encrypted message, which is visible (but encrypted) to anyone who intercepts it.
- Bob keeps the private key secret, and uses it upon receipt of the message to decrypt the message.

Notice that, since the public key can only be used to *encrypt* messages, a hacker has no useful information upon intercepting the message or the public key.

RSA encryption is an algorithm which provides one means of doing public-key cryptography using the theory of modular arithmetic. It works as follows.

Step 1. Let p and q be distinct positive prime numbers, and let $n = pq$. Then $\varphi(n) = (p-1)(q-1)$.

Step 2. Choose $e \in \mathbb{Z}$ such that $1 < e < \varphi(n)$ and $e \perp \varphi(n)$. The pair (n, e) is called the **public key**.

Step 3. Choose $d \in \mathbb{Z}$ such that $de \equiv 1 \pmod{\varphi(n)}$. The pair (n,d) is called the **private key**.

Step 4. To encrypt a message M (which is encoded as an integer), compute $K \in [n]$ such that $K \equiv M^e \pmod{n}$. Then K is the encrypted message.

Step 5. The original message M can be recovered since $M \equiv K^d \pmod{n}$.

Computing the private key (n,d) from the knowledge of (n,e) would allow a hacker to decrypt an encrypted message. However, doing so is typically very difficult when the prime factors of n are large. So if we choose p and q to be very large primes—which we can do without much hassle at all—then it becomes computationally infeasible for a hacker to compute the private key.

Example. Suppose I want to encrypt the message M, which I have encoded as the integer 32. Let $p = 13$ and $q = 17$. Then $n = 221$ and $\varphi(n) = 192$. Let $e = 7$, and note that $7 \perp 192$. Now $7 \times 55 \equiv 1 \pmod{192}$, so we can define $d = 55$.

- The public key is $(221,7)$, which Bob sends to Alice. Now Alice can encrypt the message:

$$32^7 \equiv 59 \pmod{221}$$

  Alice then sends Bob the encrypted message 59.

- The private key is $(221,55)$, so Bob can decrypt the message:

$$59^{55} \equiv 32 \pmod{221}$$

  so Bob has received Alice's message 32.

Exercise 7.3.57

Prove that the RSA algorithm is correct. Specifically, prove:

(a) If $n = pq$, for distinct positive primes p and q, then $\varphi(n) = (p-1)(q-1)$;

(b) Given $1 < e < \varphi(n)$ with $e \perp \varphi(n)$, there exists $d \in \mathbb{Z}$ with $de \equiv 1 \pmod{\varphi(n)}$.

(c) Given $M, K \in \mathbb{Z}$ with $K \equiv M^e \pmod{n}$, it is indeed the case that $K^d \equiv M \pmod{n}$.

Application: Euler's totient function

We now derive a formula for computing the totient of an arbitrary integer using the tools from Section 8.1—in particular, if you chose to read this section *before* learning about the multiplication principle, you should skip over this material.

Theorem 7.3.58 · *Multiplicativity of Euler's totient function*

Let $m,n \in \mathbb{Z}$ and let $\varphi : \mathbb{Z} \to \mathbb{N}$ be Euler's totient function (see Definition 7.3.26). If m and n are coprime, then $\varphi(mn) = \varphi(m)\varphi(n)$.

Proof
Since $\varphi(-k) = \varphi(k)$ for all $k \in \mathbb{Z}$, we may assume that $m \geqslant 0$ and $n \geqslant 0$. Moreover, if $m = 0$ or $n = 0$, then $\varphi(m)\varphi(n) = 0$ and $\varphi(mn) = 0$, so the result is immediate. Hence we may assume that $m > 0$ and $n > 0$.

Given $k \in \mathbb{Z}$, define

$$C_k = \{a \in [k] \mid a \perp k\}$$

By definition of Euler's totient function, we thus have $|C_k| = \varphi(k)$ for all $k \in \mathbb{Z}$. We will define a bijection

$$f : C_m \times C_n \to C_{mn}$$

using the Chinese remainder theorem (Theorem 7.3.46).

Given $a \in C_m$ and $b \in C_n$, let $f(a,b)$ be the element $x \in [mn]$ such that

$$\begin{cases} x \equiv a \pmod{m} \\ x \equiv b \pmod{n} \end{cases}$$

- f **is well-defined.** We check the properties of totality, existence and uniqueness.
  - ◇ **Totality.** We have accounted for all the elements of $C_m \times C_n$ in our specification of f.
  - ◇ **Existence.** By the Chinese remainder theorem, there exists $x \in \mathbb{Z}$ such that $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$. By adding an appropriate integer multiple of mn to x, we may additionally require $x \in [mn]$. It remains to check that $x \perp mn$.

    So let $d = \gcd(x, mn)$. If $d > 1$, then there is a positive prime p such that $p \mid x$ and $p \mid mn$. But then $p \mid m$ or $p \mid n$, meaning that either $p \mid \gcd(x,m)$ or $p \mid \gcd(x,n)$. But $x \equiv a \pmod{m}$, so $\gcd(x,m) = \gcd(a,m)$; and likewise $\gcd(x,n) = \gcd(b,n)$. So this contradicts the assumption that $a \perp m$ and $b \perp n$. Hence $x \perp mn$ after all.
  - ◇ **Uniqueness.** Suppose $x, y \in C_{mn}$ both satisfy the two congruences in question. By the Chinese remainder theorem, we have $x \equiv y \pmod{mn}$, and hence $x = y + kmn$ for some $k \in \mathbb{Z}$. Since $x, y \in [mn]$, we have

    $$|k|mn = |kmn| = |x - y| \leqslant mn - 1 < mn$$

    This implies $|k| < 1$, so that $k = 0$ and $x = y$.

  so f is well-defined.
- f **is injective.** Let $a, a' \in C_m$ and $b, b' \in C_n$, and suppose that $f(a,b) = f(a',b')$. Then there is an element $x \in C_{mn}$ such that

  $$\begin{cases} x \equiv a \pmod{m} \\ x \equiv a' \pmod{m} \\ x \equiv b \pmod{n} \\ x \equiv b' \pmod{n} \end{cases}$$

  Hence $a \equiv a' \pmod{m}$ and $b \equiv b' \pmod{n}$. Since $a, a' \in [m]$ and $b, b' \in [n]$, we must have $a = a'$ and $b = b'$.
- f **is surjective.** Let $x \in C_{mn}$. Let $a \in [m]$ and $b \in [n]$ be the (unique) elements such that $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$, respectively. If $a \in C_m$ and $b \in C_n$, then we'll have $f(a,b) = x$ by construction, so it remains to check that $a \perp m$ and $b \perp n$.

Suppose $d \in \mathbb{Z}$ with $d \mid a$ and $d \mid m$. We prove that $d = 1$. Since $x \equiv a \pmod{m}$, we have $d \mid x$ by Theorem 7.1.19. Since $m \mid mn$, we have $d \mid mn$. By definition of greatest common divisors, it follows that $d \mid \gcd(x, mn)$. But $\gcd(x, mn) = 1$, so that d is a unit, and so $a \perp m$ as required.

The proof that $b \perp n$ is similar.

It was a lot of work to check that it worked, but we have defined a bijection $f : C_m \times C_n \to C_{mn}$. By the multiplication principle, we have

$$\varphi(m)\varphi(n) = |C_m| \cdot |C_n| = |C_m \times C_n| = |C_{mn}| = \varphi(mn)$$

as required. □

It turns out that Theorem 7.3.58 and Exercise 7.3.28 are precisely the ingredients we need to find a general formula for the totient of a nonzero integer.

Theorem 7.3.59 · *Formula for Euler's totient function*

Let n be a nonzero integer. Then

$$\varphi(n) = |n| \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

where the product is indexed over distinct positive primes p dividing n.

Proof

Since $\varphi(n) = \varphi(-n)$ for all $n \in \mathbb{Z}$, we may assume that $n > 0$. Moreover

$$\varphi(1) = 1 = 1 \cdot \prod_{p|1} \left(1 - \frac{1}{p}\right)$$

Note that the product here is empty, and hence equal to 1, since there are no positive primes p which divide 1. So now suppose $n > 1$.

Using the fundamental theorem of arithmetic (Theorem 7.2.12), we can write

$$n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

for primes $0 < p_1 < p_2 < \cdots < p_r$ and natural numbers $k_1, k_2, \ldots, k_r \geqslant 1$.

By repeated application of Theorem 7.3.58, we have

$$\varphi(n) = \prod_{i=1}^{r} \varphi(p_i^{k_i})$$

for all $1 \leqslant i \leqslant r$. By Exercise 7.3.28, we have

$$\varphi(p_i^{k_i}) = p_i^{k_i} - p_i^{k_i - 1} = p_i^{k_i}\left(1 - \frac{1}{p_i}\right)$$

Combining these two results, it follows that

$$\varphi(n) = \prod_{i=1}^{r} p_i^{k_i}\left(1 - \frac{1}{p_i}\right) = \left(\prod_{i=1}^{r} p_i^{k_i}\right)\left(\prod_{i=1}^{r}\left(1 - \frac{1}{p_i}\right)\right) = n \cdot \prod_{i=1}^{r}\left(1 - \frac{1}{p_i}\right)$$

which is as required. □

Section 7.E

Chapter 7 exercises

Greatest common divisors

In Questions 7.E.1 to 7.E.3, use the Euclidean algorithm (Strategy 7.1.21) to find the greatest common divisor of the given pair of integers.

7.E.1. 382 and 218

7.E.2. 368475 and 26010

7.E.3. 24004512 and 10668672

7.E.4. Let $a,b,c,d \in \mathbb{Z}$ and suppose that $ad-bc=1$. Prove that $a+b$ and $c+d$ are coprime.

Definition 7.E.5
Let $f(x)$ and $g(x)$ be polynomials over $\mathbb{Z}$. We say $f(x)$ **divides** $g(x)$ if there exists a polynomial $q(x)$ over $\mathbb{Z}$ such that $g(x)=q(x)f(x)$.

7.E.6. Prove that $1-x$ divides $1-x^2$ and that $1-x$ does not divide $1+x^2$.

7.E.7. Prove that $1+x+x^2$ divides $1+x^4+x^5$.

Definition 7.E.8
Let $f(x)$ be a polynomial over $\mathbb{Z}$. The **degree** of f, written $\deg(f(x))$, is the greatest power of x in f whose coefficient is nonzero, unless f is the zero polynomial, whose degree is defined to be $-\infty$. We adopt the convention that $(-\infty)+r=-\infty=r+(-\infty)$ for all $r \in \mathbb{N}\cup\{-\infty\}$.

7.E.9. Let $f(x)$ and $g(x)$ be polynomials over $\mathbb{Z}$. Prove that $\deg(f(x),g(x)) = \deg(f(x)) + \deg(g(x))$.

7.E.10. Prove the following modified form of the division theorem (Theorem 7.1.1) for polynomials: given polynomials $f(x)$ and $g(x)$ over $\mathbb{Z}$, with $f(x)\neq 0$, prove that there exist unique polynomials $q(x)$ and $r(x)$ over $\mathbb{Z}$ such that $g(x)=q(x)f(x)+r(x)$ and $\deg(r(x)) < \deg(f(x))$.

Prime numbers

7.E.11. Prove that, for all $k \in \mathbb{N}$, the function $i : \mathbb{N}^k \to \mathbb{N}$ defined by

$$i(n_1,n_2,\ldots,n_k) = p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k}$$

for all $(n_1,n_2,\ldots,n_k) \in \mathbb{N}^k$ is an injection, where $p_1,p_2,p_3\ldots$ is an enumeration of the set of positive primes in increasing order. (Thus $p_1=2$, $p_2=3$, $p_3=5$, and so on.)

7.E.12. Use the result of Question 7.E.11 to construct a bijection

$$\bigcup_{k\in\mathbb{N}} \Big(\mathbb{N}^k \setminus \{(0,0,\ldots,0)\}\Big) \to \{n \in \mathbb{N} \mid n \geqslant 2\}$$

7.E.13. Use a method akin to that of Question 7.E.11 to define an injection $\mathbb{Z}^k \to \mathbb{N}$.

7.E.14. Define a subset $A \subseteq \mathbb{Z}$ by

$$A = \{n \in \mathbb{N} \mid \exists k > 0,\ n \mid 12^k - 1\}$$

Find all prime numbers in $\mathbb{Z} \setminus A$.

Base-b expansions

7.E.15. Let $n \in \mathbb{N}$. Prove that the number of trailing 0s in the decimal expansion of $n!$ is equal to

$$\sum_{k=1}^{d} \left\lfloor \frac{n}{5^k} \right\rfloor$$

where $d \in \mathbb{N}$ is least such that $5^{d+1} > n$, and where $\lfloor x \rfloor$ (LaTeX code: `\lfloor,\rfloor`) denotes the floor of x (see Definition 0.10).

7.E.16. Let $b \in \mathbb{N}$ with $b \geqslant 2$. Find an expression in terms of $n \in \mathbb{N}$ for the number of trailing 0s in the base-b expansion of $n!$.

True–False questions

In Questions 7.E.17 to 7.E.22, determine (with proof) whether the statement is true or false.

7.E.17. There is an integer that is not coprime to any integer.

7.E.18. Every linear Diophantine equation has a solution.

7.E.19. Every integer n is coprime to its successor $n+1$.

7.E.20. If the greatest common divisors of two integers is 3, and their least common multiple of 7, then their product is 21.

7.E.21. Every integer is congruent modulo 7 to its remainder when divided by 21.

7.E.22. Let $k \geqslant 1$. Then $15^k \equiv 1 \pmod 6$.

Always–Sometimes–Never questions

In Questions 7.E.23 to 7.E.31, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

7.E.23. Let $a, b, c \in \mathbb{Z}$ and suppose that $\gcd(a,b) = 2$ and $\gcd(b,c) = 2$. Then $\gcd(a,c) = 2$.

7.E.24. Let $a, b \in \mathbb{Z}$. Then $\gcd(a,b) = \gcd(a, \gcd(a,b))$.

7.E.25. Let $a, b \geqslant 2$. Then $\gcd(a,b) = \operatorname{lcm}(a,b)$.

7.E.26. Let $n \in \mathbb{Z}$. Then $n^2 + 1$ is coprime to $n^4 + n^2 + 1$.

7.E.27. Let $p \in \mathbb{Z}$ be prime. Then p is coprime to all integers $a \neq p$.

7.E.28. Let $p, q, r, s \in \mathbb{Z}$ be positive primes and suppose that $pq = rs$. Then $p = r$ and $q = s$.

7.E.29. Let $p_1, p_2, \ldots, p_s, q_1, q_2, \ldots, q_t \in \mathbb{Z}$ be prime and suppose that $p_1 p_2 \ldots p_s = q_1 q_2 \ldots q_t$. Then $s = t$.

7.E.30. Let $a, b \in \mathbb{Z}$ and let p be a positive prime. Then $p \mid (a+b)^p - a - b$.

7.E.31. Let $n \geqslant 0$ and let $b > 0$. Then n is congruent modulo $b-1$ to the sum of the digits in its base-b expansion.

Chapter 8

Enumerative combinatorics

Section 8.1

Counting principles

In Section 6.1 we were interested in establishing conditions under which a set is finite, and proving that we may perform certain operations on finite sets—such as unions and cartesian products—without losing the property of finiteness.

In this section, our attention turns to the task of finding the size of a set that is known to be finite. This process is called *counting* and is at the core of the mathematical field of combinatorics.

Binomials and factorials revisited

We defined binomial coefficients $\binom{n}{k}$ and factorials $n!$ *recursively* in Chapter 4, and proved elementary facts about them by induction. We will now re-define them *combinatorially*—that is, we give them meaning in terms of sizes of particular finite sets. We will prove that the combinatorial and recursive definitions are equivalent, and prove facts about them using combinatorial arguments.

The reasons for doing so are manifold. The combinatorial definitions allow us to reason about binomials and factorials with direct reference to descriptions of finite sets, which will be particularly useful when we prove identities about them using *double counting*. Moreover, the combinatorial definitions remove the seemingly arbitrary nature of the recursive definitions—for example, they provide a reason why it makes sense to define $0! = 1$ and $\binom{0}{0} = 1$.

Definition 8.1.1

Let X be a set and let $k \in \mathbb{N}$. A k-**element subset** of X is a subset $U \subseteq X$ such that $|U| = k$. The set of all k-element subsets of X is denoted $\binom{X}{k}$ (read: 'X choose k') (LaTeX code: `\binom{X}{k}`).

Intuitively, $\binom{X}{k}$ is the set of ways of picking k elements from X, without repetitions, in such a way that order doesn't matter. (If order mattered, the elements would be *sequences* instead of *subsets*.)

Example 8.1.2

We find $\binom{[4]}{k}$ for all $k \in \mathbb{N}$.

- $\binom{[4]}{0} = \{\varnothing\}$ since the only set with 0 elements is $\varnothing$;
- $\binom{[4]}{1} = \{\{0\},\{1\},\{2\},\{3\}\}$;
- $\binom{[4]}{2} = \{\{0,1\},\{0,2\},\{0,3\},\{1,2\},\{1,3\},\{2,3\}\}$;
- $\binom{[4]}{3} = \{\{0,1,2\},\{0,1,3\},\{0,2,3\},\{1,2,3\}\}$;
- $\binom{[4]}{4} = \{\{0,1,2,3\}\}$;
- If $k \geqslant 5$ then $\binom{[4]}{k} = \varnothing$, since by Exercise 6.1.16, no subset of $[4]$ can have more than 4 elements.

Proposition 8.1.3
If X is a finite set, then $\mathscr{P}(X) = \bigcup_{k \leqslant |X|} \binom{X}{k}$.

Proof
Let $U \subseteq X$. By Exercise 6.1.16, U is finite and $|U| \leqslant |X|$. Thus $U \in \binom{X}{|U|}$, and hence $U \in \bigcup_{k \leqslant |X|} \binom{X}{k}$. This proves that $\mathscr{P}(X) \subseteq \bigcup_{k \leqslant |X|} \binom{X}{k}$.

The fact that $\bigcup_{k \leqslant |X|} \binom{X}{k} \subseteq \mathscr{P}(X)$ is immediate, since elements of $\binom{X}{k}$ are defined to be subsets of X, and hence elements of $\mathscr{P}(X)$. □

Definition 8.1.4
Let $n, k \in \mathbb{N}$. Denote by $\binom{n}{k}$ (read: 'n choose k') (LaTeX code: `\binom{n}{k}`) the number of k-element subsets of $[n]$. That is, we define $\binom{n}{k} = \left|\binom{[n]}{k}\right|$. The numbers $\binom{n}{k}$ are called **binomial coefficients**.

Some authors use the notation ${}_nC_k$ or ${}^nC_k$ instead of $\binom{n}{k}$. We avoid this, as it is unnecessarily clunky.

Intuitively, $\binom{n}{k}$ is the number of ways of selecting k things from n, without repetitions, in such a way that order doesn't matter.

The value behind this notation is that it allows us to express huge numbers in a concise and meaningful way. For example,

$$\binom{4000}{11} = 103\,640\,000\,280\,154\,258\,645\,590\,429\,564\,000$$

Although these two numbers are equal, their *expressions* are very different; the expression on the left is meaningful, but the expression on the right is completely meaningless out of context.

Writing tip
When expressing the sizes of finite sets described combinatorially, it is best to *avoid* evaluating the expression; that is, leave in the powers, products, sums, binomial coefficients and factorials! The reason for this is that performing the calculations takes the meaning away from the expressions.

Example 8.1.5
In Example 8.1.2 we proved that:

$$\binom{4}{0} = 1,\ \binom{4}{1} = 4,\ \binom{4}{2} = 6,\ \binom{4}{3} = 4,\ \binom{4}{4} = 1$$

and that $\binom{4}{k} = 0$ for all $k \geqslant 5$.

Exercise 8.1.6
Fix $n \in \mathbb{N}$. Prove that $\binom{n}{0} = 1$, $\binom{n}{1} = n$ and $\binom{n}{n} = 1$.

Definition 8.1.7
Let X be a set. A **permutation** of X is a bijection $X \to X$.

Example 8.1.8
There are six permutations of the set $[3]$. Representing each bijection $f : [3] \to [3]$ by the ordered triple $(f(0), f(1), f(2))$ of its values, these permutations are thus given by

$$(0,1,2),\ (0,2,1),\ (1,0,2),\ (1,2,0),\ (2,0,1),\ (2,1,0)$$

For example, $(1,2,0)$ represents the permutation $f : [3] \to [3]$ defined by $f(0) = 1$, $f(1) = 2$ and $f(2) = 0$.

Exercise 8.1.9
List all the permutations of the set $[4]$.

Definition 8.1.10
Let $n \in \mathbb{N}$. Denote by $n!$ (read: 'n factorial') the number of permutations of $[n]$. That is,

$$n! = |\{f : [n] \to [n] \mid f \text{ is a bijection}\}|$$

The numbers $n!$ are called **factorials**.

Example 8.1.11
Example 8.1.8 shows that $3! = 6$.

The result of the following exercise allows us to interpret factorials in different ways than just permutations of sets.

Exercise 8.1.12 · *Other characterisations of the factorial*
Let $n \in \mathbb{N}$, and let X and Y be finite sets with $|X| = |Y| = n$.

Prove that the following sets all have size $n!$:

(a) The set of all n-tuples $(x_0, x_1, \ldots, x_{n-1}) \in X^n$ such that each element of X appears exactly once in the list;

(b) The set of all permutations of X;

(c) The set of all enumerations of X (that is, bijections $[n] \to X$); and

(d) The set of all bijections $X \to Y$.

Products and procedures

We saw in Proposition 6.1.22 that, given two finite sets X and Y, the product $X \times Y$ is finite. We also found a formula for its size, namely $|X \times Y| = |X| \cdot |Y|$. The *multiplication principle* (Theorem 8.1.22) generalises this formula to products that may contain any finite number of sets, not just two.

Lemma 8.1.13
Let $\{X_1, \ldots, X_n\}$ be a family of finite sets, with $n \geqslant 1$. Then $\prod_{i=1}^{n} X_i$ is finite, and

$$\left|\prod_{i=1}^{n} X_i\right| = |X_1| \cdot |X_2| \cdot \cdots \cdot |X_n|$$

Proof
We proceed by induction on n.

- (**Base case**) When $n = 1$, an element of $\prod_{i=1}^{1} X_i$ is 'officially' a sequence (x_1) with $x_1 \in X_1$. This is the same as an element of X_1, in the sense that the assignments $(x_1) \mapsto x_1$ and $x_1 \mapsto (x_1)$ define mutually inverse (hence bijective) functions between $\prod_{i=1}^{1} X_i$ and X_1, and so

$$\left|\prod_{i=1}^{1} X_i\right| = |X_1|$$

- (**Induction step**) Fix $n \in \mathbb{N}$, and suppose that

$$\left|\prod_{i=1}^{n} X_i\right| = |X_1| \cdot |X_2| \cdot \cdots \cdot |X_n|$$

for all sets X_i for $i \in [n]$. This is our induction hypothesis.
Now let $X_1, \ldots, X_n, X_{n+1}$ be sets. We define a function

$$F : \prod_{i=1}^{n+1} X_i \to \left(\prod_{i=1}^{n} X_i\right) \times X_{n+1}$$

by letting $F((x_1, \ldots, x_n, x_{n+1})) = ((x_1, \ldots, x_n), x_{n+1})$. It is again easy to check that F is a bijection, and hence

$$\left|\prod_{i=1}^{n+1} X_i\right| = \left|\prod_{i=1}^{n} X_i\right| \cdot |X_{n+1}|$$

by Proposition 6.1.22. Applying the induction hypothesis, we obtain the desired result, namely

$$\left|\prod_{i=1}^{n+1} X_i\right| = |X_1| \cdot |X_2| \cdot \cdots \cdot |X_n| \cdot |X_{n+1}|$$

By induction, we're done. □

Lemma 8.1.13 gives rise to a useful strategy for computing the size of a finite set X—see Theorem 8.1.16. Intuitively, by devising a step-by-step procedure for specifying an element of X, we are constructing a cartesian product $\prod_{k=1}^{n} X_k$, where X_k is the set of choices to be made in the k^{th} step. This establishes a bijection $\prod_{k=1}^{n} X_k \to X$, which by bijective proof (Strategy 6.1.15(c)) lets us compute $|X|$ as the product of the numbers of choices that can be made in each step.

Definition 8.1.14

A **specification procedure** for a set X is a sequence of r steps, for some $r \in \mathbb{N}$, such that:

(i) In each step, one of a predetermined set of choices is made—which choices are available may in general depend on the choices made in the previous steps;

(ii) Each sequence of choices determines an element of X; and

(iii) Each element of X is uniquely determined by a unique sequence of choices.

Example 8.1.15

Let $n \in \mathbb{N}$. A specification procedure for $\mathscr{P}([n])$ is as follows. There are n steps, numbered Step 0 through Step $n-1$. To specify an element $A \in \mathscr{P}([n])$ (that is, a subset $A \subseteq [n]$), for each $k \in [n]$, we decide in Step k whether $k \in A$ or $k \notin A$.

In the context of Definition 8.1.14, letting S be the set of all sequences of choices, condition (ii) is equivalent to the existence of a function $S \to X$, which sends a sequence of choices to the element of X determined by those choices, and condition (iii) is equivalent to this function being a bijection. So if S is finite, then so is X, and $|X| = |S|$.

If, for each $k \in [r]$, the available choices in Step k are independent of those made in other steps, this means that $S = \prod_{k \in [r]} S_k$, where S_k is the set of choices that are available in Step k.

Therefore we can apply Lemma 8.1.13 to obtain the following theorem:

Theorem 8.1.16 · *Multiplication principle (independent version)*

Let X be a set, and suppose that there is a specification procedure for X such that in each step, there are finitely many choices available, and the choices are independent of those made in other steps. Then X is finite.

Moreover, letting $r \in \mathbb{N}$ be the number of steps, and letting m_k for $k \in [r]$ be the numbers of choices available in the respective steps, we have $|X| = \prod_{k \in [r]} m_k$.

Example 8.1.17

You go to an ice cream stand selling the following flavours:

vanilla, strawberry, chocolate, rum and raisin, mint choc chip, toffee crunch

You can have your ice cream in a tub, a regular cone or a *choco-cone*. You can have one, two or three scoops. We will compute how many options you have.

To select an ice cream, you follow the following procedure:

- **Step 1.** Choose a flavour. There are 6 ways to do this.
- **Step 2.** Choose whether you'd like it in a tub, regular cone or choco-cone. There are 3 ways to do this.
- **Step 3.** Choose how many scoops you'd like. There are 3 ways to do this.

Hence there are $6 \times 3 \times 3 = 54$ options in total.

This may feel informal, but really what we are doing is establishing a bijection. Letting X be the set of options, the above procedure defines a bijection

$$F \times C \times S \to X$$

where F is the set of flavours, $C = \{\text{tub}, \text{regular cone}, \text{choco-cone}\}$ and $S = [3]$ is the set of possible numbers of scoops.

Example 8.1.18

We will prove that $|\mathscr{P}(X)| = 2^{|X|}$ for all finite sets X.

Let X be a finite set and let $n = |X|$. Write

$$X = \{x_k \mid k \in [n]\} = \{x_0, x_1, \ldots, x_{n-1}\}$$

Intuitively, specifying an element of $\mathscr{P}(X)$—that is, a subset $U \subseteq X$—is equivalent to deciding, for each $k \in [n]$, whether $x_k \in U$ or $x_k \notin U$ ('in or out'), which in turn is equivalent to specifying an element of $\{\text{in}, \text{out}\}^n$.

For example, taking $X = [7]$, the subset $U = \{1, 2, 6\}$ corresponds with the choices

$$0 \text{ out},\ 1 \text{ in},\ 2 \text{ in},\ 3 \text{ out},\ 4 \text{ out},\ 5 \text{ out},\ 6 \text{ in}$$

and hence the element $(\text{out}, \text{in}, \text{in}, \text{out}, \text{out}, \text{out}, \text{in}) \in \{\text{in}, \text{out}\}^7$.

This defines a function $i : \mathscr{P}(X) \to \{\text{in}, \text{out}\}^n$. This function is injective, since different subsets determine different sequences; and it is surjective, since each sequence determines a subset.

The above argument is sufficient for most purposes, but since this is the first bijective proof we have come across, we now give a more formal presentation of the details.

Define a function

$$i : \mathscr{P}(X) \to \{\text{in}, \text{out}\}^n$$

by letting the k^{th} component of $i(U)$ be 'in' if $x_k \in U$ or 'out' if $x_k \notin U$, for each $k \in [n]$.

We prove that i is a bijection.

- i **is injective.** To see this, take $U, V \subseteq X$ and suppose $i(U) = i(V)$. We prove that $U = V$. So fix $x \in X$ and let $k \in [n]$ be such that $x = x_k$. Then

$$\begin{aligned} x \in U &\Leftrightarrow \text{the } k^{\text{th}} \text{ component of } i(U) \text{ is 'in'} && \text{by definition of } i \\ &\Leftrightarrow \text{the } k^{\text{th}} \text{ component of } i(V) \text{ is 'in'} && \text{since } i(U) = i(V) \\ &\Leftrightarrow x \in V && \text{by definition of } i \end{aligned}$$

  so indeed we have $U = V$, as required.

- i **is surjective.** To see this, let $v \in \{\text{in}, \text{out}\}^n$, and let

$$U = \{x_k \mid \text{the } k^{\text{th}} \text{ component of } v \text{ is 'in'}\}$$

  Then $i(U) = v$, since for each $k \in [n]$ we have $x_k \in U$ if and only if the k^{th} component of v is 'in', which is precisely the definition of $i(U)$.

Hence

$$|\mathscr{P}(X)| = |\{\text{in}, \text{out}\}|^n = 2^n$$

as required.

Some authors actually write 2^X to refer to the power set of a set X. This is justified by Example 8.1.18.

Exercise 8.1.19

Let X and Y be finite sets, and recall that Y^X denotes the set of functions from X to Y. Prove that $|Y^X| = |Y|^{|X|}$.

Exercise 8.1.20

Since September 2001, car number plates on the island of Great Britain have taken the form `XX NN XXX`, where each `X` can be any letter of the alphabet except for 'I', 'Q' or 'Z', and `NN` is the last two digits of the year. [This is a slight simplification of what is really the case, but let's not concern ourselves with *too* many details!] How many possible number plates are there? Number plates of vehicles registered in the region of Yorkshire begin with the letter 'Y'. How many Yorkshire number plates can be issued in a given year?

The multiplication principle in the form of Theorem 8.1.16 does not allow for steps later in a procedure to depend on those earlier in the procedure. To see why this is a problem, suppose we want to count the size of the set $X = \{(a, b) \in [n] \times [n] \mid a \neq b\}$. A step-by-step procedure for specifying such an element is as follows:

- **Step 1.** Select an element $a \in [n]$. There are n choices.
- **Step 2.** Select an element $b \in [n]$ with $b \neq a$. There are $n - 1$ choices.

We would like to use Theorem 8.1.16 to deduce that $|X| = n(n-1)$. Unfortunately, this is not valid because the possible choices available to us in Step 2 depend on the choice made in Step 1. Elements of cartesian products do not depend on one another, and so the set of sequences of choices made cannot necessarily be expressed as a cartesian product of two sets. Thus we cannot apply Lemma 8.1.13. Oh no!

However, provided that the *number* of choices in each step remains constant, in spite of the choices themselves changing, it turns out that we can still compute the size of the set in question by multiplying together the numbers of choices.

This is what we prove next. We begin with a pairwise version (analogous to Exercise 6.1.21) and then prove the general version by induction (like in Lemma 8.1.13).

Lemma 8.1.21

Fix $m, n \in \mathbb{N}$. Let X be a finite set with $|X| = m$, and for each $a \in X$, let Y_a be a finite set with $|Y_a| = n$. Then the set

$$Z = \{(a, b) \mid a \in X \text{ and } b \in Y_a\}$$

is finite and $|Z| = mn$.

Proof

Fix bijections $f : [m] \to X$ and $g_a : [n] \to Y_a$ for each $a \in X$. Define $h : [m] \times [n] \to Z$ by letting $h(i, j) = (f(i), g_{f(i)}(j))$ for each $(i, j) \in [m] \times [n]$. Then:

- h is well-defined, since for all $i \in [m]$ and $j \in [n]$ we have $f(i) \in X$ and $g_{f(i)}(j) \in Y_{f(i)}$.
- h is injective. To see this, fix $(i, j), (k, \ell) \in [m] \times [n]$ and assume that $h(i, j) = h(k, \ell)$. Then $(f(i), g_{f(i)}(j)) = (f(k), g_{f(k)}(\ell))$, so that $f(i) = f(k)$ and $g_{f(i)}(j) = g_{f(k)}(\ell)$. Since f is injective, we have $i = k$—therefore $g_{f(i)}(j) = g_{f(i)}(\ell)$, and then since $g_{f(i)}$ is injective, we have $j = \ell$. Thus $(i, j) = (k, \ell)$, as required.
- h is surjective. To see this, let $(a, b) \in Z$. Since f is surjective and $a \in X$, we have $a = f(i)$ for some $i \in [m]$. Since g_a is surjective and $b \in Y_a$, we have $b = g_a(j)$ for some $j \in [n]$. But then

  $$(a, b) = (f(i), g_a(j)) = (f(i), g_{f(i)}(j)) = h(i, j)$$

  so that h is surjective.

Since h is a bijection, we have $|Z| = |[m] \times [n]|$ by Theorem 6.1.13(iii), which is equal to mn by Proposition 6.1.22. □

We are now ready to state and prove the theorem giving rise to the multiplication principle in its full generality.

Theorem 8.1.22 · *Multiplication principle (general version)*

Let X be a set, and suppose that there is a specification procedure for X such that in each step, there are finitely many choices available, and the number of available choices is independent of those made in other steps. Then X is finite.

Moreover, letting $r \in \mathbb{N}$ be the number of steps, and letting m_k for $k \in [r]$ be the numbers of choices available in the respective steps, we have $|X| = \prod_{k \in [r]} m_k$.

Proof
Given an r-step specification procedure for X, define a family of sets as follows:

- Let $X^{(0)}$ be the set of all choices available in Step 0, and let $m_0 = |X^{(0)}| \in \mathbb{N}$.
- For each $k \in \mathbb{N}$ with $0 < k < r$, and each sequence s of choices available in Steps 0 through $k-1$: let $X_s^{(k)}$ be the set of all choices available in Step k when the choices made in the previous steps were those in s; and let $m_k = |X_s^{(k)}| \in \mathbb{N}$. Note that the value of m_k is independent of s by the assumption in the theorem.

Finally, let P be the set of all sequences of steps. Note that there is a bijection $P \to X$ by definition of specification procedure, so it suffices to prove that P is finite and that $|P| = \prod_{k\in[r]} m_k$.

We proceed by induction on $r \geqslant 1$.

- (**Base case**) When $r = 1$, there is a single step in the specification procedure, so the function $f : X^{(0)} \to P$ defined by $f(s) = (s)$ is a bijection, and so P is finite and

$$|P| = |X^{(0)}| = m_0 = \prod_{k\in[1]} m_k$$

  as required.

- (**Induction step**) Let $r \geqslant 1$ and suppose that the theorem statement holds for r-step specification procedures.

  Suppose now that we have an $(r+1)$-step specification procedure (with steps numbered 0 through r, say), and let the sets $X_s^{(k)}$ and natural numbers m_k (for suitable natural numbers k and sequences s of choices) and P be defined as at the beginning of the theorem.

  Let Q be the set of all sequences of choices made in Steps 0 through $r-1$, and define $Z = \{(s, c_r) \mid s \in Q \text{ and } c_r \in X_s^{(r)}\}$.

  There is an evident bijection $f : Z \to P$ given by

$$f((c_0, \ldots, c_{r-1}), c_r) = (c_0, c_1, \ldots, c_{r-1}, c_r)$$

  for all $((c_0, \ldots, c_{r-1}), c_r) \in Z$.

  But now $|Q| = \prod_{k\in[r]} m_k$ by the induction hypothesis, and $|X_s^{(r)}| = m_r$ for all $s \in Q$ by definition of m_r, so by Lemma 8.1.21 we have that Z is finite and

$$|Z| = |Q| \cdot |X_s^{(r)}| = \left(\prod_{k\in[r]} m_k\right) \cdot m_r = \prod_{k\in[r+1]} m_k$$

  Since f is a bijection, P is also finite and also has this size, as required. □

Example 8.1.23

We prove that there are six bijections $[3] \to [3]$. We can specify a bijection $f : [3] \to [3]$ according to the following procedure.

- **Step 1.** Choose the value of $f(0)$. There are 3 choices: 0, 1 or 2.
- **Step 2.** Choose the value of $f(1)$. The values $f(1)$ can take depend on the chosen value of $f(0)$.
  - ◇ If $f(0) = 0$, then $f(1)$ can be equal to 1 or 2.
  - ◇ If $f(0) = 1$, then $f(1)$ can be equal to 0 or 2.
  - ◇ If $f(0) = 2$, then $f(1)$ can be equal to 0 or 1.

  In each case, there are 2 choices for the value of $f(1)$.
- **Step 3.** Choose the value of $f(2)$. The values $f(2)$ can take depend on the values of $f(0)$ and $f(1)$. We could split into the (six!) cases based on the values of $f(0)$ and $f(1)$ chosen in Steps 1 and 2; but we won't. Instead, note that $\{f(0), f(1)\}$ has two elements, and by injectivity $f(2)$ must have a distinct value, so that $[3] \setminus \{f(0), f(1)\}$ has one element. This element must be the value of $f(2)$. Hence there is only possible choice of $f(2)$.

By the multiplication principle, there are $3 \times 2 \times 1 = 6$ bijections $[3] \to [3]$.

Exercise 8.1.24

Count the number of injections $[3] \to [4]$.

Example 8.1.25

We count the number of ways we can shuffle a standard deck of cards in such a way that the colour of the cards alternate between red and black.

A procedure for choosing the order of the cards is as follows:

(i) Choose the colour of the first card. There are 2 such choices. This then determines the colours of the remaining cards, since they have to alternate.

(ii) Choose the order of the red cards. There are 26! such choices.

(iii) Choose the order of the black cards. There are 26! such choices.

By the multiplication principle, there are $2 \cdot (26!)^2$ such rearrangements. This number is huge, and in general there is no reason to write it out. Just for fun, though:

$$325\,288\,005\,235\,264\,929\,014\,077\,766\,819\,257\,214\,042\,112\,000\,000\,000\,000$$

Sums and partitions

We saw in Proposition 6.1.18 that, given two finite sets X and Y, the union $X \cup Y$ is finite. We also found formulae for their size, namely $|X \cup Y| = |X| + |Y| - |X \cap Y|$. The *addition principle* (Theorem 8.1.26) generalises this formula to any finite number of sets, provided the sets have no

elements in common with one another—that is they are *pairwise disjoint*. [The hypothesis of pairwise disjointness is removed in the *inclusion–exclusion principle*, which is studied in Section 8.2.]

If you have not covered Section 5.2 yet, you are encouraged to take a brief detour to read from Definition 5.2.21 to Exercise 5.2.25; the definition of a *partition* of a set is recalled below.

Definition 5.2.21

A **partition** of a set X is a set $\mathscr{A} \subseteq \mathscr{P}(X)$ (that is, a set of subsets of X) such that the following conditions hold:

(a) Every set $A \in \mathscr{A}$ is inhabited;

(b) $\mathscr{A}$ is a **cover** of X: for all $x \in X$, there exists $A \in \mathscr{A}$ such that $x \in A$; and

(c) $\mathscr{A}$ is **pairwise disjoint**: for all $A, B \in \mathscr{A}$, if $A \cap B$ is inhabited, then $A = B$.

In this section, we will simplify matters in two ways:

- When we say 'partition' in this section (and Section 8.2), we will allow the sets in the partition to be empty—that is, we will just need conditions (b) and (c) of Definition 5.2.21 to hold.
- Since our sets are finite, so will $\mathscr{A}$ be; so we will only ever partition our sets into finitely many pieces. That is, all of our partitions will take form $\mathscr{A} = \{A_i \mid i \in [n]\}$ for some $n \in \mathbb{N}$.

With all of this said, let's get right to it.

Theorem 8.1.26 · *Addition principle*

Let X be a set and let $\mathscr{A} = \{A_i \mid i \in [n]\}$ be a partition of X for some $n \in \mathbb{N}$, such that each set A_i is finite. Then X is finite, and

$$|X| = \sum_{i \in [n]} |A_i|$$

Exercise 8.1.27

Prove Theorem 8.1.26. The proof follows the same pattern as that of Lemma 8.1.13. Be careful to make sure you identify where you use the hypothesis that the sets U_i are pairwise disjoint!

Example 8.1.28

We will count the number of inhabited subsets of $[7]$ which either contain only even numbers, or contain only odd numbers.

Let E denote the set of inhabited subsets of $[7]$ containing only even numbers, and let O denote the set of inhabited subsets of $[7]$ containing only odd numbers. Note that $\{E, O\}$ forms a partition of the set we are counting, and so our set has $|E| + |O|$ elements.

- An element of E must be a subset of $\{0,2,4,6\}$. By Example 8.1.18 there are $2^4 = 16$ such subsets. Thus the number of *inhabited* subsets of $[7]$ containing only odd numbers is 15, since we must exclude the empty set. That is, $|E| = 15$.

- An element of O must be a subset of $\{1,3,5\}$. Again by Example 8.1.18 there are $2^3 = 8$ such subsets. Hence there are 7 inhabited subsets of $[7]$ containing only even numbers. That is, $|O| = 7$.

Hence there are $15+7 = 22$ inhabited subsets of $[7]$ containing only even or only odd numbers. And here they are:

$$\begin{array}{ccccccccc} \{0\} & \{2\} & \{4\} & \{6\} & \{0,2\} & & \{1\} & \{3\} & \{5\} \\ \{0,4\} & \{0,6\} & \{2,4\} & \{2,6\} & \{4,6\} & & \{1,3\} & \{1,5\} & \{3,5\} \\ \{0,2,4\} & \{0,2,6\} & \{0,4,6\} & \{2,4,6\} & \{0,2,4,6\} & & \{1,3,5\} & & \end{array}$$

Exercise 8.1.29

Pick your favourite integer $n \geqslant 1000$. For this value of n, how many inhabited subsets of $[n]$ contain either only even or only odd numbers? (You need not evaluate exponents.)

We now consider some examples of finite sets which use both the multiplication principle and the addition principle.

Example 8.1.30

A city has $6n$ inhabitants. The favourite colour of n of the inhabitants is orange, the favourite colour of $2n$ of the inhabitants is pink, and the favourite colour of $3n$ of the inhabitants is turquoise. The city government wishes to form a committee with equal representation from the three colour preference groups to decide how the new city hall should be painted. We count the number of ways this can be done.

Let X be the set of possible committees. First note that

$$X = \bigcup_{k=0}^{n} A_k$$

where A_k is the set of committees with exactly k people from each colour preference group. Indeed, we must have $k \leqslant n$, since it is impossible to have a committee with more than n people from the orange preference group.

Moreover, if $k \neq \ell$ then $A_k \cap A_\ell = \varnothing$, since if $k \neq \ell$ then a committee cannot simultaneously have exactly k people and exactly ℓ people from each preference group.

By the addition principle, we have

$$|X| = \sum_{k=0}^{n} |A_k|$$

We count A_k for fixed k using the following procedure:

- **Step 1.** Choose k people from the orange preference group to be on the committee. There are $\binom{n}{k}$ choices.
- **Step 2.** Choose k people from the pink preference group to be on the committee. There are $\binom{2n}{k}$ choices.

- **Step 3.** Choose k people from the turquoise preference group to be on the committee. There are $\binom{3n}{k}$ choices.

By the multiplication principle, it follows that $|A_k| = \binom{n}{k}\binom{2n}{k}\binom{3n}{k}$. Hence

$$|X| = \sum_{k=0}^{n} \binom{n}{k}\binom{2n}{k}\binom{3n}{k}$$

Exercise 8.1.31
In Example 8.1.30, how many ways could a committee be formed with a *representative* number of people from each colour preference group? That is, the proportion of people on the committee which prefer any of the three colours should be equal to the corresponding proportion of the population of the city.

Pigeonhole principle

A nice application of the addition principle is to prove the *pigeonhole principle*, which is used heavily in combinatorics.

Informally, the pigeonhole principle says that if you assign pigeons to pigeonholes, and there are more pigeons than pigeonholes, then some pigeonhole must have more than one pigeon in it. We can (and do) generalise this slightly: it says that given $q \in \mathbb{N}$, if you have more than q times as many pigeons than pigeonholes, then some pigeonhole must have more than q pigeons in it.

The proof is deceptively simple.

Theorem 8.1.32 · *Pigeonhole principle*
Let $q \in \mathbb{N}$, and let X and Y be finite sets with $|X| = m \in \mathbb{N}$ and $|Y| = n \in \mathbb{N}$. Then:

(a) If $m > qn$, then for every function $f : X \to Y$, there is some $a \in Y$ such that $|f^{-1}[\{a\}]| > q$.

(b) If $m \leqslant qn$, then there is a function $f : X \to Y$ such that $|f^{-1}[\{a\}]| \leqslant q$ for all $a \in Y$.

Proof of (a)
Suppose $m > qn$. It follows from Exercise 5.2.24 that the sets $f^{-1}[\{a\}]$ partition X. Towards a contradiction, assume $|f^{-1}[\{a\}]| \leqslant q$ for all $a \in Y$. Then by the addition principle

$$m = |X| = \left|\bigcup_{a\in Y} f^{-1}[\{a\}]\right| = \sum_{a\in Y} |f^{-1}[\{a\}]| \leqslant \sum_{a\in Y} q = |Y| \cdot q = qn$$

This contradicts the assumption that $m > qn$. □

Exercise 8.1.33
Prove part (b) of Theorem 8.1.32.

Example 8.1.34

Let $n, k \in \mathbb{N}$. Assume that you have n pairs of socks in a drawer, and each sock is one of k colours. We wish to know how many socks you must take out of the drawer before you can guarantee that you have a matching pair.

Let C be set of colours of the socks, so that $|C| = k$, and let X be the set of socks that you have selected. We obtain a function $f : X \to C$ that assigns to each sock x its colour $f(x) \in C$. Given a colour $c \in C$, the preimage $f^{-1}[\{c\}]$ is the set of socks of colour c that we have selected.

Thus the question becomes: what size must X be in order to have $|f^{-1}[\{c\}]| \geqslant 2$ for some $c \in C$? [The English translation of this question is: how many socks must we have picked in order for two of the socks to have the same colour?]

Well, by the pigeonhole principle, we can guarantee $|f^{-1}[\{c\}]| \geqslant 2$ (or equivalently > 1) if and only if $|X| > |C|$. That is, we need to select at least $k+1$ socks to guarantee a matching pair.

Exercise 8.1.35

Six people are in a room. The atmosphere is tense, since each pair of people is either friends or enemies. There are no allegiances, so for example it is possible for a friend of a friend to be an enemy, or an enemy of a friend to be a friend, and so on. Prove that there is some set of three people that are either all each other's friends or all each other's enemies.

Double counting

Double counting (also known as *counting in two ways*) is a proof technique that allows us to prove that two natural numbers are equal by establishing they are two expressions for the size of the same set. (More generally, by Theorem 6.1.13(iii), we can relate them to the sizes of two sets which are in bijection.)

The proof of Proposition 8.1.36 illustrates this proof very nicely. We proved it already by induction in Example 4.2.13; the combinatorial proof we now provide is much shorter and cleaner.

Proposition 8.1.36

Let $n \in \mathbb{N}$. Then $2^n = \sum_{k=0}^{n} \binom{n}{k}$.

Proof

We know that $|\mathscr{P}([n])| = 2^n$ by Example 8.1.18 and that $\mathscr{P}([n]) = \bigcup_{k=0}^{n} \binom{[n]}{k}$ by Proposition 8.1.3.

Moreover, the sets $\binom{[n]}{k}$ are pairwise disjoint, so by the addition principle it follows that

$$2^n = |\mathscr{P}([n])| = \left|\bigcup_{k=0}^{n} \binom{[n]}{k}\right| = \sum_{k=0}^{n} \left|\binom{[n]}{k}\right| = \sum_{k=0}^{n} \binom{n}{k}$$

Strategy 8.1.37 · *Double counting*
In order to prove that two expressions involving natural numbers are equal, it suffices to define a set X and devise two counting arguments to show that $|X|$ is equal to both expressions.

The next example counts elements of *different* sets and puts them in bijection to establish an identity.

Proposition 8.1.38
Let $n,k \in \mathbb{N}$ with $n \geqslant k$. Then

$$\binom{n}{k} = \binom{n}{n-k}$$

Proof
First note that $\binom{n}{k} = \left|\binom{[n]}{k}\right|$ and $\binom{n}{n-k} = \left|\binom{[n]}{n-k}\right|$, so in order to prove $\binom{n}{k} = \binom{n}{n-k}$, it suffices by Strategy 6.1.15 to find a bijection $f : \binom{[n]}{k} \to \binom{[n]}{n-k}$. Intuitively, this bijection arises because choosing k elements from $[n]$ to *put into* a subset is equivalent to choosing $n-k$ elements from $[n]$ to *leave out of* the subset. Specifically, we define

$$f(U) = [n] \setminus U \text{ for all } U \in \binom{[n]}{k}$$

Note first that f is well-defined, since if $U \subseteq [n]$ with $|U| = k$, then $[n] \setminus U \subseteq [n]$ and $|[n] \setminus U| = |[n]| - |U| = n-k$ by Exercise 6.1.20. We now prove f is a bijection:

- f **is injective.** Let $U,V \subseteq [n]$ and suppose $[n] \setminus U = [n] \setminus V$. Then for all $k \in [n]$, we have

$$\begin{aligned} k \in U &\Leftrightarrow k \notin [n] \setminus U && \text{by definition of set difference} \\ &\Leftrightarrow k \notin [n] \setminus V && \text{since } [n] \setminus U = [n] \setminus V \\ &\Leftrightarrow k \in V && \text{by definition of set difference} \end{aligned}$$

  so $U = V$, as required.

- f **is surjective.** Let $V \in \binom{[n]}{n-k}$. Then $|[n] \setminus V| = n - (n-k) = k$ by Exercise 6.1.20, so that $[n] \setminus V \in \binom{[n]}{k}$. But then

$$f([n] \setminus V) = [n] \setminus ([n] \setminus V) = V$$

  by Exercise 2.2.43.

Since f is a bijection, we have

$$\binom{n}{k} = \left|\binom{[n]}{k}\right| = \left|\binom{[n]}{n-k}\right| = \binom{n}{n-k}$$

as required. □

We put a lot of detail into this proof. A slightly less formal proof might simply say that $\binom{n}{k} = \binom{n}{n-k}$ since choosing k elements from $[n]$ to put into a subset is equivalent to choosing $n-k$ elements from $[n]$ to leave out of the subset. This would be fine as long as the members of the intended audience of your proof could reasonably be expected to construct the bijection by themselves.

The proof of Proposition 8.1.39 follows this more informal format.

Proposition 8.1.39

Let $n, k, \ell \in \mathbb{N}$ with $n \geqslant k \geqslant \ell$. Then

$$\binom{n}{k}\binom{k}{\ell} = \binom{n}{\ell}\binom{n-\ell}{k-\ell}$$

Proof

Let's home in on the left-hand side of the equation. By the multiplication principle, $\binom{n}{k}\binom{k}{\ell}$ is the number of ways of selecting a k-element subset of $[n]$ and an ℓ-element subset of $[k]$. Equivalently, it's the number of ways of selecting a k-element subset of $[n]$ and then an ℓ-element subset *of the k-element subset that we just selected*. To make this slightly more concrete, let's put it this way:

> $\binom{n}{k}\binom{k}{\ell}$ is the number of ways of painting k balls red from a bag of n balls, and painting ℓ of the red balls blue. This leaves us with ℓ blue balls and $k-\ell$ red balls.

Now we need to find an equivalent interpretation of $\binom{n}{\ell}\binom{n-\ell}{k-\ell}$. Well, suppose we pick the ℓ elements to be coloured blue first. To make up the rest of the k-element subset, we now have to select $k-\ell$ elements, and there are now $n-\ell$ to choose from. Thus

> $\binom{n}{\ell}\binom{n-\ell}{k-\ell}$ is the number of ways of painting ℓ balls from a bag of n balls blue, and painting $k-\ell$ of the remaining balls red.

Thus, both numbers represent the number of ways of painting ℓ balls blue and $k-\ell$ balls red from a bag of n balls. Hence they are equal. □

Exercise 8.1.40

Make the proof of Proposition 8.1.39 more formal by defining a bijection between sets of the appropriate sizes.

Exercise 8.1.41

Provide a combinatorial proof that if $n, k \in \mathbb{N}$ with $n \geqslant k$, then

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$$

Deduce that the combinatorial definition of binomial coefficients (Definition 8.1.4) is equivalent to the recursive definition (Definition 4.1.15).

The following proposition demonstrates that the combinatorial definition of factorials (Definition 8.1.10) is equivalent to the recursive definition (Definition 4.1.14).

Theorem 8.1.42
$0! = 1$, and $(n+1)! = (n+1)\cdot n!$ for all $n \in \mathbb{N}$.

Proof
The only permutation of $\varnothing$ is the empty function $e : \varnothing \to \varnothing$. Hence $S_0 = \{e\}$ and $0! = |S_0| = 1$.

Let $n \in \mathbb{N}$. A permutation of $[n+1]$ is a bijection $f : [n+1] \to [n+1]$. Specifying such a bijection is equivalent to carrying out the following procedure:

- Choose the (unique!) element $k \in [n+1]$ such that $f(k) = n$. There are $n+1$ choices for k.
- Choose the values of f at each $\ell \in [n+1]$ with $\ell \neq k$. This is equivalent to finding a bijection $[n+1] \setminus \{k\} \to [n]$. Since $|[n+1] \setminus \{k\}| = |[n]| = n$, there are $n!$ such choices.

By the multiplication principle, we have

$$(n+1)! = |S_{n+1}| = (n+1)\cdot n!$$

so we're done. □

We now revisit Theorem 4.2.14; this time, our proof will be combinatorial, rather than inductive.

Theorem 8.1.43
Let $n, k \in \mathbb{N}$. Then

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & \text{if } k \leqslant n \\ 0 & \text{if } k > n \end{cases}$$

Proof
Suppose $k > n$. By Exercise 6.1.16, if $U \subseteq [n]$ then $|U| \leqslant n$. Hence if $k > n$, then $\binom{[n]}{k} = \varnothing$, and so $\binom{n}{k} = 0$, as required.

Now suppose $k \leqslant n$. We will prove that $n! = \binom{n}{k} \cdot k! \cdot (n-k)!$; the result then follows by dividing through by $k!(n-k)!$. We prove this equation by counting the number of elements of S_n.

A procedure for defining an element of S_n is as follows:

(i) Choose which elements will appear in the first k positions of the list. There are $\binom{n}{k}$ such choices.

(ii) Choose the order of these k elements. There are $k!$ such choices.

(iii) Choose the order of the remaining $n-k$ elements. There are $(n-k)!$ such choices.

By the multiplication principle, $n! = \binom{n}{k} \cdot k! \cdot (n-k)!$. □

Note that the proof of Theorem 8.1.43 relied only on the combinatorial definitions of binomial coefficients and factorials; we didn't need to know how to compute them at all! The proof was

much shorter, cleaner and, in some sense, more meaningful, than the inductive proof we gave in Theorem 4.2.14.

We conclude this section with some more examples and exercises in which double counting can be used.

Exercise 8.1.44

Let $n,k \in \mathbb{N}$ with $k \leqslant n+1$. Prove that

$$k\binom{n}{k} = (n-k+1)\binom{n}{k-1}$$

Example 8.1.45

Let $m,n,k \in \mathbb{N}$. We prove that

$$\sum_{\ell=0}^{k} \binom{m}{\ell}\binom{n}{k-\ell} = \binom{m+n}{k}$$

by finding a procedure for counting the number of k-element subsets of an appropriate $(m+n)$-element set. Specifically, let X be a set containing m cats and n dogs. Then $\left|\binom{m+n}{k}\right|$ is the number of k-element subsets $U \subseteq X$. We can specify such a subset according to the following procedure.

- **Step 1.** Split into cases based on the number ℓ of cats in U. Note that we must have $0 \leqslant \ell \leqslant k$, since the number of cats must be a natural number and cannot exceed k as $|U| = k$. Moreover, these cases are mutually exclusive. Hence by the addition principle we have

  $$\binom{m+n}{k} = \sum_{\ell=0}^{k} a_\ell$$

  where a_ℓ is the number of subsets of X containing ℓ cats and $k-\ell$ dogs.
- **Step 2.** Choose ℓ cats from the m cats in X to be elements of U. There are $\binom{m}{\ell}$ such choices.
- **Step 3.** Choose $k-\ell$ dogs from the n dogs in X to be elements of U. There are $\binom{n}{k-\ell}$ such choices.

The multiplication principle shows that $a_\ell = \binom{m}{\ell}\binom{n}{k-\ell}$. Hence

$$\binom{m+n}{k} = \sum_{\ell=0}^{k} \binom{m}{\ell}\binom{n}{k-\ell}$$

as required.

Exercise 8.1.46

Given natural numbers n,a,b,c with $a+b+c=n$, define the **trinomial coefficient** $\binom{n}{a,b,c}$ to be the number of ways of partitioning $[n]$ into three sets of sizes a, b and c, respectively. That is,

$\binom{n}{a,b,c}$ is the size of the set

$$\left\{ (A,B,C) \;\middle|\; \begin{array}{ccc} A \subseteq [n], & B \subseteq [n], & C \subseteq [n], \\ |A| = a, & |B| = b, & |C| = c, \\ & \text{and } A \cup B \cup C = [n] & \end{array} \right\}$$

By considering trinomial coefficients, prove that if $a,b,c \in \mathbb{N}$, then $(a+b+c)!$ is divisible by $a! \cdot b! \cdot c!$.

Section 8.2

Alternating sums

Using the addition principle, together with double counting, turned out to be very useful for proving combinatorial identities involving sums in Section 8.1. In this section, we turn our attention to *alternating sums*, which are sums whose terms alternate between positive and negative. As we will see later, sums of this kind can be used to computing sizes of unions of not-necessarily-disjoint sets—this has all manner of uses and applications.

An example of such a sum is the following.

$$\binom{6}{0} - \binom{6}{1} + \binom{6}{2} - \binom{6}{3} + \binom{6}{4} - \binom{6}{5} + \binom{6}{6}$$

We can express such sums more succinctly by observing that, given $k \in \mathbb{N}$, we have

$$(-1)^k = \begin{cases} 1 & \text{if } k \text{ is even} \\ -1 & \text{if } k \text{ is odd} \end{cases}$$

For example, the sum above could be expressed as $\sum_{k=0}^{6} (-1)^k \binom{6}{k}$. It so happens that this sum evaluates to zero:

$$1 - 6 + 15 - 20 + 15 - 6 + 1 \;=\; 0$$

The goal of the following exercise is to demonstrate how… *annoying*… it is to prove identities involving alternating sums using induction.

Exercise 8.2.1
Prove by induction that $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$ for all $n \in \mathbb{N}$.

Evidently we need a better approach.

If you stare at the equation in Exercise 8.2.1 for long enough, you should be able to convince yourself that

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \;=\; \sum_{\text{even } k} \binom{n}{k} - \sum_{\text{odd } k} \binom{n}{k}$$

and it suffices to prove that $\sum_{\text{even } k} \binom{n}{k} = \sum_{\text{odd } k} \binom{n}{k}$. This will be our strategy in the proof of Proposition 8.2.2, which serves as our prototype for the abstract material to come.

For the sake of readability, we left implicit that k is varying over (the even or odd elements of) the set $\{0, 1, \ldots, n\}$ in each sum—we shall adopt this practice throughout this section.

Proposition 8.2.2
Let $n \in \mathbb{N}$ with $n > 0$. Then $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$.

Proof
As we observed, it suffices to prove

$$\sum_{\text{even } k} \binom{n}{k} = \sum_{\text{odd } k} \binom{n}{k}$$

So define

$$\mathscr{E} = \{U \subseteq [n] \mid |U| \text{ is even}\} \quad \text{and} \quad \mathscr{O} = \{U \subseteq [n] \mid |U| \text{ is odd}\}$$

That is, $\mathscr{E}$ is the set of all even-sized subsets of $[n]$, and $\mathscr{O}$ is the set of all odd-sized subsets of $[n]$.

Note that the sets $\binom{[n]}{k}$ for even $k \leqslant n$ partition $\mathscr{E}$, and the sets $\binom{[n]}{k}$ for odd $k \leqslant n$ partition $\mathscr{O}$. So by the addition principle, we have

$$|\mathscr{E}| = \left|\bigcup_{\text{even } k} \binom{[n]}{k}\right| = \sum_{\text{even } k} \binom{n}{k} \quad \text{and} \quad |\mathscr{O}| = \left|\bigcup_{\text{odd } k} \binom{[n]}{k}\right| = \sum_{\text{odd } k} \binom{n}{k}$$

It suffices to show that $|\mathscr{E}| = |\mathscr{O}|$. To do this, define a function $f : \mathscr{E} \to \mathscr{O}$ for $U \in \mathscr{E}$ by

$$f(U) = \begin{cases} U \cup \{0\} & \text{if } 0 \notin U \\ U \setminus \{0\} & \text{if } 0 \in U \end{cases}$$

That is, f puts 0 into a subset if it wasn't already there, and removes it if it was. Then:

- f **is well-defined.** Given $U \in \mathscr{E}$, note that $|f(U)| = |U| \pm 1$; since $|U|$ is even, we have that $|f(U)|$ is odd, so that $f(U) \in \mathscr{O}$.
- f **is bijective.** Define $g : \mathscr{O} \to \mathscr{E}$ by letting

  $$g(V) = \begin{cases} V \cup \{0\} & \text{if } 0 \notin V \\ V \setminus \{0\} & \text{if } 0 \in V \end{cases}$$

  for all $V \in \mathscr{O}$. The proof that g is well-defined is identical to that of f. Moreover, given $U \in \mathscr{E}$, we have:
  - ◇ If $0 \in U$, then $f(U) = U \setminus \{0\}$, so that $g(f(U)) = (U \setminus \{0\}) \cup \{0\} = U$.
  - ◇ If $0 \notin U$, then $f(U) = U \cup \{0\}$, so that $g(f(U)) = (U \cup \{0\}) \setminus \{0\} = U$.

  Hence $g(f(U)) = U$ for all $U \in \mathscr{E}$. An identical computation reveals that $f(g(V)) = V$ for all $V \in \mathscr{O}$, and so g is an inverse for f.

Putting all of this together, it follows form the fact that $f : \mathscr{E} \to \mathscr{O}$ is a bijection that $|\mathscr{E}| = |\mathscr{O}|$, and so

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \sum_{\text{even } k} \binom{n}{k} - \sum_{\text{odd } k} \binom{n}{k} = |\mathscr{E}| - |\mathscr{O}| = 0$$

as required. □

Wait a minute—didn't I say this would be a *better* approach than induction? That proof felt like a lot of work. The reason for working through this proof is that it highlights the ideas that we will use throughout this section. These ideas will allow us to derive a general proof strategy, called the *involution principle* (Strategy 8.2.24), which greatly simplifies proofs of results of this nature—indeed, we will prove Proposition 8.2.2 again using the involution principle in Example 8.2.25.

With that said, Proposition 8.2.2 highlights the following general strategy for proving that an alternating sum evaluates to zero.

❖ **Strategy 8.2.3** · *Proving that an alternating sum evaluates to zero*
Let $a_0, a_1, \ldots, a_n \in \mathbb{N}$. In order to prove that $\sum_{k=0}^{n} (-1)^k a_k = 0$, it suffices to find:

(i) A partition $U_0, U_2, \ldots$ of a set $\mathscr{E}$, with $|U_k| = a_k$ for all even k;

(ii) A partition $U_1, U_3, \ldots$ of a set $\mathscr{O}$, with $|U_k| = a_k$ for all odd k; and

(iii) A bijection $\mathscr{E} \to \mathscr{O}$.

✎ **Exercise 8.2.4**
Use Strategy 8.2.3 to prove that

$$\sum_{k=0}^{n} (-1)^k \cdot k \cdot \binom{n}{k} = 0$$

for all $n \geqslant 2$.

Unfortunately Strategy 8.2.3 is still somewhat limited. For a start, it tells us nothing about how to evaluate an alternating sum that *doesn't* end up being equal to zero. Also, it ignores a key clue from the proof of Proposition 8.2.2: namely, the function $f : \mathscr{E} \to \mathscr{O}$ and its inverse $g : \mathscr{O} \to \mathscr{E}$ were defined identically. They are both restrictions of a function $h : \mathscr{P}([n]) \to \mathscr{P}([n])$ defined in the same way:

$$h(U) = \begin{cases} U \cup \{0\} & \text{if } 0 \notin U \\ U \setminus \{0\} & \text{if } 0 \in U \end{cases}$$

This function has the property that $h(h(U)) = U$ for all $U \subseteq [n]$ (that is, h is an *involution*), and h restricts to a bijection between the set of even-sized subsets of $[n]$ and the set of odd-sized subsets of $[n]$ (that is, h *swaps parity*).

This property of being a parity-swapping involution will be the key to deriving the involution principle.

Involutions

An involution is a function that is its own inverse.

✦ **Definition 8.2.5**
Let X be a set. An **involution** of X is a function $h : X \to X$ such that $h \circ h = \mathrm{id}_X$.

Example 8.2.6
Consider the function $h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = 1 - x$ for each $x \in \mathbb{R}$. Then h is an involution, since for all $x \in \mathbb{R}$ we have $h(h(x)) = 1 - (1 - x) = x$.

Exercise 8.2.7
Given a set X, prove that the relative complement function $r : \mathscr{P}(X) \to \mathscr{P}(X)$, defined by $r(U) = X \setminus U$ for all $U \subseteq X$, is an involution.

Exercise 8.2.8
Prove that every involution is a bijection.

Exercise 8.2.9
Let $h : X \to X$ be an involution and let $a \in X$. Prove that h either fixes a—that is, $h(a) = a$—or swaps it with another element $b \in X$—that is, $h(a) = b$ and $h(b) = a$.

The involution that we used in the proof of Proposition 8.2.2 was an instance of *toggling* an element in a subset—that is, removing it if it is there, and putting it in if it is not.

Toggling is so useful that we assign special notation.

Definition 8.2.10
Let X be a set. The **toggle** operation $\oplus$ (LaTeX code: `\oplus`) is defined by letting

$$U \oplus a = \begin{cases} U \cup \{a\} & \text{if } a \notin U \\ U \setminus \{a\} & \text{if } a \in U \end{cases}$$

for each $U \subseteq X$ and each $a \in X$.

Example 8.2.11
Taking $X = [3]$ and $a = 2$, we have:

$$\varnothing \oplus 2 = \{2\} \quad \{0\} \oplus 2 = \{0,2\} \quad \{1\} \oplus 2 = \{1,2\} \quad \{0,1\} \oplus 2 = \{0,1,2\}$$

$$\{2\} \oplus 2 = \varnothing \quad \{0,2\} \oplus 2 = \{0\} \quad \{1,2\} \oplus 2 = \{1\} \quad \{0,1,2\} \oplus 2 = \{0,1\}$$

The next two exercises are generalisations of facts that we showed in the proof of Proposition 8.2.2.

Exercise 8.2.12
Let X be a set and let $a \in X$. Prove that the function $T_a : \mathscr{P}(X) \to \mathscr{P}(X)$ defined by $T_a(U) = U \oplus a$ for all $U \subseteq X$ is an involution.

Exercise 8.2.13
Let X be a finite set and let $a \in X$. Prove that, for all $U \subseteq X$, if $|U|$ is even then $|U \oplus a|$ is odd, and if $|U|$ is odd then $|U \oplus a|$ is even.

The property of the toggle operation that you proved in Exercise 8.2.13 is an instance of *parity-swapping*. While toggling swaps the parity of the size of a subset, we can generalise the notion of parity more generally, provided we have a notion of what it means for an element of a set to be 'even' or 'odd'.

A first attempt to define 'even' and 'odd' might be to simply partition a set X as $X = E \cup O$, for disjoint subsets $E, O \subseteq X$—the elements of E will be deemed to be 'even' and the elements of O will be deemed to be 'odd'. But it will be helpful later on to go one step further than this: we will partition X into finitely many pieces, indexed by natural numbers, and the natural number will determine the parities of the elements of X.

Definition 8.2.14
Let X be a set and let $\mathscr{U} = \{U_0, U_1, \ldots, U_n\}$ be a partition of X for some $n \in \mathbb{N}$. The **parity** of an element $a \in X$ (relative to $\mathscr{U}$) is the parity—*even* or *odd*—of the unique $k \in \{0, 1, \ldots, n\}$ such that $a \in U_k$.

Write $X^+ = \{a \in X \mid a \text{ has even parity}\}$ (LaTeX code: `X^+`) and $X^- = \{a \in X \mid a \text{ has odd parity}\}$ (LaTeX code: `X^-`).

Note that, with notation as in Definition 8.2.14, we have partitions of X^+ and X^- as

$$X^+ = \bigcup_{\text{even } k} U_k \quad \text{and} \quad X^- = \bigcup_{\text{odd } k} U_k$$

Example 8.2.15
Let X be a finite set, and consider the partition of $\mathscr{P}(X)$ given by $U_k = \binom{X}{k}$ for all $0 \leqslant k \leqslant n$. With respect to this partition, an element $U \in \mathscr{P}(X)$ has even parity if and only if $|U|$ is even, and odd parity if and only if $|U|$ is odd.

For example, we have

$$\mathscr{P}([2])^+ = \{\varnothing, \{0,1\}\} \quad \text{and} \quad \mathscr{P}([2])^- = \{\{0\}, \{1\}\}$$

Example 8.2.16
Let $m, n \in \mathbb{N}$ and let X be the set of all functions $[n] \to [n]$. For each $k \leqslant n$, define

$$X_k = \{f : [n] \to [n] \mid |\{a \in [n] \mid f(a) = a\}| = k\}$$

That is, for each $k \leqslant n$, the set X_k is the set of all functions $f : [n] \to [n]$ that fix exactly k elements of $[n]$.

A function $f : [n] \to [n]$ has even parity with respect to this partition if it fixes an even number of elements, and odd parity if it fixes an odd number of elements.

Definition 8.2.17

Let X be a set and let $\{U_0, U_1, \ldots U_n\}$ be a partition of X for some $n \in \mathbb{N}$. A function $f : X \to X$ **swaps parity** (or is **parity-swapping**) if, for all $a \in X$, if a has even parity then $f(a)$ has odd parity, and if a has odd parity then $f(a)$ has even parity.

Example 8.2.18

With parity defined as in Example 8.2.15, the result of Exercise 8.2.12 says precisely that, for every set X and element $a \in X$, the toggle function $T_a : \mathscr{P}(X) \to \mathscr{P}(X)$ swaps parity, where T_a is defined by $T_a(U) = U \oplus a$ for all $U \subseteq X$.

Exercise 8.2.19

Let X be a finite set. Under what conditions does the involution $r : \mathscr{P}(X) \to \mathscr{P}(X)$ given by $r(U) = X \setminus U$ for all $U \subseteq X$ swap parity?

Exercise 8.2.20

Let $n \in \mathbb{N}$ and let X be the set of all functions $[n] \to [n]$, partitioned as in Example 8.2.16, so that a function $f : [n] \to [n]$ has even parity if it fixes an even number of elements, and odd parity if it fixes an odd number of elements. Find a parity-swapping function $X \to X$.

The next two following technical results will be used fundamentally in the proof of Theorem 8.2.23.

Lemma 8.2.21

Let X be a finite set, let $\{U_0, U_1, \ldots, U_n\}$ be a partition of X for some $n \in \mathbb{N}$, and let $h : X \to X$ be a parity-swapping involution. Then h induces a bijection $f : X^+ \to X^-$ defined by $f(x) = h(x)$ for all $x \in X^+$.

Proof

First note that the definition $f : X^+ \to X^-$ by letting $f(x) = h(x)$ for all $x \in X^+$ is well-defined since h swaps parity. Indeed, if $x \in X^+$, then x has even parity, so that $f(x) = h(x)$ has odd parity, meaning that $f(x) \in X^-$.

To see that f is a bijection, define a function $g : X^- \to X^+$ by $g(x) = h(x)$ for all $x \in X^-$. Again, g is well-defined since h swaps parity.

Finally note that g is an inverse for f—given $x \in X^+$, we have

$$g(f(x)) = h(h(x)) = x$$

and likewise $f(g(x)) = x$ for all $x \in X^-$.

Since f has an inverse, it is a bijection. □

Lemma 8.2.22

Let X be a finite set, let $\{U_k \mid k \in [n]\}$ be a partition of X for some $n \in \mathbb{N}$, and let $h : X \to X$ be a parity-swapping involution. Then

$$\sum_{k \in [n]} (-1)^k |U_k| = 0$$

Proof
By Lemma 8.2.21 we know that $h : X \to X$ restricts to a bijection $X^+ \to X^-$, and so we have $|X^+| = |X^-|$. By the addition principle, we have

$$\sum_{k\in[n]} (-1)^k |U_k| \;=\; \sum_{\text{even } k\in[n]} |U_k| - \sum_{\text{odd } k\in[n]} |U_k| \;=\; |X^+| - |X^-| \;=\; 0$$

as required. □

Lemma 8.2.22 gets us well on our way to deriving the involution principle. In fact, it already makes Strategy 8.2.3 obsolete: we can now prove that an alternating sum is equal to zero simply by finding a parity-swapping involution from a suitably partitioned set to itself!

But in practice, it might not be easy (or even possible) to define a parity-swapping involution $h : X \to X$ on the whole set X. In such cases, we do the best that we can: define h on some subset $D \subseteq X$, and worry about what is left over afterwards.

Theorem 8.2.23
Let X be a finite set, let $\{U_k \mid k \in [n]\}$ be a partition of X for some $n \in \mathbb{N}$, let $D \subseteq X$, let $h : D \to D$ be a parity-swapping involution, and let $F_k = U_k \setminus D$ for each $k \in [n]$. Then

$$\sum_{k\in[n]} (-1)^k |U_k| = \sum_{k\in[n]} (-1)^k |F_k|$$

Proof
Note first that the sets $U_k \cap D$ for $k \in [n]$ partition D, with the elements of D having the same parities as they did when they were considered as elements of X.

It follows from Lemma 8.2.22 that

$$\sum_{k\in[n]} |U_k \cap D| = 0$$

Moreover $|U_k| = |U_k \cap D| + |U_k \setminus D|$ for each $k \in [n]$ by the addition principle. Since $F_k = U_k \setminus D$ for each $k \in [n]$, we have

$$\sum_{k\in[n]} (-1)^k |U_k| \;=\; \underbrace{\sum_{k\in[n]} (-1)^k |U_k \cap D|}_{=0} + \sum_{k\in[n]} (-1)^k |U_k \setminus D| \;=\; \sum_{k\in[n]} (-1)^k |F_k|$$

as required. □

We have suggestively used the letter D to refer to where the involution is defined, and the letter F to refer to the elements where the involution fails.

Strategy 8.2.24 · *Involution principle*

In order to evaluate an alternating sum $\sum_{k\in[n]}(-1)^k a_k$, where $a_k \in \mathbb{N}$ for all $k \in [n]$, it suffices to follow the following steps:

(i) Find a set X with a partition $\{U_k \mid k \in [n]\}$, such that $|U_k| = a_k$ for all $k \in [n]$.

(ii) Find a parity-swapping involution $h : D \to D$ for some subset $D \subseteq X$—often it is easiest to specify the values of h first, and take D to be the set of elements of X for which the specification makes sense.

(iii) Evaluate $\sum_{k\in[n]}(-1)^k|F_k|$, where $F_k = U_k \setminus D$ for all $k \in [n]$—that is, count the elements of each U_k where the involution *failed* to be well-defined, and add them positively or negatively according to their parity.

It will often be the case that many of the sets F_k are empty, simplifying matters greatly.

This is rather abstract, so let's see some examples of the involution principle in action.

Example 8.2.25

Here is a succinct proof that $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$ for all $n \in \mathbb{N}$ using the involution principle.

Let $n \in \mathbb{N}$ and define $U_k = \binom{[n]}{k}$ for all $0 \leqslant k \leqslant n$—these sets form a partition of $\mathscr{P}([n])$, and $|U_k| = \binom{n}{k}$ for each $0 \leqslant k \leqslant n$.

By Exercise 8.2.12, the function $h : \mathscr{P}([n]) \to \mathscr{P}([n])$ defined by $h(U) = U \oplus n$ is a parity-swapping involution. By the involution principle (Strategy 8.2.24) with $D = \mathscr{P}([n])$, we have $U_k \setminus D = \varnothing$ for each $0 \leqslant k \leqslant n$, and hence

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$$

as required.

Exercise 8.2.26

Repeat Exercise 8.2.4 using the involution principle—that is, use the involution principle to prove that

$$\sum_{k=0}^{n}(-1)^k \cdot k \cdot \binom{n}{k} = 0$$

for all $n \geqslant 2$.

Exercise 8.2.27

Use the involution principle to prove that

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}\binom{k}{\ell} = 0$$

for all $n, \ell \in \mathbb{N}$ with $\ell < n$.

The next example is slightly more involved, because we find an involution that is not defined on the whole set being counted. This generalises the result of Example 8.2.25.

Proposition 8.2.28
Let $n, r \in \mathbb{N}$ with $r \leqslant n$. Then $\displaystyle\sum_{k=0}^{r} (-1)^k \binom{n}{k} = (-1)^r \binom{n-1}{r}$.

Proof
Let X be the set of subsets of $[n]$ of size $\leqslant r$, and for each $0 \leqslant k \leqslant r$, let $U_k = \binom{[n]}{k}$. Note that the sets U_k partition X for $0 \leqslant k \leqslant r$.

Define $h(U) = U \oplus 0$ for all $U \in X$. Since h is defined by toggling 0, it is a parity-swapping involution.

The only way that h can fail to be well-defined is if $|h(U)| > r$. Since $|U \oplus 0| = |U| \pm 1$ for all $U \in X$, the only way we can have $|h(U)| > r$ is if $|U| = r$ and $0 \notin U$, in which case $h(U) = U \cup \{0\}$ has size $r+1$.

Hence $F_k = \varnothing$ for all $k < r$, and $F_r = \{U \subseteq [n] \mid |U| = r \text{ and } n \notin U\}$. Specifying an element of F_r is therefore equivalent to specifying a subset of $[n-1]$ of size r, so that $|F_r| = \binom{n-1}{r}$.

Putting this all together, we obtain

$$\sum_{k=0}^{r} (-1)^k \binom{n}{k} \;=\; \sum_{k=0}^{r} (-1)^k |F_k| \;=\; (-1)^r \binom{n-1}{r}$$

as required. □

The next example is slightly more colourful.

Example 8.2.29
Let $a, b, r \in \mathbb{N}$ with $a \leqslant r \leqslant b$. We prove that

$$\sum_{k=0}^{r} (-1)^k \binom{a}{k} \binom{b}{r-k} = \binom{b-a}{r}$$

Consider a population of b animals, of which exactly a are cats. A government of exactly r animals must be formed, and a Feline Affairs Committee—which is a branch of the government—must be chosen from amongst the cats. The Feline Affairs Committee may have any size, but its size is bounded by the size of the government.

Let X be the set of all pairs (G, C), where G is a government and $C \subseteq G$ is the Feline Affairs Committee.

For $k \leqslant r$, let U_k be the set of all government–committee pairs (G,C) such that $|C| = k$—that is, such that exactly k cats sit on the Feline Affairs Committee. Note that parity is determined by the number of cats on the Feline Affairs Committee: indeed, (G,C) has even parity if $|C|$ is even, and odd parity if $|C|$ is odd.

Given a government–committee pair (G,C), let $h(G,C) = (G, C \oplus x)$, where $x \in G$ is the youngest cat on the government. That is, if the youngest cat on the government is on the Feline Affairs Committee, then that cat is removed from the committee; and if the youngest cat on the government is not on the Feline Affairs Committee, then that cat is added to the committee.

Evidently h is an involution, and it swaps parity since it adds or removes one cat to or from the Feline Affairs Committee.

The only way that h can fail to be well-defined is if there are no cats on the government, in which case $k = 0$. Thus by the involution principle

$$\sum_{k=0}^{r} (-1)^k \binom{a}{k}\binom{b}{r-k} = (-1)^0 \cdot \left| \left\{ (G,\varnothing) \in X \;\middle|\; G \text{ contains no cats} \right\} \right|$$

But there are exactly $b-a$ non-cats in the animal population, so that

$$\left| \left\{ (G,\varnothing) \in X \;\middle|\; G \text{ contains no cats} \right\} \right| = \binom{b-a}{r}$$

and hence we have $\displaystyle\sum_{k=0}^{r} (-1)^k \binom{a}{k}\binom{b}{r-k} = \binom{b-a}{r}$, as required.

If you dislike reasoning about animals, Example 8.2.29 could be reformulated by taking:

- $X = \{(A,B) \mid A \subseteq B \cap [a],\ B \subseteq [b],\ |B| = r\}$;
- $U_k = \{(A,B) \in X \mid |A| = k\}$ for all $k \leqslant r$; and
- $h(A,B) = h(A \oplus x, B)$, where x is the least element of $B \cap [a]$.

You are encouraged to verify the details!

Exercise 8.2.30

Let $n \in \mathbb{N}$ and consider the set

$$X = \{(k,i) \mid k \in [n] \text{ and } i \leqslant k\}$$

For example, if $n = 3$ then $X = \{(0,0),(1,0),(1,1),(2,0),(2,1),(2,2)\}$.

(a) Prove that $|X| = \displaystyle\sum_{k=0}^{n} k$.

(b) Use the involution principle to prove that

$$\sum_{k=0}^{n} (-1)^k k = \begin{cases} \dfrac{n}{2} & \text{if } n \text{ is even} \\ -\dfrac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$

Inclusion–exclusion principle

Our final application of the involution principle will be to prove the *inclusion–exclusion principle*, which is used for computing the sizes of unions of sets that are not necessarily pairwise disjoint.

We saw in Proposition 6.1.18 how to compute the size of a union of two not-necessarily-disjoint sets:

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

So far so good. But what if we have three or four sets instead of just two?

Exercise 8.2.31
Let X, Y, Z be sets. Show that

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

Let W be another set. Derive a similar formula for $|W \cup X \cup Y \cup Z|$.

The inclusion–exclusion principle is a generalisation of Exercise 8.2.31 to arbitrary finite collections of finite sets, but it is stated in a slightly different way in order to make the proof more convenient.

Theorem 8.2.32 · *Inclusion–exclusion principle*
Let $n \in \mathbb{N}$, let X_i be a finite set for each $i \in [n]$, and let $X = \bigcup_{i \in [n]} X_i$. Then

$$\sum_{I \subseteq [n]} (-1)^{|I|} |X_I| = 0$$

where $X_I = \{a \in X \mid a \in X_i \text{ for all } i \in I\}$.

The statement of Theorem 8.2.32 looks fairly abstract, so before we prove it, let's examine its content. The sum is over all subsets $I \subseteq [n]$, and then the power $(-1)^{|I|}$ is equal to 1 if I has an even number of elements, and -1 if I has an odd number of elements. Moreover, if I is inhabited then X_I is the intersection of the sets X_i for $i \in I$—for example $X_{\{2,3,5\}} = X_2 \cap X_3 \cap X_5$; on the other hand, a careful examination of the definition of X_I reveals that $X_\varnothing = X$.

Thus when $n = 3$, the sum $\sum_{I \subseteq [3]} (-1)^{|I|} |X_I|$ can be evaluated as

$$|X| - |X_0| - |X_1| - |X_2| + |X_0 \cap X_1| + |X_0 \cap X_2| + |X_1 \cap X_2| - |X_0 \cap X_1 \cap X_2|$$

The theorem says that this sum is equal to zero, and solving for $|X| = |X_0 \cup X_1 \cup X_2|$ yields an equivalent equation to that in Exercise 8.2.31.

Proof of Theorem 8.2.32
We will prove the inclusion–exclusion principle using the involution principle.

First we introduce some notation:

- Define $S = \{(I, a) \mid I \subseteq [n],\ a \in X_I\}$. We can think of an element $(I, a) \in S$ as being an element $a \in X$ together with a label I indicating that $a \in X_i$ for all $i \in I$.

- For each $0 \leqslant k \leqslant n$, define $S_k = \{(I,a) \in S \mid |I| = k\}$.
- For each $a \in X$, let $i_a = \min\{k \in [n] \mid a \in X_k\}$.

Note that the sets $S_0, S_1, S_2, \ldots, S_n$ form a partition of S, so we can consider the parity of an element $(I,a) \in S$—namely, the parity of (I,a) is even if $|I|$ is even, and odd if $|I|$ is odd.

Define a function $f : S \to S$ by letting

$$f(I,a) = (I \oplus i_a, a)$$

for each $I \subseteq [n]$ and each $a \in X_I$. Then:

- f is an involution since by Exercise 8.2.12 we have

$$f(f(I,a)) = f(I \oplus i_a, a) = ((I \oplus i_a) \oplus i_a, a) = (I,a)$$

- f is parity-swapping, since $|I \oplus i_a|$ and $|I|$ have opposite parity for each $a \in X$.

By the involution principle, we have

$$\sum_{k=0}^{n} (-1)^k |S_k| = 0$$

Now for fixed $I \subseteq [n]$, let $T_I = \{(I,a) \mid a \in X\}$. Then for each $0 \leqslant k \leqslant n$, the sets T_I for $|I| = k$ partition S_k, and moreover $(-1)^k = (-1)^{|I|}$, so that by the addition principle we have

$$\sum_{k=0}^{n} (-1)^k |S_k| = \sum_{k=0}^{n} \sum_{I \in \binom{[n]}{k}} (-1)^{|I|} |T_I| = \sum_{I \subseteq [n]} (-1)^{|I|} |T_I| = 0$$

Finally note that, for each $I \subseteq [n]$, the function $g_I : X_I \to T_I$ defined by $g_I(a) = (I,a)$ for all $a \in X_I$ is a bijection, with inverse given by $g_I^{-1}(I,a) = a$ for all $(I,a) \in T_I$.

Hence $|X_I| = |T_I|$, and the result is proved. □

It is more common to see the inclusion–exclusion principle stated in one two equivalent forms, stated here as Corollaries 8.2.33 and 8.2.34.

Corollary 8.2.33

Let $n \in \mathbb{N}$ and for each $i \in [n]$ let X_i be a set. Then

$$\left| \bigcup_{i \in [n]} X_i \right| = \sum_{k=1}^{n} \left(\sum_{i_1 < i_2 < \cdots < i_k < n} (-1)^{k-1} |X_{i_1} \cap X_{i_2} \cap \cdots \cap X_{i_k}| \right)$$

Proof
Moving all terms to the left-hand side of the equation and observing that $-(-1)^{k-1} = (-1)^k$, the statement is equivalent to

$$\left|\bigcup_{i\in[n]} X_i\right| - \sum_{k=1}^{n}\left(\sum_{i_1<i_2<\cdots<i_k<n} (-1)^k |X_{i_1}\cap X_{i_2}\cap\cdots\cap X_{i_k}|\right) = 0$$

But using the notation of Theorem 8.2.32, we have

$$\left|\bigcup_{i\in[n]} X_i\right| = |X| = (-1)^{|\varnothing|}|X_\varnothing|$$

and for all $i_1 < i_2 < \cdots < i_k < n$, we have

$$(-1)^k |X_{i_1}\cap X_{i_2}\cap\cdots\cap X_{i_k}| = (-1)^{|\{i_1,i_2,\ldots,i_k\}|}|X_{\{i_1,i_2,\ldots,i_k\}}|$$

and so we see that this is just a restatement of Theorem 8.2.32. □

Corollary 8.2.34
Let X be a set, let $n\in\mathbb{N}$, and for each $i\in[n]$ let $U_i\subseteq X$. Then

$$\left|X\setminus\bigcup_{i\in[n]} U_i\right| = |X| + \sum_{k=1}^{n}\left(\sum_{i_1<i_2<\cdots<i_k<n} (-1)^k |U_{i_1}\cap U_{i_2}\cap\cdots\cap U_{i_k}|\right)$$

Proof
Since $\bigcup_{i\in[n]} U_i \subseteq X$, we have

$$\left|X\setminus\bigcup_{i\in[n]} U_i\right| = |X| - \left|\bigcup_{i\in[n]} U_i\right|$$

The result then follows immediately from Corollary 8.2.33. □

Strategy 8.2.35 · *Finding the size of a union by inclusion–exclusion*
In order to find the size of a union $\bigcup_{i\in[n]} X_i$, it suffices to:

- Add the sizes of the individual sets X_i;
- Subtract the sizes of the double-intersections $X_i\cap X_j$;
- Add the sizes of the triple-intersections $X_i\cap X_j\cap X_k$;
- Subtract the sizes of the quadruple-intersections $X_i\cap X_j\cap X_k\cap X_\ell$;
- . . . and so on . . .

Continue alternating until the intersection of all the sets is covered.

Example 8.2.36

We count how many subsets of $[12]$ contain a multiple of 3. Precisely, we count the number of elements of the set

$$X_0 \cup X_3 \cup X_6 \cup X_9$$

where $X_k = \{S \subseteq [12] \mid k \in S\}$. We will apply the inclusion–exclusion principle:

(i) An element $S \in X_0$ is precisely a set of the form $\{0\} \cup S'$, where $S' \subseteq [12] \setminus \{0\}$. Since $[12] \setminus \{0\}$ has 11 elements, there are 2^{11} such subsets. So $|X_0| = 2^{11}$, and likewise $|X_3| = |X_6| = |X_9| = 2^{11}$.

(ii) An element $S \in X_0 \cap X_3$ is a set of the form $\{0,3\} \cup S'$, where $S' \subseteq [12] \setminus \{0,3\}$. Thus there are 2^{10} such subsets, so $|X_0 \cap X_3| = 2^{10}$. And likewise

$$|X_0 \cap X_6| = |X_0 \cap X_9| = |X_3 \cap X_6| = |X_3 \cap X_9| = |X_6 \cap X_9| = 2^{10}$$

(iii) Reasoning as in the last two cases, we see that

$$|X_0 \cap X_3 \cap X_6| = |X_0 \cap X_3 \cap X_9| = |X_0 \cap X_6 \cap X_9| = |X_3 \cap X_6 \cap X_9| = 2^9$$

(iv) ... and $|X_0 \cap X_3 \cap X_6 \cap X_9| = 2^8$.

Thus the number of subsets of $[12]$ which contain a multiple of 3 is

$$\underbrace{4 \times 2^{11}}_{\text{by (i)}} - \underbrace{6 \times 2^{10}}_{\text{by (ii)}} + \underbrace{4 \times 2^9}_{\text{by (iii)}} - \underbrace{2^8}_{\text{by (iv)}}$$

which is equal to 3840.

Exercise 8.2.37

How many natural numbers less than 1000 are multiples of 2, 3, 5 or 7?

Section 8.E

Chapter 8 exercises

Finite sets

8.E.1. Let $n \in \mathbb{N}$ and let $f : [n] \to [n]$ be a function. Prove that f is injective if and only if f is surjective.

8.E.2. Prove that $|\mathbb{Z}/n\mathbb{Z}| = n$ for all integers $n > 0$.

8.E.3. Let A, B and C be sets. Prove that if $A \triangle B$ and $B \triangle C$ are finite, then $A \triangle C$ is finite, where the notation $\triangle$ refers to the *symmetric difference* operation (Definition 2.E.8)—that is, $A \triangle B = (A \setminus B) \cup (B \setminus A)$, and so on.

Counting

8.E.4. Let X and Y be finite sets with $|X| = m \in \mathbb{N}$ and $|Y| = n \in \mathbb{N}$. Prove that there are 2^{mn} relations from X to Y.

8.E.5. Let X be a set and let R be a relation on X. Prove that R is reflexive if and only if $\Delta_X \subseteq \mathrm{Gr}(R)$, where Δ_X is the diagonal subset of $X \times X$ (see Definition 5.1.17). Deduce that if X is finite and $|X| = n \in \mathbb{N}$, then there are $2^{n(n-1)}$ reflexive relations on X.

8.E.6. Let X be a finite set with $|X| = n \in \mathbb{N}$. Prove that there are $2^{\binom{n}{2}} \cdot 2^n$ symmetric relations on X.

8.E.7. Let X be a finite set with $|X| = n \in \mathbb{N}$. Prove that there are $3^{\binom{n}{2}} \cdot 2^n$ antisymmetric relations on X.

8.E.8. Let X be a finite set with $|X| = n \in \mathbb{N}$, let $\sim$ be an equivalence relation on X, and suppose that there is some natural number k such that $|[a]_\sim| = k$ for all $a \in X$. Prove that k divides n, and that $|X/\sim| = \dfrac{n}{k}$.

8.E.9. Let $n, k \in \mathbb{N}$ with $n > 0$ and $k \leqslant n$. Prove that the number of functions $f : [n] \to [n]$ that fix exactly k elements of $[n]$ is equal to $\binom{n}{k}(n-1)^{n-k}$.

Double counting

8.E.10. Let $a, b, m, n \in \mathbb{N}$. Prove each of the following by double counting.

(a) $a(m+n) = am + an$

(b) $a^{m+n} = a^m \cdot a^n$

(c) $(a^m)^n = a^{mn}$

(d) $(ab)^n = a^n \cdot b^n$

8.E.11. Prove that $\displaystyle\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$ for all $n \in \mathbb{N}$

8.E.12. Prove that $\sum_{k=m}^{n} \binom{n}{k}\binom{k}{m} = 2^{n-m}\binom{n}{m}$ for all $m,n \in \mathbb{N}$ with $m \leqslant n$.

8.E.13. Prove that $\sum_{j=0}^{k} \binom{n-j}{k-j} = \binom{n+1}{k}$ for all $n,k \in \mathbb{N}$ with $k \leqslant n$.

8.E.14. Prove that $\sum_{k=1}^{n}\sum_{\ell=0}^{k} k\binom{n}{k}\binom{n-k}{\ell} = n \cdot 3^{n-1}$ for all $n \in \mathbb{N}$.

8.E.15. Prove that $\binom{n}{r+s+1} = \sum_{k=r+1}^{n-s} \binom{k-1}{r}\binom{n-k}{s}$ for all $n,r,s \in \mathbb{N}$.

8.E.16. Let $a_1,a_2,\ldots,a_r \in \mathbb{N}$ and let $n = a_1+a_2+\cdots+a_r$. Prove that

$$\binom{n}{a_1,a_2,\ldots,a_r} = \prod_{k=0}^{r-1} \binom{n-\sum_{i=1}^{k} a_i}{a_{k+1}}$$

where $\binom{n}{a_1,a_2,\ldots,a_r}$ is the number of ordered r-tuples $(U_1,U_2,\ldots,U_r)$ such that $U_1,U_2,\ldots,U_r$ is a partition of $[n]$ and $|U_k| = a_k$ for all $1 \leqslant k \leqslant r$.

Alternating sums

8.E.17. Let X be a finite set. Prove that if $|X|$ is odd then there is no parity-swapping involution $X \to X$.

8.E.18. Find the number of subsets of $[100]$ that do not contain a multiple of 8.

Miscellaneous exercises

8.E.19. Let $n \in \mathbb{N}$. For each $k \in \mathbb{N}$, find the coefficient of x^k in the polynomial $p(x)$ defined by $p(x) = (1+x+x^2)^n$.

8.E.20. Fix $n \in \mathbb{N}$. For each $k \in \mathbb{N}$, find the coefficient of x^k in the polynomial $p(x)$ defined by $p(x) = \sum_{i=0}^{n} x^{n-i}(1+x)^i$.

True–False questions

In Questions 8.E.21 to 8.E.25, determine (with proof) whether the statement is true or false.

8.E.21. Every finite set has a unique enumeration.

8.E.22. The union of a finite number of finite sets is finite.

8.E.23. Every finite set is a proper subset of another finite set.

8.E.24. There is a finite set with infinitely many subsets.

8.E.25. For every partition $\mathscr{A}$ of a finite set X, there is a function $f : X \to X$ that is a parity-swapping involution with respect to $\mathscr{A}$.

Always–Sometimes–Never questions

In Questions 8.E.26 to 8.E.32, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

8.E.26. Let X and Y be finite sets and let $f : X \to Y$ be a function. If $|X| \leqslant |Y|$, then f is injective.

8.E.27. Let X be a finite set. Then there is some proper subset $U \subsetneqq X$ such that $|U| = |X|$.

8.E.28. Let X be an infinite set. Then $\mathbb{N} \subseteq X$.

8.E.29. Let X be a finite set. Then $\mathscr{P}(X)$ is finite.

8.E.30. Let $a, b \in \mathbb{R}$ with $a < b$. Then the set $[a, b] \cap \mathbb{Q}$ is finite.

8.E.31. Let X and Y be sets with X finite, and let $f : X \to Y$ be an injection. Then Y is finite.

8.E.32. Let X and Y be sets with X finite, and let $f : X \to Y$ be a surjection. Then Y is finite.

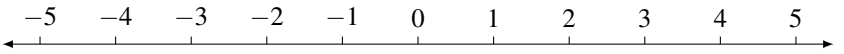
Chapter 9

Real numbers

Section 9.1

Inequalities and means

We first encountered the real numbers in Chapter 0, when the real numbers were introduced using a vague (but intuitive) notion of an *infinite number line* (Definition 0.4):



This section will scrutinise the set of real numbers in its capacity as a *complete ordered field*. Decomposing what this means:

- A *field* is a set with a notion of 'zero' and 'one', in which it makes sense to talk about addition, subtraction, multiplication, and division by everything except zero. Examples are $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}/p\mathbb{Z}$ when p is a prime number (but not when p is composite). However, $\mathbb{Z}$ is not a field, since we can't freely divide by nonzero elements—for example, $1 \in \mathbb{Z}$ and $2 \in \mathbb{Z}$, but no integer n satisfies $2n = 1$.
- An *ordered field* is a field which is equipped with a well-behaved notion of order. Both $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields, but $\mathbb{Z}/p\mathbb{Z}$ is not. We'll see why soon.
- A *complete ordered field* is an ordered field in which every set with an upper bound has a *least* upper bound. As we will see, $\mathbb{Q}$ is not a complete ordered field, but $\mathbb{R}$ is.

This is made (extremely) precise in Section B.2.

Magnitude and scalar product

In this part of the section, we home in on sets of the form $\mathbb{R}^n$, for $n \in \mathbb{N}$. Elements of $\mathbb{R}^n$ are sequences of the form $(x_1, x_2, \ldots, x_n)$, where each $x_i \in \mathbb{R}$. With our interpretation of the reals $\mathbb{R}$ as a *line*, we can interpret a sequence $(x_1, x_2, \ldots, x_n)$ as a point in *n-dimensional space*:

- 0-dimensional space is a single point. The set $\mathbb{R}^0$ has one element, namely the empty sequence $()$, so this makes sense.
- 1-dimensional space is a line. This matches our intuition that $\mathbb{R} = \mathbb{R}^1$ forms a line.
- 2-dimensional space is a *plane*. The elements of $\mathbb{R}^2$ are pairs (x, y), where x and y are both real numbers. We can interpret the pair (x, y) as *coordinates* for a point which is situated x units to the right of $(0, 0)$ and y units above $(0, 0)$ (where negative values of x or y reverse this direction)—see Figure 9.1.

With this intuition in mind, we set up the following notation.

Notation 9.1.1

Let $n \in \mathbb{N}$. Elements of $\mathbb{R}^n$ will be denoted $\vec{x}, \vec{y}, \vec{z}, \ldots$ (LaTeX code: `\vec`) and called (***n*-dimensional**) **vectors**. Given a vector $\vec{x} \in \mathbb{R}^n$, we write x_i for the i^{th} **component** of $\vec{x}$, so that

$$\vec{x} = (x_1, x_2, \ldots, x_n)$$

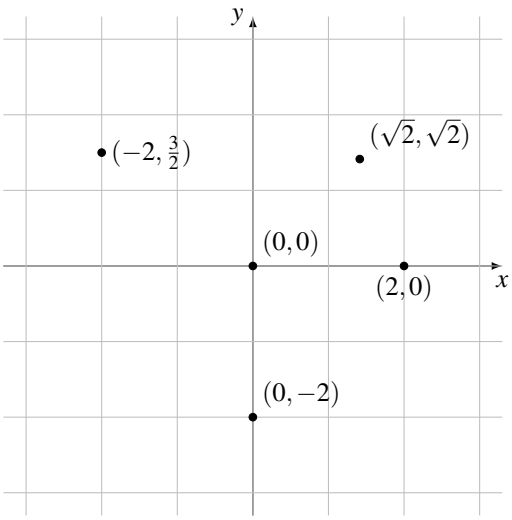


Figure 9.1: Some points in $\mathbb{R}^2$

The element $(0,0,\ldots,0) \in \mathbb{R}^n$ is called the **origin** or **zero vector** of $\mathbb{R}^n$, and is denoted by $\vec{0}$.

Moreover, if $\vec{x},\vec{y} \in \mathbb{R}^n$ and $a \in \mathbb{R}$ we write

$$\vec{x}+\vec{y} = (x_1+y_1, x_2+y_2, \ldots, x_n+y_n) \quad \text{and} \quad a\vec{x} = (ax_1, ax_2, \ldots, ax_n)$$

Example 9.1.2

For all $\vec{x} \in \mathbb{R}^n$, we have

$$\vec{x}+\vec{0} = \vec{x} \quad \text{and} \quad 1\vec{x} = \vec{x}$$

Definition 9.1.3

Let $\vec{x} \in \mathbb{R}^n$. The **magnitude** of $\vec{x}$ is the real number $\|\vec{x}\|$ (LaTeX code: `\lVert, \rVert`) defined by

$$\|\vec{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2+x_2^2+\cdots+x_n^2}$$

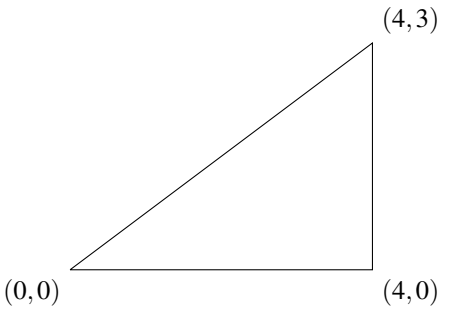
Given vectors $\vec{x},\vec{y} \in \mathbb{R}^n$, the **distance** from $\vec{x}$ to $\vec{y}$ is defined to be $\|\vec{y}-\vec{x}\|$. Thus the magnitude of a vector can be thought of as the distance from that vector to the origin.

Example 9.1.4

In $\mathbb{R}^2$, Definition 9.1.3 says that

$$\|(x,y)\| = \sqrt{x^2+y^2}$$

This matches the intuition obtained from the Pythagorean theorem on the sides of right-hand triangles. For example, consider the triangle with vertices $(0,0)$, $(4,0)$ and $(4,3)$:



The hypotenuse of the triangle has magnitude

$$\|(4,3)\| = \sqrt{4^2+3^2} = \sqrt{25} = 5$$

Exercise 9.1.5
Let $\vec{x},\vec{y} \in \mathbb{R}^n$. Prove that $\|\vec{x}-\vec{y}\| = \|\vec{y}-\vec{x}\|$. That is, the distance from $\vec{x}$ to $\vec{y}$ is equal to the distance from $\vec{y}$ to $\vec{x}$.

Exercise 9.1.6
Prove that if $x \in \mathbb{R}$ then the magnitude $\|(x)\|$ is equal to the absolute value $|x|$.

Exercise 9.1.7
Let $\vec{x} \in \mathbb{R}^n$. Prove that $\|\vec{x}\| = 0$ if and only if $\vec{x} = \vec{0}$.

The triangle inequality and the Cauchy–Schwarz inequality

The first, and simplest, inequality that we investigate is the (one-dimensional version of the) *triangle inequality* (Theorem 9.1.9). It is so named because of a generalisation to higher dimensions (Theorem 9.1.19), which can be interpreted geometrically as saying that the sum of two side lengths of a triangle is greater than or equal to the third side length.

The triangle inequality is used very frequently in mathematical proofs—you will encounter it repeatedly in this chapter—yet its proof is surprisingly simple.

Before we can prove the triangle inequality, we need the following fact about square roots of real numbers.

Lemma 9.1.8
Let $x,y \in \mathbb{R}$. If $0 \leqslant x \leqslant y$, then $\sqrt{x} \leqslant \sqrt{y}$.

Proof
Suppose $0 \leqslant x \leqslant y$. Note that, by definition of the square root symbol, we have $\sqrt{x} \geqslant 0$ and $\sqrt{y} \geqslant 0$.

Suppose $\sqrt{x} > \sqrt{y}$. By two applications of Theorem B.2.36(d), we have

$$y = \sqrt{y} \cdot \sqrt{y} < \sqrt{x} \cdot \sqrt{y} < \sqrt{x} \cdot \sqrt{x} = x$$

so that $y < x$. But this contradicts the assumption that $x \leqslant y$. Hence $\sqrt{x} \leqslant \sqrt{y}$, as required. □

Theorem 9.1.9 · *Triangle inequality in one dimension*
Let $x, y \in \mathbb{R}$. Then $|x+y| \leqslant |x| + |y|$. Moreover, $|x+y| = |x| + |y|$ if and only if x and y have the same sign.

Proof
Note first that $xy \leqslant |xy|$; indeed, xy and $|xy|$ are equal if xy is non-negative, and otherwise we have $xy < 0 < |xy|$. Also $x^2 = |x|^2$ and $y^2 = |y|^2$. Hence

$$(x+y)^2 = x^2 + 2xy + y^2 \leqslant |x|^2 + 2|xy| + |y|^2 = (|x| + |y|)^2$$

Taking (nonnegative) square roots yields

$$|x+y| \leqslant ||x| + |y||$$

by Lemma 9.1.8. But $|x| + |y| \geqslant 0$, so $||x| + |y|| = |x| + |y|$. This completes the first part of the proof.

Equality holds in the above if and only if $xy = |xy|$, which occurs if and only if $xy \geqslant 0$. But this is true if and only if x and y are both non-negative or both non-positive—that is, they have the same sign. □

Example 9.1.10
Let $x, y \in \mathbb{R}$. We prove that

$$\frac{|x+y|}{1+|x+y|} \leqslant \frac{|x|}{1+|x|} + \frac{|y|}{1+|y|}$$

First note that, if $0 \leqslant s \leqslant t$, then

$$\frac{s}{1+s} \leqslant \frac{t}{1+t}$$

To see this, note that

$$\begin{aligned}
s \leqslant t &\Rightarrow 1+s \leqslant 1+t && \text{rearranging} \\
&\Rightarrow \frac{1}{1+t} \leqslant \frac{1}{1+s} && \text{since } 1+s, 1+t > 0 \\
&\Rightarrow 1 - \frac{1}{1+s} \leqslant 1 - \frac{1}{1+t} && \text{rearranging} \\
&\Rightarrow \frac{s}{1+s} \leqslant \frac{t}{1+t} && \text{rearranging}
\end{aligned}$$

Now letting $s = |x+y|$ and $t = |x|+|y|$, we have $s \leqslant t$ by the triangle inequality, and hence

$$\frac{|x+y|}{1+|x+y|} \leqslant \frac{|x|}{1+|x|+|y|} + \frac{|y|}{1+|x|+|y|} \leqslant \frac{|x|}{1+|x|} + \frac{|y|}{1+|y|}$$

as required.

Exercise 9.1.11
Let $n \in \mathbb{N}$ and let $x_i \in \mathbb{R}$ for each $i \in [n]$. Prove that

$$\left|\sum_{i=1}^{n} x_i\right| \leqslant \sum_{i=1}^{n} |x_i|$$

with equality if and only if the numbers x_i are either all non-positive or all non-negative.

Exercise 9.1.12
Let $x, y \in \mathbb{R}$. Prove that

$$||x| - |y|| \leqslant |x-y|$$

We will generalise the triangle inequality to arbitrary dimensions in Theorem 9.1.19. Our proof will go via the *Cauchy–Schwarz inequality* (Theorem 9.1.16). To motivate the Cauchy–Schwarz inequality, we introduce another geometric notion called the *scalar product* of two vectors.

Definition 9.1.13
Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. The **scalar product** (or **dot product**) of $\vec{x}$ with $\vec{y}$ is the real number $\vec{x} \cdot \vec{y}$ (LaTeX code: `\cdot`) defined by

$$\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Example 9.1.14
Let $\vec{x} \in \mathbb{R}^n$. Then $\vec{x} \cdot \vec{x} = \|\vec{x}\|^2$. Indeed

$$\vec{x} \cdot \vec{x} = \sum_{i=1}^{n} x_i^2 = \|\vec{x}\|^2$$

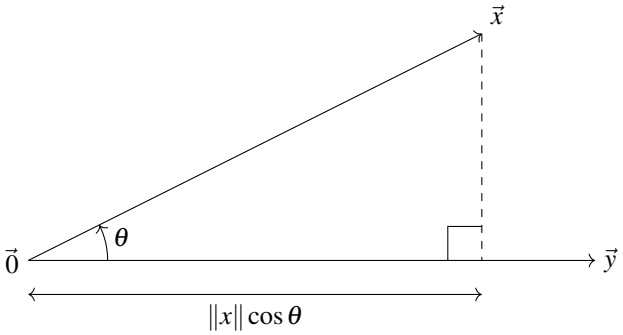
Exercise 9.1.15
Let $\vec{x}, \vec{y}, \vec{z}, \vec{w} \in \mathbb{R}^n$ and let $a, b, c, d \in \mathbb{R}$. Prove that

$$(a\vec{x} + b\vec{y}) \cdot (c\vec{z} + d\vec{w}) = ac(\vec{x} \cdot \vec{z}) + ad(\vec{x} \cdot \vec{w}) + bc(\vec{y} \cdot \vec{z}) + bd(\vec{y} \cdot \vec{w})$$

Intuitively, the scalar product of two vectors $\vec{x}$ and $\vec{y}$ measures the extent to which $\vec{x}$ and $\vec{y}$ fail to be *orthogonal*. In fact, if θ is the acute angle formed between the lines ℓ_1 and ℓ_2, where ℓ_1 passes

through $\vec{0}$ and $\vec{x}$ and ℓ_2 passes through $\vec{0}$ and $\vec{y}$, then a formula for the scalar product of $\vec{x}$ and $\vec{y}$ is given by

$$\vec{x}\cdot\vec{y} = \|\vec{x}\|\|\vec{y}\|\cos\theta$$



Evidently, $\vec{x}$ and $\vec{y}$ are orthogonal if and only if $\cos\theta = 0$, in which case $\vec{x}\cdot\vec{y} = 0$ as well. We cannot prove this yet, though, as we have not yet defined trigonometric functions or explored their properties, but hopefully this provides some useful intuition.

The Cauchy–Schwarz inequality provides a useful comparison of the size of a scalar product of two vectors with the magnitudes of the vectors.

Theorem 9.1.16 · *Cauchy–Schwarz inequality*
Let $n \in \mathbb{N}$ and let $x_i, y_i \in \mathbb{R}$ for each $i \in [n]$. Then

$$|\vec{x}\cdot\vec{y}| \leqslant \|\vec{x}\|\|\vec{y}\|$$

with equality if and only if $a\vec{x} = b\vec{y}$ for some $a, b \in \mathbb{R}$ which are not both zero.

Proof
If $\vec{y} = \vec{0}$, then this is trivial: both sides of the equation are equal to zero! So assume that $\vec{y} \neq \vec{0}$. In particular, by Exercise 9.1.7, we have $\|\vec{y}\| > 0$.

Define $k = \dfrac{\vec{x}\cdot\vec{y}}{\|\vec{y}\|^2}$. Then

$$\begin{aligned}
0 &\leqslant \|\vec{x} - k\vec{y}\|^2 && \text{since squares are nonnegative}\\
&= (\vec{x} - k\vec{y})\cdot(\vec{x} - k\vec{y}) && \text{by Example 9.1.14}\\
&= (\vec{x}\cdot\vec{x}) - 2k(\vec{x}\cdot\vec{y}) + k^2(\vec{y}\cdot\vec{y}) && \text{by Exercise 9.1.15}\\
&= \|\vec{x}\|^2 - \frac{(\vec{x}\cdot\vec{y})^2}{\|y\|^2} && \text{by definition of } k
\end{aligned}$$

Multiplying through by $\|\vec{y}\|^2$, which is non-negative and therefore doesn't change the sign of the inequality, yields

$$0 \leqslant \|\vec{x}\|^2\|\vec{y}\|^2 - (\vec{x}\cdot\vec{y})^2$$

which is equivalent to what was to be proved.

Evidently, equality holds if and only if $\|\vec{x}-k\vec{y}\|=0$, which by Exercise 9.1.7 occurs if and only if $\vec{x}-k\vec{y}=0$. Now:

- If $\vec{x}-k\vec{y}=0$, then we have

$$\begin{aligned}
&\vec{x}-k\vec{y}=0 \\
&\Leftrightarrow \vec{x}-\frac{\vec{x}\cdot\vec{y}}{\|\vec{y}\|^2}\vec{y}=0 && \text{by definition of } k \\
&\Leftrightarrow \|\vec{y}\|^2\vec{x}=(\vec{x}\cdot\vec{y})\vec{y} && \text{rearranging}
\end{aligned}$$

  If $\vec{y}\neq\vec{0}$ then let $a=\|\vec{y}\|^2$ and $b=\vec{x}\cdot\vec{y}$; otherwise, let $a=0$ and $b=1$. In both cases, we have $a\vec{x}=b\vec{y}$ and a,b are not both zero.

  If $a\vec{x}=b\vec{y}$ for some $a,b\in\mathbb{R}$ not both zero, then either:

  - $a=0$ and $b\neq 0$, in which case $\vec{y}=0$ and we have equality in the Cauchy–Schwarz inequality; or
  - $a\neq 0$, in which case $\vec{y}=\frac{b}{a}\vec{x}$. Write $c=\frac{b}{a}$. Then

$$\begin{aligned}
|\vec{x}\cdot\vec{y}| &= |\vec{x}\cdot(c\vec{x})| \\
&= |c(\vec{x}\cdot\vec{x})| && \text{by Exercise 9.1.15} \\
&= |c|\|\vec{x}\|^2 && \text{by Example 9.1.14} \\
&= \|\vec{x}\|\|c\vec{x}\| && \text{rearranging} \\
&= \|\vec{x}\|\|\vec{y}\|
\end{aligned}$$

  In either case, we have equality in the Cauchy–Schwarz inequality.

So equality holds if and only if $a\vec{x}=b\vec{y}$ for some $a,b\in\mathbb{R}$ not both zero. □

Example 9.1.17

Let $a,b,c\in\mathbb{R}$. We'll prove that

$$ab+bc+ca\leqslant a^2+b^2+c^2$$

and examine when equality holds.

Letting $\vec{x}=(a,b,c)$ and $\vec{y}=(b,c,a)$ yields

$$\vec{x}\cdot\vec{y}=ab+bc+ca$$

and

$$\|\vec{x}\|=\sqrt{a^2+b^2+c^2}=\sqrt{b^2+c^2+a^2}=\|\vec{y}\|$$

Hence $\|\vec{x}\|\|\vec{y}\|=a^2+b^2+c^2$. By the Cauchy–Schwarz inequality, it follows that

$$\vec{x}\cdot\vec{y}=ab+bc+ca\leqslant a^2+b^2+c^2=\|\vec{x}\|\|\vec{y}\|$$

as required. Equality holds if and only if $k(a,b,c)=\ell(b,c,a)$ for some $k,\ell\in\mathbb{R}$ not both zero. We may assume $k\neq 0$—otherwise, swap the vectors $\vec{x}$ and $\vec{y}$ in what follows. Then, letting $t=\frac{\ell}{k}$, we

have

$$\begin{aligned} & k(a,b,c) = \ell(b,c,a) \\ \Leftrightarrow\ & (a,b,c) = (tb,tc,ta) && \text{by definition of } t \\ \Leftrightarrow\ & (a,b,c) = (t^2c,t^2a,t^2b) && \text{substituting } a = tb \text{ etc.} \\ \Leftrightarrow\ & (a,b,c) = (t^3a,t^3b,t^3c) && \text{substituting } a = tb \text{ etc. again} \\ \Leftrightarrow\ & \vec{x} = t^3\vec{x} \end{aligned}$$

This occurs if and only if either $(a,b,c) = (0,0,0)$, or $t = 1$, in which case

$$(a,b,c) = (tb,tc,ta) = (b,c,a)$$

So equality holds if and only if $a = b = c$.

Exercise 9.1.18

Let $r \in \mathbb{N}$ and let $a_1, a_2, \ldots, a_r \in \mathbb{R}$ be such that $a_1^2 + a_2^2 + \cdots + a_n^2 = 6$. Prove that

$$(a_1 + 2a_2 + \cdots + na_n)^2 \leqslant n(n+1)(2n+1)$$

and determine when equality holds.

We now use the Cauchy–Schwarz inequality to generalise the one-dimensional version of the triangle inequality (Theorem 9.1.9) to arbitrary (finite) dimensions.

Theorem 9.1.19 · *Triangle inequality*

Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then

$$\|\vec{x} + \vec{y}\| \leqslant \|\vec{x}\| + \|\vec{y}\|$$

with equality if and only if $a\vec{x} = b\vec{y}$ for some real numbers $a, b \geqslant 0$.

Proof

We proceed by calculation:

$$\begin{aligned} \|\vec{x}+\vec{y}\|^2 &= (\vec{x}+\vec{y}) \cdot (\vec{x}+\vec{y}) && \text{by Example 9.1.14} \\ &= (\vec{x}\cdot\vec{x}) + 2(\vec{x}\cdot\vec{y}) + (\vec{y}\cdot\vec{y}) && \text{by Exercise 9.1.15} \\ &\leqslant (\vec{x}\cdot\vec{x}) + 2|\vec{x}\cdot\vec{y}| + (\vec{y}\cdot\vec{y}) && \text{since } a \leqslant |a| \text{ for all } a \in \mathbb{R} \\ &\leqslant \|\vec{x}\|^2 + 2\|x\|\|y\| + \|\vec{y}\|^2 && \text{by Example 9.1.14 and Cauchy–Schwarz} \\ &= (\|\vec{x}\| + \|\vec{y}\|)^2 && \text{rearranging} \end{aligned}$$

Taking (nonnegative) square roots of both sides yields

$$\|\vec{x} + \vec{y}\| \leqslant \|\vec{x}\| + \|\vec{y}\|$$

by Lemma 9.1.8, as required.

Equality holds if and only if the two '$\leqslant$' symbols in the above derivation are in fact '$=$' symbols.

- The first inequality is equality if and only if $\vec{x}\cdot\vec{y} = |\vec{x}\cdot\vec{y}|$, which holds if and only if $\vec{x}\cdot\vec{y} \geqslant 0$.
- The second inequality is equality if and only if equality holds in the Cauchy–Schwarz inequality. In turn, this occurs if and only if $a\vec{x} = b\vec{y}$ for some $a, b \in \mathbb{R}$. We may, moreover, assume that $a \geqslant 0$—if not, replace a and b by their negatives.

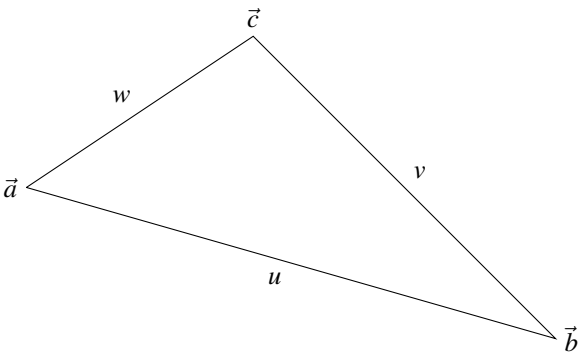
If $a = 0$ then we can take $b = 0$. If $a > 0$, then by Example 9.1.14 and Exercise 9.1.15, we have

$$\vec{x}\cdot\left(\frac{b}{a}\vec{x}\right) = \frac{b}{a}\|\vec{x}\|^2$$

which is non-negative if and only if $b \geqslant 0$, since we are assuming that $a \geqslant 0$.

Thus, equality holds in the triangle inequality if and only if $a\vec{x} = b\vec{y}$ for some $a, b \geqslant 0$. □

This general version of the triangle inequality has a geometric interpretation in terms of—you guessed it—triangles. Any three points $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ form a (potentially flat) triangle:



The side lengths u, v, w are given by the following equations:

$$u = \|\vec{b} - \vec{a}\|, \quad v = \|\vec{c} - \vec{b}\|, \quad w = \|\vec{a} - \vec{c}\|$$

The triangle inequality says tells us that $u \leqslant v + w$. Indeed:

$$\begin{aligned}
u &= \|\vec{b} - \vec{a}\| && \text{by definition of } u \\
&= \|(\vec{b} - \vec{c}) + (\vec{c} - \vec{a})\| && \text{rearranging} \\
&\leqslant \|\vec{b} - \vec{c}\| + \|\vec{c} - \vec{a}\| && \text{by the triangle inequality} \\
&= \|\vec{c} - \vec{b}\| + \|\vec{a} - \vec{c}\| && \text{by Exercise 9.1.5} \\
&= v + w && \text{by definition of } v \text{ and } w
\end{aligned}$$

That is, the triangle inequality says that the sum of two side lengths of a triangle is greater than or equal to the third side length. Moreover, it tells us $u = v + w$ precisely when $k(\vec{a} - \vec{c}) = \ell(\vec{c} - \vec{b})$ for some $k, \ell \geqslant 0$. If $k = 0$ then

$$\vec{c} \quad = \quad \vec{b} \quad = \quad 0\vec{a} + (1 - 0)\vec{b}$$

if $k > 0$, then $k + \ell > 0$, so we have

$$\vec{c} \quad = \quad \frac{k}{k+\ell}\vec{a} + \frac{\ell}{k+\ell}\vec{b} \quad = \quad \frac{k}{k+\ell}\vec{a} + \left(1 - \frac{k}{k+\ell}\right)\vec{b}$$

Examining this a bit more closely yields that $u = v + w$ if and only if

$$\vec{c} = t\vec{a} + (1-t)\vec{b}$$

for some $0 \leqslant t \leqslant 1$, which is to say precisely that $\vec{c}$ lies on the line segment between $\vec{a}$ and $\vec{b}$. In other words, equality holds in the triangle inequality only if the three vertices of the triangle are *collinear*, which is to say that the triangle whose vertices are the points $\vec{a}$, $\vec{b}$ and $\vec{c}$, is flat.

Inequalities of means

Our goal now is to explore different kinds of average—specifically, *means*—of finite sets of non-negative real numbers. We will compare the relative sizes of these means with respect to one-another, with emphasis on three particular kinds of mean: the *arithmetic mean* (Definition 9.1.20), the *geometric mean* (Definition 9.1.21) and the *harmonic mean* (Definition 9.1.29). These means in fact assemble into a continuum of means, called *generalised means* (Definition 9.1.37), all of which can be compared with one another.

Definition 9.1.20

Let $n \geqslant 1$. The (**arithmetic**) **mean** of real numbers $x_1, \ldots, x_n$ is

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Definition 9.1.21

Let $n \geqslant 1$. The **geometric mean** of non-negative real numbers $x_1, \ldots, x_n$ is

$$\sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \cdots \cdot x_n}$$

The following theorem is commonly known as the **AM–GM inequality**.

Theorem 9.1.22 · *Inequality of arithmetic and geometric means*

Let $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \geqslant 0$. Then

$$\underbrace{\sqrt[n]{x_1 \cdots x_n}}_{\text{geometric mean}} \leqslant \underbrace{\frac{x_1 + \cdots + x_n}{n}}_{\text{arithmetic mean}}$$

with equality if and only if $x_1 = \cdots = x_n$.

Proof when $n = 2$

We need to show that, if $x, y \in \mathbb{R}$ with $x, y \geqslant 0$, then

$$\sqrt{xy} \leqslant \frac{x+y}{2}$$

with equality if and only if $x = y$.

First note that the square roots of x and y exist since they are non-negative. Now

$$\begin{aligned} 0 &\leqslant (\sqrt{x} - \sqrt{y})^2 && \text{since squares are nonnegative} \\ &= (\sqrt{x})^2 - 2\sqrt{x}\sqrt{y} + (\sqrt{y})^2 && \text{expanding} \\ &= x - 2\sqrt{xy} + y && \text{rearranging} \end{aligned}$$

Rearranging the inequality $0 \leqslant x - 2\sqrt{xy} + y$ yields the desired result.

If $\sqrt{xy} = \frac{x+y}{2}$, then we can rearrange the equation as follows:

$$\begin{aligned} \sqrt{xy} = \frac{x+y}{2} &\Rightarrow 2\sqrt{xy} = x + y && \text{multiplying by 2} \\ &\Rightarrow 4xy = x^2 + 2xy + y^2 && \text{squaring both sides} \\ &\Rightarrow x^2 - 2xy + y^2 = 0 && \text{rearranging} \\ &\Rightarrow (x-y)^2 = 0 && \text{factorising} \\ &\Rightarrow x - y = 0 && \text{since } a^2 = 0 \Rightarrow a = 0 \text{ for } a \in \mathbb{R} \\ &\Rightarrow x = y && \text{rearranging} \end{aligned}$$

So we have proved both parts of the theorem. □

Example 9.1.23

We use the AM–GM inequality to prove that the area of a rectangle with fixed perimeter is maximised when the rectangle is a square.

Indeed, fix a perimeter $p > 0$, and let $x, y > 0$ be side lengths of a rectangle with perimeter p—that is, x and y satisfy the equation $2x + 2y = p$. The area a of the rectangle satisfies $a = xy$. By the AM–GM inequality, we have

$$a = xy \leqslant \left(\frac{x+y}{2}\right)^2 = \frac{p^2}{16}$$

Equality holds if and only if $x = y$, in other words, if and only if the rectangle is a square.

Exercise 9.1.24

Let $a, b > 0$ be real numbers. Prove that $\dfrac{a^2 + b^2}{2} \geqslant ab$.

Example 9.1.25

Let $x > 0$. We find the minimum possible value of $x + \frac{9}{x}$. By the AM–GM inequality, we have

$$x + \frac{9}{x} \geqslant 2\sqrt{x \cdot \frac{9}{x}} = 2\sqrt{9} = 6$$

with equality if and only if $x = \frac{9}{x}$, which occurs if and only if $x = 3$. Hence the minimum value of $x + \frac{9}{x}$ when $x > 0$ is 6.

Exercise 9.1.26

Let $x > 0$ and let $n \in \mathbb{N}$. Find the minimum possible value of $\sum_{k=-n}^{n} x^k$.

Exercises 9.1.27 and 9.1.28 complete the proof of the AM–GM inequality (Theorem 9.1.22). Before proceeding with the exercises, let's fix some notation: for each $n \in \mathbb{N}$, let $p_{\text{AM–GM}}(n)$ be the assertion that the AM–GM inequality holds for collections of n numbers; that is, $p_{\text{AM–GM}}(n)$ is the assertion:

> For all $x_1, x_2, \ldots, x_n \geqslant 0$, we have
>
> $$\sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{1}{n} \sum_{i=1}^{n} x_i$$
>
> with equality if and only if $x_1 = x_2 = \cdots = x_n$.

Note that we already proved $p_{\text{AM–GM}}(2)$.

Exercise 9.1.27

Let $r \in \mathbb{N}$ and let $x_1, x_2, \ldots, x_{2r} \in \mathbb{R}$. Write

$$a = \frac{1}{r} \sum_{i=1}^{r} x_i \quad \text{and} \quad g = \sqrt[r]{\prod_{i=1}^{r} x_i}$$

for the arithmetic and geometric means, respectively, of the numbers $x_1, \ldots, x_r$; write

$$a' = \frac{1}{r} \sum_{i=r+1}^{2r} x_i \quad \text{and} \quad g' = \sqrt[r]{\prod_{i=r+1}^{2r} x_i}$$

for the arithmetic and geometric means, respectively, of the numbers $x_{r+1}, \ldots, x_{2r}$; and write

$$A = \frac{1}{2r} \sum_{i=1}^{2r} x_i \quad \text{and} \quad G = \sqrt[2r]{\prod_{i=1}^{2r} x_i}$$

for the arithmetic and geometric means, respectively, of all the numbers $x_1, \ldots, x_{2r}$.

Prove that

$$A = \frac{a + a'}{2} \quad \text{and} \quad G = \sqrt{gg'}$$

Deduce that, for each $r \in \mathbb{N}$, if $p_{\text{AM–GM}}(r)$ is true then $p_{\text{AM–GM}}(2r)$ is true. Deduce further than $p_{\text{AM–GM}}(2^m)$ is true for all $m \in \mathbb{N}$.

Exercise 9.1.28

Let $r \geqslant 2$ and let $x_1, \ldots, x_{r-1} \in \mathbb{N}$. Define

$$x_r = \frac{1}{r-1} \sum_{i=1}^{r-1} x_i$$

Prove that

$$\frac{1}{r} \sum_{i=1}^{r} x_i = x_r$$

Assuming $p_{\text{AM–GM}}(r)$, deduce that

$$x_r^r \geqslant \prod_{i=1}^{r} x_i = \left(\prod_{i=1}^{r-1} x_i \right) \cdot x_r$$

with equality if and only if $x_1 = x_2 = \cdots = x_r$. Deduce that $p_{\text{AM–GM}}(r)$ implies $p_{\text{AM–GM}}(r-1)$. Use Exercise 9.1.27 to deduce further that $p_{\text{AM–GM}}(n)$ is true for all $n \geqslant 1$.

We now introduce another kind of mean, called the *harmonic mean*.

Definition 9.1.29

Let $n \in \mathbb{N}$. The **harmonic mean** of nonzero real numbers $x_1, x_2, \ldots, x_n$ is

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} \right)^{-1} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

The harmonic mean of two nonzero real numbers x and y has a simpler expression:

$$\left(\frac{x^{-1} + y^{-1}}{2} \right)^{-1} = \frac{2xy}{x+y}$$

The harmonic mean arises naturally when considering rates of change of quantities over fixed amounts of time.

Example 9.1.30

The cities of York and Leeds are located $d > 0$ miles apart. Two cars drive from York to Leeds, then immediately turn around and drive back. The two cars depart from York at the same time and arrive back in York at the same time.

- The first car drives from York to Leeds at a constant speed of u miles per hour, and drives back to York at a constant speed of v miles per hour.
- The second car drives from York to Leeds and back again at the same constant speed of w miles per hour.

According to the following formula from physics:

$$\text{speed} \times \text{time} = \text{distance}$$

the time spent driving by the first car is $\frac{d}{u}+\frac{d}{v}$, and the time spent driving by the second car is $\frac{2d}{w}$.

Since the cars spend the same amount of time driving, it follows that

$$\frac{2d}{w}=\frac{d}{u}+\frac{d}{v} \qquad \Rightarrow \qquad w=\frac{2uv}{u+v}$$

That is, the second car's speed is the harmonic mean of the two speeds driven by the first car.

As might be expected, we now prove a theorem relating the harmonic means with the other means we have established so far—this theorem is known as the **GM–HM inequality**.

Theorem 9.1.31 · *Inequality of geometric and harmonic means*
Let $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n > 0$. Then

$$\underbrace{\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}}_{\text{harmonic mean}} \leqslant \underbrace{\sqrt[n]{x_1 x_2 \cdots x_n}}_{\text{geometric mean}}$$

with equality if and only if $x_1 = \cdots = x_n$.

Proof when $n = 2$
We need to prove that if $x, y > 0$, then

$$\frac{2}{\frac{1}{x}+\frac{1}{y}} \leqslant \sqrt{xy}$$

This is almost immediate from the AM–GM inequality (Theorem 9.1.22). Indeed, since all numbers in sight are positive, we can take reciprocals to see that this inequality is equivalent to the assertion that

$$\frac{1}{\sqrt{xy}} \leqslant \frac{x^{-1}+y^{-1}}{2}$$

But $\frac{1}{\sqrt{xy}} = \sqrt{x^{-1}y^{-1}}$, so this is immediate from the AM–GM inequality. □

Exercise 9.1.32
Prove the GM–HM inequality for general values of $n \in \mathbb{N}$.

Another example of a mean that has applications in probability theory and statistics is that of the *quadratic mean*.

Definition 9.1.33
Let $n \in \mathbb{N}$. The **quadratic mean** (or **root-mean-square**) of real numbers $x_1, x_2, \ldots, x_n$ is

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = \sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}}$$

The following theorem is, predictably, known as the **QM–AM inequality** (or **RMS–AM inequality**); it is a nice application of the Cauchy–Schwarz inequality.

Theorem 9.1.34 · *Inequality of quadratic and arithmetic means*
Let $n > 0$ and $x_1, x_2, \ldots, x_n \geqslant 0$. Then

$$\underbrace{\frac{x_1 + \cdots + x_n}{n}}_{\text{arithmetic mean}} \leqslant \underbrace{\sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}}_{\text{quadratic mean}}$$

with equality if and only if $x_1 = \cdots = x_n$.

Proof
Define

$$\vec{x} = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad \vec{y} = (1, 1, \ldots, 1)$$

Then

$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= \vec{x} \cdot \vec{y} && \text{by definition of scalar product} \\ &\leqslant \|\vec{x}\| \|\vec{y}\| && \text{by Cauchy–Schwarz} \\ &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \cdot \sqrt{n} && \text{evaluating the magnitudes} \end{aligned}$$

Dividing through by n yields

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \leqslant \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$$

as required. Equality holds if and only if equality holds in the Cauchy–Schwarz inequality, which occurs if and only if

$$(ax_1, ax_2, \ldots, ax_n) = (b, b, \ldots, b)$$

for some $a, b \in \mathbb{R}$ not both zero. If $a = 0$ then $b = 0$, so we must have $a \neq 0$. Hence equality holds if and only if $x_i = \frac{b}{a}$ for all $i \in [n]$—in particular, if and only if $x_1 = x_2 = \cdots = x_n$. □

To summarise, what we have proved so far is

$$\text{harmonic mean} \overset{(9.1.31)}{\leqslant} \text{geometric mean} \overset{(9.1.22)}{\leqslant} \text{arithmetic mean} \overset{(9.1.34)}{\leqslant} \text{quadratic mean}$$

with equality in each case if and only if the real numbers whose means we are taking are all equal.

The following exercise allows us to bookend our chain of inequalities with the minimum and maximum of the collections of numbers.

Exercise 9.1.35
Let $n > 0$ and let $x_1, x_2, \ldots, x_n$ be positive real numbers. Prove that

$$\min\{x_1, x_2, \ldots, x_n\} \leqslant \left(\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} \right)^{-1} \quad \text{and} \quad \max\{x_1, x_2, \ldots, x_n\} \geqslant \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

with equality in each case if and only if $x_1 = x_2 = \cdots = x_n$.

⋆ Generalised means

We conclude this section by mentioning a generalisation of the results we have proved about means. We are not yet ready to prove the results that we mention; they are only here for the sake of interest.

Definition 9.1.36

The **extended real number line** is the (ordered) set $[-\infty,\infty]$, defined by

$$[-\infty,\infty] = \mathbb{R} \cup \{-\infty,\infty\}$$

where $\mathbb{R}$ is the set of real numbers with its usual ordering, and $-\infty,\infty$ are new elements ordered in such a way that $-\infty < x < \infty$ for all $x \in \mathbb{R}$.

Note that the extended real line does *not* form a field—the arithmetic operations are not defined on $-\infty$ or ∞, and we will at no point treat $-\infty$ and ∞ as real numbers; they are merely elements which extend the reals to add a least element and a greatest element.

Definition 9.1.37

Let $p \in [-\infty,\infty]$, let $n \in \mathbb{N}$, and let $x_1, x_2, \ldots, x_n$ be positive real numbers. The **generalised mean with exponent** p (or simply p-**mean**) $x_1, x_2, \ldots, x_n$ is the real number $M_p(x_1, x_2, \ldots, x_n)$ defined by

$$M_p(x_1, x_2, \ldots, x_n) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^p \right)^{\frac{1}{p}}$$

if $p \notin \{-\infty, 0, \infty\}$, and by

$$M_p(x_1, x_2, \ldots, x_n) = \lim_{q \to p} M_q(x_1, x_2, \ldots, x_n)$$

if $p \in \{-\infty, 0, \infty\}$, where the notation $\lim_{q \to p}$ refers to the *limit* as q tends to p, as discussed in Section B.3.

We can see immediately that the harmonic, arithmetic and quadratic means of a finite set of positive real numbers are the p-means for a suitable value of p: the harmonic mean is the (-1)-mean, the arithmetic mean is the 1-mean, and the quadratic mean is the 2-mean. Furthermore, Proposition 9.1.38 exhibits the *minimum* as the $(-\infty)$-mean, the *geometric mean* as the 0-mean, and the *maximum* as the ∞-mean.

Proposition 9.1.38

Let $n > 0$ and let $x_1, x_2, \ldots, x_n \geqslant 0$. Then

- $M_{-\infty}(x_1, x_2, \ldots, x_n) = \min\{x_1, x_2, \ldots, x_n\}$;
- $M_0(x_1, x_2, \ldots, x_n) = \sqrt[n]{x_1 x_2 \cdots x_n}$; and
- $M_{\infty}(x_1, x_2, \ldots, x_n) = \max\{x_1, x_2, \ldots, x_n\}$.

All of the inequalities of means we have seen so far will be subsumed by Theorem 9.1.39, which compares the p-mean and q-mean of a set of numbers for all values of $p, q \in [-\infty, \infty]$.

Theorem 9.1.39

Let $n > 0$, let $x_1, x_2, \ldots, x_n \geqslant 0$ and let $p, q \in [-\infty, \infty]$ with $p < q$. Then

$$M_p(x_1, x_2, \ldots, x_n) \leqslant M_q(x_1, x_2, \ldots, x_n)$$

with equality if and only if $x_1 = x_2 = \cdots = x_n$.

Theorem 9.1.39 implies each of the following:

- **HM–min inequality** (Exercise 9.1.35): take $p = -\infty$ and $q = -1$;
- **GM–HM inequality** (Theorem 9.1.31): take $p = -1$ and $q = 0$;
- **AM–GM inequality** (Theorem 9.1.22): take $p = 0$ and $q = 1$;
- **QM–AM inequality** (Theorem 9.1.34): take $p = 1$ and $q = 2$;
- **max–QM inequality** (Exercise 9.1.35): take $p = 2$ and $q = \infty$.

Section 9.2

Completeness and convergence

For most of the results that we proved in Section 9.1, it did not matter that we were talking about real numbers. We could just as well have been working with any other ordered field, such as the rational numbers—that is, most of the results in Section 9.1 remain true by replacing $\mathbb{R}$ by $\mathbb{Q}$ (or any other ordered field) throughout.

From here onwards, we isolate the property of $\mathbb{R}$ that separates it from $\mathbb{Q}$—namely, *completeness*. It is completeness that will allow us to define and explore the fundamental concepts of mathematical analysis: sequences, functions, convergence, limits, continuity, differentiability, and so on.

The property of completeness concerns least upper bounds for certain sets of real numbers.

Definition 9.2.1

Let $A \subseteq \mathbb{R}$. A real number m is an **upper bound** for A if $a \leqslant m$ for all $a \in A$. A **supremum** of A is a *least* upper bound of A; that is, a real number m such that:

(i) m is an upper bound of A—that is, $a \leqslant m$ for all $a \in A$; and

(ii) m is least amongst all upper bounds for A—that is, for all $x \in \mathbb{R}$, if $a \leqslant x$ for all $a \in A$, then $m \leqslant x$.

Example 9.2.2

We prove that 1 is a supremum of the open interval $(0,1)$.

(i) Let $a \in (0,1)$. Then $a < 1$, so that 1 is an upper bound of $(0,1)$.

(ii) Let $x \in \mathbb{R}$ be another upper bound of $(0,1)$. If $x < 1$, then we have

$$x = \frac{x+x}{2} < \frac{x+1}{2} < \frac{1+1}{2} = 1$$

and so $x < \dfrac{x+1}{2} \in (0,1)$. This contradicts the assumption that x is an upper bound of $(0,1)$. It follows that $x \geqslant 1$, as required.

Hence 1 is indeed a supremum of $(0,1)$.

Exercise 9.2.3

Define the notions of **lower bound** and **infimum**, and find the infimum of the open interval $(0,1)$.

The following proposition provides a convenient way of testing whether a real number is a supremum of a subset.

Proposition 9.2.4

Let $A \subseteq \mathbb{R}$ and suppose $m \in \mathbb{R}$ is an upper bound of A. Then m is a supremum of A if and only if, for all $\varepsilon > 0$, there exists $a \in A$ such that $a > m - \varepsilon$.

Proof

- ($\Rightarrow$). Suppose m is a supremum of A, and let $\varepsilon > 0$. If there is no $a \in A$ such that $a > m - \varepsilon$, then $a \leqslant m - \varepsilon$ for all $a \in A$. But this contradicts the assumption that m is a supremum of a, since $m - \varepsilon$ is an upper bound of A that is less than m. So there exists $a \in A$ with $a > m - \varepsilon$, as required.
- ($\Leftarrow$). Suppose that, for all $\varepsilon > 0$, there exists $a \in A$ with $a > m - \varepsilon$, and let $x \in \mathbb{R}$ be an upper bound of A. In order to prove that m is a supremum of A, we must prove that $m \leqslant x$.

  Suppose $x < m$, and define $\varepsilon = m - x$. Then $\varepsilon > 0$, so there exists $a \in A$ such that

  $$a > m - \varepsilon = m - (m - x) = x$$

  But this contradicts the assumption that x is an upper bound of A. So we must have $m \leqslant x$, as required. □

Theorem 9.2.5 · *Uniqueness of suprema*
Let A be a subset of $\mathbb{R}$. If m_1 and m_2 are suprema of A, then $m_1 = m_2$.

Proof
Since m_1 is an upper bound of A and m_2 is a supremum of A, we have $m_2 \leqslant m_1$ by Definition 9.2.1(ii). Likewise, since m_2 is an upper bound of A and m_1 is a supremum of A, we have $m_1 \leqslant m_2$ by Definition 9.2.1(ii) again. But this implies that $m_1 = m_2$. □

An analogous result proves that a subset of $\mathbb{R}$ may have at most one infimum. This allows us to introduce the following notation.

Definition 9.2.6
Let $A \subseteq \mathbb{R}$. The supremum of A, if it exists is denoted by $\sup(A)$ (LaTeX code: `\mathrm{sup}`); the infimum of A, if it exists, is denoted by $\inf(A)$ (LaTeX code: `\mathrm{inf}`).

Now that we are more familiar with suprema, here is the completeness axiom in its full glory.

Axiom 9.2.7 · *Completeness axiom*
Let $A \subseteq \mathbb{R}$ be inhabited. If A has an upper bound, then A has a supremum.

The true power of the completeness axiom will become apparent later in the section when we discuss the existence of limits of sequences of real numbers.

Before we embark on that adventure, we first prove that the rational numbers are *not* complete, by exhibiting a subset of $\mathbb{Q}$ that has no rational supremum.

Proposition 9.2.8
Let $A = \{x \in \mathbb{Q} \mid x^2 < 2\}$. Then A does not have a rational supremum.

A quick proof of Proposition 9.2.8 would be to verify that $\sqrt{2}$, which is irrational, is a supremum of A, and use uniqueness of suprema to deduce that there can be no rational supremum. However, this is cheating. Failure of completeness is an *intrinsic* property—we should be able to prove Proposition 9.2.8 without venturing outside of the realm of rational numbers at all. That is, we cannot use irrational numbers in our proof. This makes the proof significantly longer, but significantly more satisfying.

Proof of Proposition 9.2.8
Towards a contradiction, suppose that A has a supremum q.

First note that $q > 0$. Indeed, $1^2 < 2$, so that $1 \in A$, and so $q \geqslant 1 > 0$.

Next, we prove that $q^2 = 2$. Indeed:

- Assume $q^2 < 2$, so that $2 - q^2 > 0$. For each $n \geqslant 1$, we have
$$\left(q + \frac{1}{n}\right)^2 = q^2 + \frac{2q}{n} + \frac{1}{n^2}$$
Choose n sufficiently large that $\dfrac{2q}{n} < \dfrac{2-q^2}{2}$ and $\dfrac{1}{n^2} < \dfrac{2-q^2}{2}$. Then by the above, we observe that
$$\left(q + \frac{1}{n}\right)^2 < q^2 + \frac{2-q^2}{2} + \frac{2-q^2}{2} = q^2 + (2-q^2) = 2$$
and so $q + \frac{1}{n} \in A$. But $q + \frac{1}{n} > q$, so this contradicts the assumption that q is an upper bound of A.
- Assume $q^2 > 2$, so that $q^2 - 2 > 0$. For each $n \geqslant 1$, we have
$$\left(q - \frac{1}{n}\right)^2 = q^2 - \frac{1}{n}\left(2q - \frac{1}{n}\right)$$
Choose n sufficiently large that $\frac{1}{n} < q$ $(< 2q)$ and $\frac{2q}{n} < q^2 - 2$. Then by the above work, we observe that
$$\left(q - \frac{1}{n}\right)^2 > q^2 - \frac{2q}{n} > q^2 - (q^2 - 2) = 2$$
Moreover $q - \frac{1}{n} > 0$ since $\frac{1}{n} < q$.
Suppose that $q - \frac{1}{n}$ is *not* an upper bound for A. Then there is some $x \in A$ with $x > q - \frac{1}{n} > 0$. But then $(q - \frac{1}{n})^2 < x^2 < 2$, contradicting the fact that $\left(q - \frac{1}{n}\right)^2 > 2$.
So $q - \frac{1}{n}$ is an upper bound for A, contradicting the fact that q is a supremum of A.

So we must have $q^2 = 2$. But this is impossible—the proof is identical to that of Proposition 4.3.12, but with all instances of '$\sqrt{2}$' replaced by 'q' in the proof.

So $\{x \in \mathbb{Q} \mid x^2 < 2\}$ has no rational supremum. □

Sequences of real numbers

The rest of this chapter is dedicated to studying *convergence* of sequences of real numbers. We will use the completeness axiom to find sufficient conditions for a sequence to converge.

Definition 9.2.9

A **sequence of real numbers** is a function $x : \mathbb{N} \to \mathbb{R}$. Given a sequence x, we write x_n instead of $x(n)$ and write $(x_n)_{n\geqslant 0}$, or even just (x_n), instead of $x : \mathbb{N} \to \mathbb{R}$. The values x_n are called the **terms** of the sequence, and the variable n is called the **index** of the term x_n.

Example 9.2.10

Some very basic but very boring examples of sequences are *constant sequences*. For example, the constant sequence with value 0 is

$$(0,0,0,0,0,0,\dots)$$

More generally, for fixed $a \in \mathbb{R}$, the constant sequence with value a is defined by $x_n = a$ for all $n \in \mathbb{N}$.

Example 9.2.11

Sequences can be defined just like functions. For example, there is a sequence defined by $x_n = 2^n$ for all $n \in \mathbb{N}$. Writing out the first few terms, this sequence is

$$(1,2,4,8,16,\dots)$$

Sometimes it will be convenient to start the indexing of our sequence from numbers other than 0, particularly when an expression involving a variable n isn't defined when $n = 0$. We'll denote such sequences by $(x_n)_{n\geqslant 1}$ or $(x_n)_{n\geqslant 2}$, and so on.

Example 9.2.12

Let $(z_n)_{n\geqslant 2}$ be the sequence defined by $z_n = \frac{(n+1)(n+2)}{(n-1)n}$ for all $n \geqslant 2$:

$$\left(6, \frac{10}{3}, \frac{5}{2}, \frac{21}{10}, \dots\right)$$

The indexing of this sequence begins at 2, rather than 0, since when $n = 0$ or $n = 1$, the expression $\frac{(n+1)(n+2)}{(n-1)n}$ is undefined. We could *reindex* the sequence: by letting $z'_n = z_{n+2}$ for all $n \geqslant 0$, we obtain a new sequence $(z'_n)_{n\geqslant 0}$ defined by $z'_n = \frac{(n+3)(n+4)}{(n+1)(n+2)}$ whose indexing starts from 0. Fortunately for us, such matters won't cause any problems—it's just important to make sure that whenever we define a sequence, we make sure the terms make sense for all of the indices.

Convergence of sequences

Of particular interest to us will be sequences whose terms get closer and closer to a fixed real number. This phenomenon is called *convergence*.

Example 9.2.13

Consider the sequence $(y_n)_{n\geqslant 1}$ defined by $y_n = \frac{1}{n}$ for all $n \geqslant 1$:

$$\left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots\right)$$

The terms y_n become closer and closer to 0 as n grows.

Example 9.2.14

Define a sequence $(r_n)_{n\geqslant 0}$ by $r_n = \frac{2n}{n+1}$ for all $n \in \mathbb{N}$. Some of the values of this sequence are illustrated in the following table:

n	r_n	decimal expansion
0	0	0
1	1	1
2	$\frac{4}{3}$	$1.333\dots$
3	$\frac{3}{2}$	1.5
10	$\frac{20}{11}$	$1.818\dots$
100	$\frac{200}{101}$	$1.980\dots$
1000	$\frac{2000}{1001}$	$1.998\dots$
$\vdots$	$\vdots$	$\vdots$

As n increases, the values of r_n become closer and closer to 2.

The precise sense in which the terms of the sequences in Examples 9.2.13 and 9.2.14 'get closer' to 0 and 2, respectively, is called *convergence*, which we will define momentarily in Definition 9.2.15.

First, let's try to work out what the definition *should be* for a sequence (x_n) to converge to a real number a.

A naïve answer might be to say that the sequence is 'eventually equal to a'—that is, after some point in the sequence, all terms are equal to a. Unfortunately, this isn't quite good enough: if it were, then the values $r_n = \frac{2n}{n+1}$ from Example 9.2.14 would be equal to 2 for sufficiently large n. However, if for some $n \in \mathbb{N}$ we have $\frac{2n}{n+1} = 2$, then it follows that $2n = 2(n+1)$; rearranging this gives $1 = 0$, which is a contradiction.

However, this answer isn't too far from giving us what we need. Instead of saying that the terms x_n are eventually *equal* to a, we might want to say that they become *infinitely close* to a, whatever that means.

We can't really make sense of an 'infinitely small positive distance' (e.g. Exercise 1.1.39), so we might instead make sense of 'infinitely close' by saying that the terms x_n eventually become as close to a as we could possibly want them to be. Spelling this out, this means that for any positive distance ε (LaTeX code: `\varepsilon`) (read: 'epsilon')[a] no matter how small, the terms x_n are eventually within distance ε of a. In summary:

[a] The lower case Greek letter *epsilon* (ε) is traditionally used in analysis to denote a positive quantity whose value can be made arbitrarily small. We will encounter this letter frequently in this section and the next when discussing convergence.

Definition 9.2.15

Let (x_n) be a sequence and let $a \in \mathbb{R}$. We say that (x_n) **converges** to a, and write $(x_n) \to a$ (LaTeX code: `\to`), if the following condition holds:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, |x_n - a| < \varepsilon$$

The value a is called a **limit** of (x_n). Moreover, we say that a sequence (x_n) **converges** if it has a limit, and diverges otherwise.

Sometimes, we may write '$x_n \to a$ as $n \to \infty$' to mean $(x_n) \to a$; this indicates that the terms x_n approach a as n increases without bound. Take heed of the fact that the symbol '∞' in here does not have meaning on its own—it is simply a means of suggesting that as the index n gets greater, the values x_n of the terms in the sequence get closer to the limit.

Before we move onto some examples, let's quickly digest the definition of the expression $(x_n) \to a$. The following table presents a suggestion of how you might read the expression '$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, |x_n - a| < \varepsilon$' in English.

Symbols	English
$\forall \varepsilon > 0 \ldots$	For any positive distance ε (no matter how small)…
$\ldots \exists N \in \mathbb{N} \ldots$	…there is a stage in the sequence…
$\ldots \forall n \geqslant N \ldots$	…after which all terms in the sequence…
$\ldots \lvert x_n - a \rvert < \varepsilon.$	…are within distance ε of a.

Thus, a sequence (x_n) converges to a if '*for any positive distance ε (no matter how small), there is a stage in the sequence after which all terms in the sequence are within ε of a*'. After reading this a few times, you should hopefully be content that this definition captures what is meant by saying that the terms in the sequence are eventually as close to a as we could possibly want them to be.

We are now ready to see some examples of convergent (and divergent) sequences. When reading the following proofs, keep in mind the logical structure—that is, the alternating quantifiers $\forall \varepsilon \ldots \exists N \ldots \forall n \ldots$—in the definition of $(x_n) \to a$.

Proposition 9.2.16

The sequence (y_n) defined by $y_n = \frac{1}{n}$ for all $n \geqslant 1$ converges to 0.

Proof

By Definition 9.2.15, we need to prove

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, \left| \frac{1}{n} - 0 \right| < \varepsilon$$

So fix $\varepsilon > 0$. Our goal is to find $N \in \mathbb{N}$ such that $\left|\frac{1}{n}\right| < \varepsilon$ for all $n \geqslant N$.

Let N be any natural number which is greater than $\frac{1}{\varepsilon}$. Then for all $n \geqslant N$, we have

$$\begin{aligned}
\left|\frac{1}{n}\right| &= \frac{1}{n} && \text{since } \frac{1}{n} > 0 \text{ for all } n \geqslant 1 \\
&\leqslant \frac{1}{N} && \text{since } n \geqslant N \\
&< \frac{1}{1/\varepsilon} && \text{since } N > \frac{1}{\varepsilon} \\
&= \varepsilon
\end{aligned}$$

Hence $|y_n| < \varepsilon$ for all $n \geqslant N$. Thus we have proved that $(y_n) \to 0$. □

Remark 9.2.17

The value of N you need to find in the proof of convergence will usually depend on the parameter ε. (For instance, in Proposition 9.2.16, we defined N to be some natural number greater than $\frac{1}{\varepsilon}$.) This is to be expected—remember that ε is the distance away from the limit that the terms are allowed to vary after the N^{th} term. If you make this distance smaller, you'll probably have to go further into the sequence before your terms are all close enough to a. In particular, the value of N will usually grow as the value of ε gets smaller. This was the case in Proposition 9.2.16: note that $\frac{1}{\varepsilon}$ increases as ε decreases.

Example 9.2.18

Let (r_n) be the sequence from Example 9.2.14 defined by $r_n = \dfrac{2n}{n+1}$ for all $n \in \mathbb{N}$. We'll prove that $(r_n) \to 2$. So fix $\varepsilon > 0$. We need to find $N \in \mathbb{N}$ such that

$$\left|\frac{2n}{n+1} - 2\right| < \varepsilon \text{ for all } n \geqslant N$$

To find such a value of n, we'll first do some algebra. Note first that for all $n \in \mathbb{N}$ we have

$$\left|\frac{2n}{n+1} - 2\right| = \left|\frac{2n - 2(n+1)}{n+1}\right| = \left|\frac{-2}{n+1}\right| = \frac{2}{n+1}$$

Rearranging the inequality $\frac{2}{n+1} < \varepsilon$ gives $\frac{n+1}{2} > \frac{1}{\varepsilon}$, and hence $n > \frac{2}{\varepsilon} - 1$.

To be clear, what we've shown so far is that a *necessary* condition for $|r_n - 2| < \varepsilon$ to hold is that $n > \frac{2}{\varepsilon} - 1$. This informs us what the desired value of N might look like—we will then verify that the desired inequality holds.

So define $N = \frac{2}{\varepsilon} - 1$. For all $n \geqslant N$, we have

$$\begin{aligned}
\left|\frac{2n}{n+1} - 2\right| &= \frac{2}{n+1} && \text{by the above work} \\
&\leqslant \frac{2}{N+1} && \text{since } n \geqslant N \\
&< \frac{2}{\left(\frac{2}{\varepsilon} - 1\right) + 1} && \text{since } N > \frac{2}{\varepsilon} - 1 \\
&= \frac{2}{2/\varepsilon} && \text{rearranging} \\
&= \varepsilon && \text{rearranging}
\end{aligned}$$

Thus, as claimed, we have $|r_n - 2| < \varepsilon$ for all $n \geqslant N$. It follows that $(r_n) \to 2$, as required.

Exercise 9.2.19
Let (x_n) be the constant sequence with value $a \in \mathbb{R}$. Prove that $(x_n) \to a$.

Exercise 9.2.20
Prove that the sequence (z_n) defined by $z_n = \frac{n+1}{n+2}$ converges to 1.

Here's a slightly more involved example.

Proposition 9.2.21
Let $r \in (-1, 1)$. Then $(r^n) \to 0$.

Proof
If $r = 0$, then $r^n = 0$ for all $n \geqslant 1$, and so for any $\varepsilon > 0$ and $n \geqslant 1$ we have

$$|r^n - 0| = |0| = 0 < \varepsilon$$

so that $(r^n) \to 0$ as required.

So assume $r \neq 0$ and let $a = \dfrac{1}{|r|} > 1$. Then $a = 1 + \delta$ for some $\delta > 0$, so that by the binomial theorem we have

$$a^n \;=\; (1+\delta)^n \;=\; 1 + n\delta + \sum_{k=2}^{n} \binom{n}{k} \delta^{n-k} \;\geqslant\; 1 + n\delta$$

for all $n \geqslant 1$.

Now let $\varepsilon > 0$, and let $N \geqslant 2$ be such that $1 + N\delta > \dfrac{1}{\varepsilon}$; any $N > \dfrac{1-\varepsilon}{\delta\varepsilon}$ will do.

Then for all $n \geqslant N$, we have

$$|r^n| \;=\; \frac{1}{a^n} \;\leqslant\; \frac{1}{a^N} \;\leqslant\; \frac{1}{1+N\delta} \;<\; \frac{1}{1/\varepsilon} \;=\; \varepsilon$$

and so $(r^n) \to 0$, as required. □

Divergence

Before we go too much further, let's see some examples of sequences which *diverge*. Recall (Definition 9.2.15) that a sequence (x_n) converges if $(x_n) \to a$ for some $a \in \mathbb{R}$. Spelling this out symbolically, to say '(x_n) converges' is to say the following:

$$\exists a \in \mathbb{R}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, |x_n - a| < \varepsilon$$

We can negate this using the tools of Section 1.3: to say that a sequence (x_n) diverges is to say the following:

$$\forall a \in \mathbb{R}, \exists \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \geqslant N, |x_n - a| \geqslant \varepsilon$$

In more intuitive terms: for all possible candidates for a limit $a \in \mathbb{R}$, there is a positive distance ε such that, no matter how far down the sequence you go (say x_N), you can find a term x_n beyond that point which is at distance $\geqslant \varepsilon$ away from a.

Example 9.2.22

Let (x_n) be the sequence defined by $x_n = (-1)^n$ for all $n \in \mathbb{N}$:

$$(1, -1, 1, -1, 1, -1, \dots)$$

We'll prove that (x_n) diverges. Fix $a \in \mathbb{R}$. Intuitively, if a is non-negative, then it must be at distance $\geqslant 1$ away from -1, and if a is negative, then it must be at distance $\geqslant 1$ away from 1. We'll now make this precise.

So let $\varepsilon = 1$, and fix $N \in \mathbb{N}$. We need to find $n \geqslant N$ such that $|(-1)^n - a| \geqslant 1$. We'll split into cases based on whether a is non-negative or negative.

- Suppose $a \geqslant 0$. Then $-1 - a \leqslant -1 < 0$, so that we have

  $$|-1 - a| = a - (-1) = a + 1 \geqslant 1$$

  So let $n = 2N + 1$. Then $n \geqslant N$ and n is odd, so that

  $$|x_n - a| = |(-1)^n - a| = |-1 - a| \geqslant 1$$

- Suppose $a < 0$. Then $1 - a > 1 > 0$, so that we have

  $$|1 - a| = 1 - a > 1$$

  So let $n = 2N$. Then $n \geqslant N$ and n is even, so that

  $$|x_n - a| = |(-1)^n - a| = |1 - a| \geqslant 1$$

In both cases, we've found $n \geqslant N$ such that $|x_n - a| \geqslant 1$. It follows that (x_n) diverges.

Example 9.2.22 is an example of a *periodic* sequence—that is, it's a sequence that repeats itself. It is difficult for such sequences to converge since, intuitively speaking, they jump up and down a lot. (In fact, the only way that a period sequence *can* converge is if it is a constant sequence!)

Exercise 9.2.23
Let (y_n) be the sequence defined by $y_n = n$ for all $n \in \mathbb{N}$:

$$(0, 1, 2, 3, \dots)$$

Prove that (y_n) diverges.

Exercise 9.2.24
Let $r \in \mathbb{R}$. Recall that $(r^n) \to 0$ if $|r| < 1$ (this was Proposition 9.2.21) and that (r^n) diverges if $r = -1$ (this was Example 9.2.22). Prove that (r^n) diverges if $|r| > 1$.

'Eventually'

Consider the following sequence:

$$\left(1,\ 2,\ -10,\ 7,\ \frac{1}{\sqrt{2}},\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ \dots\right)$$

It takes some nonzero values initially, but after the 5^{th} term in the sequence, it remains constant with the value 0. For most intents and purposes, we can treat it as a constant sequence: after a certain point, it is constant, and so any properties involving *limits* of constant sequences will also be true of this sequence.

Situations like this arise frequently. For example, we might not need a sequence to be *increasing* (Definition 9.2.44)—we might just need it to be increasing after some finite stage.

We use the word 'eventually' to refer to this phenomenon. (In fact, the word 'eventually' is a new kind of quantifier!)

Definition 9.2.25
Let $p(x)$ be a logical formula with free variable x ranging over sequences of real numbers. We say $p((x_n)_{n\geqslant 0})$ is **eventually** true if $p((x_n)_{n\geqslant k})$ is true for some $k \in \mathbb{N}$.

Example 9.2.26
Some examples of the word 'eventually' include:

- A sequence (x_n) is *eventually constant* if $(x_n)_{n\geqslant k}$ is constant for some $k \in \mathbb{N}$—that is, if there is some $k \in \mathbb{N}$ such that $x_m = x_n$ for all $m, n \geqslant k$.
- A sequence (x_n) is *eventually nonzero* if there is some $k \in \mathbb{N}$ such that $x_n \neq 0$ for all $n \geqslant k$.
- Two sequences (x_n) and (y_n) are *eventually equal* if there is some $k \in \mathbb{N}$ such that $x_n = y_n$ for all $n \geqslant k$.

Example 9.2.27
The definition of $(x_n) \to a$ can be equivalently phrased as:

> For all $\varepsilon > 0$, the sequence (x_n) eventually satisfies $|x_n - a| < \varepsilon$.
>
> This is because '$\exists N \in \mathbb{N}, \forall n \geqslant N, |x_n - a| < \varepsilon$' means precisely that $|x_n - a|$ is eventually less than ε.

✎ **Exercise 9.2.28**

Prove that if a sequence (x_n) converges to a nonzero limit, then (x_n) is eventually nonzero. Find a sequence (x_n) that converges to zero, but is not eventually nonzero.

✎ **Exercise 9.2.29**

Let (x_n) be a sequence and let $p(x)$ be a logical formula. What does it mean to say that $p(x_n)$ is *not* eventually true? Find a sentence involving the phrase 'not eventually' that is equivalent to the assertion that (x_n) diverges.

The next theorem will allow us to use the word 'eventually' in our proofs, without worrying about whether we're being precise.

Theorem 9.2.30 · *'Eventually' preserves conjunction and disjunction*

Let (x_n) be a sequence, and let $p(x)$ and $q(x)$ be logical formula with free variable x ranging over sequences of real numbers.

(a) If $p(x_n)$ is eventually true and $q(x_n)$ is eventually true, then $p(x_n) \wedge q(x_n)$ is eventually true.

(b) If $p(x_n)$ is eventually true or $q(x_n)$ is eventually true, then $p(x_n) \vee q(x_n)$ is eventually true.

Proof

(a) Let $k, \ell \in \mathbb{N}$ be such that $p(x_n)$ is true for all $n \geqslant k$ and $q(x_n)$ is true for all $n \geqslant \ell$. Define $N = \max\{k, \ell\}$. Then for all $n \geqslant N$, we have $p(x_n)$ is true since $n \geqslant N \geqslant k$, and $q(x_n)$ is true since $n \geqslant N \geqslant \ell$, so that $p(x_n) \wedge q(x_n)$ is true for all $n \geqslant N$. Hence $p(x_n) \wedge q(x_n)$ is eventually true.

(b) Assume that $p(x_n)$ is eventually true. Then there is some $k \in \mathbb{N}$ such that $p(x_n)$ is true for all $n \geqslant k$. But then $p(x_n) \vee q(x_n)$ is true for all $n \geqslant k$, so that $p(x_n) \vee q(x_n)$ is eventually true. Likewise, if $q(x_n)$ is eventually true, then $p(x_n) \vee q(x_n)$ is eventually true. □

The next exercise urges you not to become too complacent with your use of the word 'eventually'.

✎ **Exercise 9.2.31** · *'Eventually' does not preserve negation*

Find a sequence (x_n) and a logical formula $p(x)$ such that $p(x_n)$ is neither eventually true nor eventually false. (Thus '$p(x_n)$ is not eventually true' does not imply '$\neg p(x_n)$ is eventually true'.)

The following proposition justifies our use of 'eventually' in proofs regarding limits—it implies that limiting behaviour of a sequence is not affected by changing (or completely disregarding) the finitely many terms at the beginning of the sequence.

Theorem 9.2.32

Let (x_n) and (y_n) be sequences. If (x_n) and (y_n) are eventually equal, then (x_n) converges if and only if (y_n) converges and, if $(x_n) \to a \in \mathbb{R}$, then $(y_n) \to a$ as well.

Proof

- First assume that (x_n) converges to $a \in \mathbb{R}$. We prove that $(y_n) \to a$.

  So fix $\varepsilon > 0$. Since $(x_n) \to a$, eventually we have $|x_n - a| < \varepsilon$ by Example 9.2.27. But eventually $x_n = y_n$, and so we eventually have

  $$|y_n - a| = |x_n - a| < \varepsilon$$

  as required.

- Now assume that (x_n) diverges. We prove that (y_n) diverges. So let $a \in \mathbb{R}$, and fix $\varepsilon > 0$ such that, for all $N \in \mathbb{N}$, we have $|x_n - a| \geqslant \varepsilon$ for some $n \geqslant N$.

  Let $M \in \mathbb{N}$ and define $N = \max\{k, N\}$, where $k \in \mathbb{N}$ is such that $x_n = y_n$ for all $n \geqslant k$.

  Since (x_n) diverges, there is some $n \geqslant N$ such that $|x_n - a| \geqslant \varepsilon$. But then $n \geqslant N \geqslant M$ and

  $$|y_n - a| = |x_n - a| \geqslant \varepsilon$$

  so that (y_n) diverges. □

Computing limits

Finding limits of sequences can be tricky. Theorem 9.2.34 makes it slightly easier by saying that if a sequence is built up using arithmetic operations—addition, subtraction, multiplication and division—from sequences whose limits you know, then you can simply apply those arithmetic operations to the limits.

In order to prove part of Theorem 9.2.34, however, the following lemma will be useful.

Lemma 9.2.33

Let (x_n) be a sequence of real numbers. If (x_n) converges, then (x_n) is bounded—that is, there is some real number k such that $|x_n| \leqslant k$ for all $n \in \mathbb{N}$.

Proof

Let $a \in \mathbb{R}$ be such that $(x_n) \to a$. Letting $\varepsilon = 1$ in the definition of convergence, it follows that there exists some $N \in \mathbb{N}$ such that $|x_n - a| < 1$ for all $n \geqslant N$. It follows that $-1 < x_n - a < 1$ for all $n \geqslant N$, and hence $-(1-a) < x_n < 1 + a$ for all $n \geqslant N$.

Now define

$$k = \max\{|x_0|, |x_1|, \ldots, |x_{N-1}|, |1-a|, |1+a|\} + 1$$

For all $n < N$, we have

$$-k < -|x_n| \leqslant x_n \leqslant |x_n| < k$$

so that $|x_n| < k$. For all $n \geqslant N$, we have

$$-k < -|1-a| \leqslant -(1-a) < x_n < 1+a \leqslant |1+a| < k$$

so that $|x_n| < k$.

Hence $|x_n| < k$ for all $n \in \mathbb{N}$, as required. □

Theorem 9.2.34

Let (x_n) and (y_n) be sequences of real numbers, let $a, b \in \mathbb{R}$, and suppose that $(x_n) \to a$ and $(y_n) \to b$. Then

(a) $(x_n + y_n) \to a + b$;

(b) $(x_n - y_n) \to a - b$;

(c) $(x_n y_n) \to ab$; and

(d) $(\frac{x_n}{y_n}) \to \frac{a}{b}$, so long as $b \neq 0$.

Proof of (a) and (c)

(a). Fix $\varepsilon > 0$. We need to prove that eventually $|(x_n + y_n) - (a+b)| < \varepsilon$.

- Since $(x_n) \to a$, we eventually have $|x_n - a| < \frac{\varepsilon}{2}$;
- Since $(y_n) \to b$, we eventually have $|x_n - b| < \frac{\varepsilon}{2}$.

It follows from the triangle inequality (Theorem 9.1.9) that we eventually have

$$|(x_n + y_n) - (a+b)| = |(x_n - a) + (y_n - b)| \leqslant |x_n - a| + |y_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

as required.

(c). This one is a little harder. Fix $\varepsilon > 0$. Since (x_n) converges, it follows from Lemma 9.2.33 that there is some real number k with $|x_n| < k$ for all $n \in \mathbb{N}$.

- Since $(x_n) \to a$, we eventually have $|x_n - a| < \frac{\varepsilon}{2|b|}$;
- Since $(y_n) \to b$, we eventually have $|x_n - b| < \frac{\varepsilon}{2k}$.

Then using the triangle inequality again, eventually we have:

$$\begin{aligned}
|x_n y_n - ab| &= |x_n(y_n - b) + b(x_n - a)| && \text{rearranging} \\
&\leqslant |x_n(y_n - b)| + |b(x_n - a)| && \text{by the triangle inequality} \\
&= |x_n||y_n - b| + |b||x_n - a| && \text{rearranging} \\
&< k|y_n - b| + |b||x_n - a| && \text{since } |x_n| < k \text{ for all } n \\
&< k\frac{\varepsilon}{2k} + |b|\frac{\varepsilon}{2|b|} && \text{(eventually)} \\
&= \varepsilon && \text{rearranging}
\end{aligned}$$

Hence $(x_n y_n) \to ab$, as required. □

Exercise 9.2.35
Prove parts (b) and (d) of Theorem 9.2.34.

Theorem 9.2.34 *appears* obvious, but as you can see in the proof, it is more complicated than perhaps expected. It was worth the hard work, though, because we can now compute more complicated limits formed in terms of arithmetic operations by taking the limits of the individual components.

The following example uses Theorem 9.2.34 to prove that $\left(\frac{2n}{n+1}\right) \to 2$ in a much simpler way than we saw in Example 9.2.18.

Example 9.2.36
We provide another proof that the sequence (r_n) of Example 9.2.14, defined by $r_n = \frac{2n}{n+1}$ for all $n \in \mathbb{N}$, converges to 2.

For all $n \geqslant 1$, dividing by the top and bottom gives

$$r_n = \frac{2}{1+\frac{1}{n}}$$

The constant sequences (2) and (1) converge to 2 and 1, respectively; and by Proposition 9.2.16, we know that $(\frac{1}{n}) \to 0$. It follows that

$$(r_n) \to \frac{2}{1+0} = 2$$

as required.

Exercise 9.2.37
Let (x_n) be a sequence of real numbers converging to a real number a, and let $p(x) = a_0 + a_1 x + \cdots + a_d x^d$ be a polynomial function. Prove that $(p(x_n)) \to p(a)$, and that $\left(\frac{1}{p(x_n)}\right) \to \frac{1}{p(a)}$ if $p(a) \neq 0$.

The so-called *squeeze theorem* provides another means of computing limits. It says that if we can eventually 'squeeze' the terms of a sequence (y_n) between terms of two other sequences that converge to the same limit, then we can deduce that (y_n) converges to the same limit.

Theorem 9.2.38 · *Squeeze theorem*
Let (x_n), (y_n) and (z_n) be sequences of real numbers such that:

(i) $(x_n) \to a$ and $(z_n) \to a$; and

(ii) Eventually $x_n \leqslant y_n \leqslant z_n$.

Then $(y_n) \to a$.

Proof
Fix $\varepsilon > 0$. We need to prove that, eventually, $|y_n - a| < \varepsilon$.

Since $(x_n) \to a$ and $(z_n) \to a$, we eventually have $|x_n - a| < \varepsilon$ and $|z_n - a| < \varepsilon$.

Fix $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $|y_n - a| < \varepsilon$, $|z_n - a| < \varepsilon$ and $x_n < y_n < z_n$. Given $n \geqslant N$:

- If $y_n \geqslant a$, then we have $a \leqslant y_n \leqslant z_n$, and so
$$|y_n - a| = y_n - a \leqslant z_n - a = |z_n - a| < \varepsilon$$
- If $y_n < a$, then we have $x_n \leqslant y_n \leqslant a$, and so
$$|y_n - a| = a - y_n \leqslant a - x_n = |x_n - a| < \varepsilon$$

In both cases we have proved $|y_n - a| < \varepsilon$. It follows that $(y_n) \to a$. □

Example 9.2.39
Fix $k \geqslant 1$. We prove that the sequence $\left(\frac{1}{n^k}\right)_{n \geqslant 1}$ converges to zero.

Note that $n^k > n$, so that we have $0 < \frac{1}{n^k} \leqslant \frac{1}{n}$ for all $n \in \mathbb{N}$. We know that $(\frac{1}{n}) \to 0$ by Example 9.2.13, and $(0) \to 0$ since it is a constant sequence, so the squeeze theorem implies that $(\frac{1}{n^k}) \to 0$.

Exercise 9.2.40
Fix $r \in \mathbb{N}$, and let $p(x) = a_0 + a_1 x + \cdots + a_r x^r$ and $q(x) = b_0 + b_1 x + \cdots + b_r x^r$ be polynomials with real coefficients. Prove that if $b_r \neq 0$, then $\left(\frac{p(n)}{q(n)}\right) \to \frac{a_r}{b_r}$.

Uniqueness of limits

We now prove that a sequence can have at most one limit. This will allow us to talk about 'the' limit of a sequence, and introduce notation for the limit of a sequence.

Theorem 9.2.41 · *Uniqueness of limits*
Let (x_n) be a sequence and let $a, b \in \mathbb{R}$. If $(x_n) \to a$ and $(x_n) \to b$, then $a = b$.

Proof
We'll prove that $|a - b| = 0$, which will imply that $a = b$. To do this, we'll prove that $|a - b|$ is not positive: we already know it's non-negative, so this will imply that it is equal to zero. To prove that $|a - b|$ is not positive, we'll prove that it is less than every positive number.

So fix $\varepsilon > 0$. Then also $\frac{\varepsilon}{2} > 0$. The definition of convergence (Definition 9.2.15) tells us that eventually $|x_n - a| < \frac{\varepsilon}{2}$ and $|x_n - b| < \frac{\varepsilon}{2}$.

By the triangle inequality (Theorem 9.1.9), it follows that eventually

$$\begin{aligned} |a-b| &= |(a-x_n)+(x_n-b)| && \text{by cancelling the } x_n \text{ terms} \\ &\leqslant |a-x_n|+|x_n-b| && \text{by the triangle inequality} \\ &= |x_n-a|+|x_n-b| && \text{by Exercise 9.1.5} \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon && \text{(eventually)} \end{aligned}$$

Since $|a-b|<\varepsilon$ for all $\varepsilon>0$, it follows that $|a-b|$ is a non-negative real number that is less than every positive real number, so that it is equal to zero.

Since $|a-b|=0$, we have $a-b=0$, and so $a=b$. □

Theorem 9.2.41 justifies the following notation.

Definition 9.2.42

Let (x_n) be a convergent sequence. The limit of (x_n) is denoted by $\lim_{n\to\infty}(x_n)$ (LaTeX code: `\lim_{n \to \infty}`).

[The usual warnings about the symbol ∞ apply.]

Example 9.2.43

Proposition 9.2.16 and Example 9.2.18 tell us that

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)=0 \quad \text{and} \quad \lim_{n\to\infty}\left(\frac{2n}{n+1}\right)=2$$

Existence of limits

It is often useful to know *that* a sequence converges, but not necessary to go to the arduous lengths of computing its limit. However, as it currently stands, we don't really have any tools for proving that a sequence converges other than finding a limit for it! The remainder of this section is dedicated to deriving tools for finding out when a sequence does or does not converge, without needing to know exactly what the limit is.

Perhaps the most fundamental result is the *monotone convergence theorem* (Theorem 9.2.48), since it underlies the proofs of all the other results that we will prove. What it says is that if the terms in a sequence always increase, or always decrease, and the set of terms in the sequence is bounded, then the sequence converges to a limit.

The sequence (r_n) from Example 9.2.14, defined by $r_n=\frac{2n}{n+1}$ for all $n\in\mathbb{N}$, is an example of such a sequence. We proved that it converged by computing its limit in Example 9.2.18 and again in Example 9.2.36. We will soon (Example 9.2.51) use the monotone convergence theorem to give *yet another proof* that it converges, but this time without going to the trouble of first finding its limit.

Before we can state the monotone convergence theorem, we must first define what we mean by a *monotonic sequence*.

Definition 9.2.44

A sequence of real numbers (x_n) is...

- ...**increasing** if $m \leqslant n$ implies $x_m \leqslant x_n$ for all $m, n \in \mathbb{N}$;
- ...**decreasing** if $m \leqslant n$ implies $x_m \geqslant x_n$ for all $m, n \in \mathbb{N}$.

If a sequence is either increasing or decreasing, we say it is **monotonic**.

Example 9.2.45

The sequence (x_n) defined by $x_n = n^2$ for all $n \in \mathbb{N}$ is increasing, since for all $m, n \in \mathbb{N}$, if $m \leqslant n$, then $m^2 \leqslant n^2$. To see this, note that if $m \leqslant n$, then $n - m \geqslant 0$ and $n + m \geqslant 0$, so that

$$n^2 - m^2 = (n-m)(n+m) \geqslant 0 \cdot 0 = 0$$

and hence $n^2 \geqslant m^2$, as required.

Example 9.2.46

The sequence (r_n) from Examples 9.2.14 and 9.2.36, defined by $r_n = \frac{2n}{n+1}$ for all $n \in \mathbb{N}$, is increasing. To see this, suppose $m \leqslant n$. Then $n = m + k$ for some $k \geqslant 0$. Now

$$\begin{aligned}
& 0 \leqslant k && \text{by assumption} \\
\Leftrightarrow\ & m^2 + km + m \leqslant m^2 + km + m + k && \text{adding } m^2 + km + m \text{ to both sides} \\
\Leftrightarrow\ & m(m+k+1) \leqslant (m+1)(m+k) && \text{factorising} \\
\Leftrightarrow\ & m(n+1) \leqslant (m+1)n && \text{since } n = m+k \\
\Leftrightarrow\ & \frac{m}{m+1} \leqslant \frac{n}{n+1} && \text{dividing both sides by } (m+1)(n+1) \\
\Leftrightarrow\ & r_m \leqslant r_n && \text{by definition of } (r_n)
\end{aligned}$$

Note that the step where we divided through by $(m+1)(n+1)$ is justified since this quantity is positive.

It is perhaps useful to add that to *come up with* this proof, it is more likely that you would start with the assumption $r_m \leqslant r_n$ and derive that $k \geqslant 0$—noting that all steps are reversible then allows us to write it in the 'correct' order.

Exercise 9.2.47

Prove that the sequence $(5^n - n^5)_{n \geqslant 0}$ is *eventually* increasing—that is, there is some $k \in \mathbb{N}$ such that $(5^n - n^5)_{n \geqslant k}$ is an increasing sequence.

The monotone convergence theorem underlies all of the other tools for proving convergence of sequences that are to follow. It makes essential use of the completeness axiom.

Theorem 9.2.48 · *Monotone convergence theorem*
Let (x_n) be a sequence of real numbers.

(a) If (x_n) is increasing and has an upper bound, then it converges;

(b) If (x_n) is decreasing and has a lower bound, then it converges.

Proof of (a)
We prove (a) here—part (b) is Exercise 9.2.49.

So suppose (x_n) is increasing and has an upper bound. Then:

(i) $x_m \leqslant x_n$ for all $m \leqslant n$; and

(ii) There is some real number u such that $u \geqslant x_n$ for all $n \in \mathbb{N}$.

Condition (ii) tells us that the set $\{x_n \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$ has an upper bound. By the completeness axiom, it has a supremum a. We prove that $(x_n) \to a$.

So let $\varepsilon > 0$. We need to find $N \in \mathbb{N}$ such that $|x_n - a| < \varepsilon$ for all $n \geqslant N$.

Since a is a supremum of $\{x_n \mid n \in \mathbb{N}\}$, there is some $N \in \mathbb{N}$ such that $x_N > a - \varepsilon$.

Since (x_n) is increasing, by (i) we have $x_N \leqslant x_n$ for all $n \geqslant N$. Moreover, since a is an upper bound of the sequence, we actually have $x_N \leqslant x_n \leqslant a$ for all $n \geqslant N$.

Putting this together, for all $n \geqslant N$, we have

$$\begin{aligned} |x_n - a| &= a - x_n && \text{since } x_n - a \leqslant 0 \\ &\leqslant a - x_N && \text{since } x_N \leqslant x_n \text{ for all } n \geqslant N \\ &< \varepsilon && \text{since } x_N > a - \varepsilon \end{aligned}$$

It follows that $(x_n) \to a$, as required. □

Exercise 9.2.49
Prove part (b) of the monotone convergence theorem (Theorem 9.2.48). That is, prove that if a sequence (x_n) is decreasing and has a lower bound, then it converges.

Example 9.2.50
The monotone convergence theorem can be used to show that many of the sequences that we have already seen converge, although it doesn't tell us what their limit is. For example, $\left(\frac{1}{n}\right)$ converges since it is a decreasing sequence that is bounded below by 0.

Example 9.2.51
Let (r_n) be our recurring example sequence from Examples 9.2.14, 9.2.36 and 9.2.46, defined by $r_n = \frac{2n}{n+1}$ for all $n \in \mathbb{N}$. We proved in Example 9.2.46 that (r_n) is increasing. Moreover, for all $n \in \mathbb{N}$ we have

$$r_n = \frac{2n}{n+1} < \frac{2(n+1)}{n+1} = 2$$

and so (r_n) is bounded above by 2. By the monotone convergence theorem, the sequence (r_n) converges. Unfortunately, the monotone convergence theorem does not tell us what the limit of (r_n) is, but we have already computed it twice!

Exercise 9.2.52
Use the monotone convergence theorem to prove that the sequence $\left(\frac{n!}{n^n}\right)$ converges.

Exercise 9.2.53
A sequence (x_n) is defined recursively by $x_0 = 0$ and $x_{n+1} = \sqrt{2+x_n}$ for all $n \geqslant 0$. That is,

$$x_n = \underbrace{\sqrt{2+\sqrt{2+\sqrt{\cdots+\sqrt{2}}}}}_{n \text{ '2's}}$$

Prove that (x_n) converges.

We now define the notion of a *subsequence* of a sequence. A subsequence of a sequence is just like a subset of a set, except we can only pick out terms in a sequence in the order they appear.

Definition 9.2.54
Let (x_n) be a sequence of real numbers. A **subsequence** of (x_n) is a sequence of the form $(x_{n_i})_{i\geqslant 0}$, where $n_i < n_j$ for all $0 \leqslant i < j$.

In Definition 9.2.54 we were careful to write $(x_{n_i})_{i\geqslant 0}$ rather than just (x_{n_i}), because we wanted to emphasise that the indexing variable is i, rather than n. This is good practice in any situation where confusion might arise over which variable is the indexing variable.

Example 9.2.55
Define a sequence (x_n) by $x_n = (-1)^n$ for all $n \geqslant 0$.

$$(x_n)_{n\geqslant 0} = (1,-1,1,-1,1,-1,\dots)$$

The subsequence (x_{2i}) is the constant sequence with value 1, since for each $i \geqslant 0$ we have $x_{2i} = (-1)^{2i} = 1$, and the subsequence (x_{2i+1}) is the constant sequence with value -1, since for each $i \geqslant 0$ we have $x_{2i+1} = (-1)^{2i+1} = -1$.

Theorem 9.2.56
Let (x_n) be a sequence, let $a \in \mathbb{R}$, and suppose $(x_n) \to a$. Then every subsequence of (x_n) converges to a.

Proof
Let $(x_{n_i})_{i\geqslant 0}$ be a subsequence of (x_n). We need to prove that $(x_{n_i}) \to a$ as $i \to \infty$. To this end, fix $\varepsilon > 0$. We need to find $I \geqslant 0$ such that $|x_{n_i} - a| < \varepsilon$ for all $i \geqslant I$.

Since $(x_n) \to a$ as $n \to \infty$, there exists some $N \geqslant 0$ such that $|x_n - a| < \varepsilon$ for all $n \geqslant N$. Let $I \geqslant 0$ be least such that $n_I \geqslant N$. We know that I exists since we have $0 \leqslant n_0 < n_1 < n_2 < \ldots$.

But then for all $i \geqslant I$, we have $n_i \geqslant n_I \geqslant N$, and hence $|x_{n_i} - a| < \varepsilon$ by definition of N.

Hence the subsequence (x_{n_i}) converges to a, as required. □

Exercise 9.2.57

Prove that a subsequence of an increasing sequence is increasing, that a subsequence of a decreasing sequence is decreasing, and that a subsequence of a constant sequence is constant.

We can use the monotone convergence theorem and the squeeze theorem to prove the following very powerful result, which is related to a notion in the field of topology known as *sequential compactness*.

Theorem 9.2.58 · *Bolzano–Weierstrass theorem*

Every bounded sequence of real numbers has a convergent subsequence.

Proof

Let (x_n) be a sequence of real numbers and let $a, b \in \mathbb{R}$ be such that $a < x_n < b$ for each $n \geqslant 0$—the numbers a and b exist since the sequence (x_n) is bounded.

Our strategy is as follows. The sequence (x_n) is entirely contained inside the interval $[a,b]$, which has length $\ell = b - a$. Letting $c = \frac{a+b}{2}$ be the (arithmetic) mean of a and b, we see that one of the intervals $[a,c]$ or $[c,b]$, or possibly both, must contain infinitely many terms of the sequence (x_n)—but then this defines a subsequence of (x_n) which is entirely contained inside a sub-interval of $[a,b]$ whose length is $\frac{\ell}{2}$. We iterate this process inductively, obtaining smaller and smaller intervals that contain infinitely many terms in the sequence (x_n). The end-points of these intervals are then bounded monotone sequences—the sequence of lower end-points is increasing, and the sequence of upper end-points is decreasing. The monotone convergence theorem implies that both sequences converge. We will prove that they converge to the same limit, thereby 'trapping' a subsequence of (x_n), which will converge by the squeeze theorem.

Now let's put our strategy into action. We will define the terms n_i, a_i and b_i by induction on i, and then verify that the resulting subsequence $(x_{n_i})_{i \geqslant 0}$ converges.

First, define $n_0 = 0$, $a_0 = a$ and $b_0 = b$.

Now fix $i \geqslant 0$ and suppose that the numbers n_i, a_i and b_i have been defined in such a way that:

(i) $x_{n_i} \in [a_i, b_i]$;

(ii) $x_n \in [a_i, b_i]$ for infinitely many $n > n_i$;

(iii) $a_j \leqslant a_i < b_i \leqslant b_j$ for all $j \leqslant i$; and

(iv) $b_i - a_i = \frac{\ell}{2^i}$.

Write $c_i = \frac{a_i+b_i}{2}$. By condition (ii), it must be case that infinitely many of the terms x_n, for $n > n_i$, are contained in either $[a_i, c_i]$ or in $[c_i, b_i]$. In the former case, define $a_{i+1} = a_i$ and $b_{i+1} = c_i$; and in the latter case define $a_{i+1} = c_i$ and $b_{i+1} = b_i$; and then define $n_{i+1} > n_i$ such that $x_{n_{i+1}} \in [a_{i+1}, b_{i+1}]$.

Note that conditions (i)–(iv) are satisfied, with i now replaced by $i+1$. Indeed, (i) and (ii) are satisfied by definition of a_{i+1}, b_{i+1} and n_{i+1}. Condition (iii) is satisfied since either $a_{i+1} = a_i$ or $a_{i+1} = \frac{a_i+c_i}{2} \geqslant a_i$, and likewise for b_{i+1}. Condition (iv) is satisfied since

$$c_i - a_i = \frac{a_i+b_i}{2} - a_i = \frac{b_i-a_i}{2} = \frac{\ell/2^i}{2} = \frac{\ell}{2^{i+1}}$$

and likewise $b_i - c_i = \frac{\ell}{2^{i+1}}$.

Since by construction we have $n_i < n_{i+1}$ for each $i \geqslant 0$, we have defined a subsequence $(x_{n_i})_{i\geqslant 0}$ of (x_n).

Now the sequence (a_i) is increasing and is bounded above by b, and the sequence (b_i) is decreasing and is bounded below by a. By the monotone convergence theorem $(a_i) \to a^\star$ and $(b_i) \to b^\star$ for some $a^\star, b^\star \in \mathbb{R}$. But moreover we have

$$\frac{\ell}{2^i} = b_i - a_i \to b^\star - a^\star$$

Since $\frac{\ell}{2^i} \to 0$, we have $b^\star - a^\star = 0$ by uniqueness of limits, and so $a^\star = b^\star$. Write $x^\star$ for the common value of $a^\star$ and $b^\star$.

Finally, we have $a_i \leqslant x_{n_i} \leqslant b_i$ for all $i \geqslant 0$, so that $x_{n_i} \to x^\star$ by the squeeze theorem. □

The Bolzano–Weierstrass theorem can be used to prove that a sequence converges by verifying that its terms get arbitrarily close together. Such sequences are called *Cauchy* sequences, and the fact that all Cauchy sequences converge is proved in Theorem 9.2.62.

Definition 9.2.59

A **Cauchy sequence** is a sequence (x_n) of real numbers such that, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|x_m - x_n| < \varepsilon$ for all $m, n \geqslant N$.

Example 9.2.60

Let (r_n) be our favourite recurring example sequence from Examples 9.2.14, 9.2.36, 9.2.46 and 9.2.51, defined by $r_n = \dfrac{2n}{n+1}$ for all $n \in \mathbb{N}$. We prove that (r_n) is Cauchy.

First note that, given $m, n \geqslant 1$, we have

$$|r_m - r_n| = \left|\frac{2m}{m+1} - \frac{2n}{n+1}\right| = \frac{2|m-n|}{(m+1)(n+1)} = \frac{2|\frac{1}{n} - \frac{1}{m}|}{(1+\frac{1}{m})(1+\frac{1}{n})}$$

Now fix $\varepsilon > 0$, and let $N \in \mathbb{N}$ be such that $\frac{1}{m} < \frac{\varepsilon}{2}$ and $\frac{1}{n} < \frac{\varepsilon}{2}$ for all $m, n > N$. Note that such a value of N exists by Example 9.2.13.

Now let $m,n \geqslant N$. Then $|\frac{1}{n} - \frac{1}{m}| < \frac{\varepsilon}{2}$ since both $\frac{1}{m}$ and $\frac{1}{n}$ are elements of $(0, \frac{\varepsilon}{2})$. Moreover $1 + \frac{1}{m} > 1$ and $1 + \frac{1}{n} > 1$. It follows that, for all $m,n \geqslant N$, we have

$$|r_m - r_n| < \frac{2 \cdot \frac{\varepsilon}{2}}{1 \cdot 1} = \varepsilon$$

Hence (r_n) is Cauchy, as claimed.

The following exercise generalises the previous example.

Exercise 9.2.61

Prove that every convergent sequence is a Cauchy sequence.

Theorem 9.2.62 · *Cauchy criterion*

Every Cauchy sequence of real numbers converges.

Proof

Let (x_n) be a Cauchy sequence of real numbers.

First note that (x_n) is bounded. To see this, note that by definition of Cauchy sequences, there is some $N \in \mathbb{N}$ such that $|x_m - x_n| < 1$ for all $m,n \geqslant N$. In particular, $|x_m - x_N| < 1$ for all $m \geqslant N$. This means that the sequence (x_n) is bounded below by

$$a = \min\{x_0, x_1, \ldots, x_{N-1}, x_N - 1\}$$

and is bounded above by

$$b = \max\{x_0, x_1, \ldots, x_{N-1}, x_N + 1\}$$

By the Bolzano–Weierstrass theorem (Theorem 9.2.58), the sequence (x_n) has a convergent subsequence (x_{n_i}). Let $x^\star = \lim_{i\to\infty}(x_{n_i})$. We prove that $(x_n) \to x^\star$.

So let $\varepsilon > 0$. Fix M sufficiently large that:

- $|x_{n_i} - x^\star| < \frac{\varepsilon}{3}$ for all $n_i \geqslant M$; and
- $|x_n - x_m| < \frac{\varepsilon}{3}$ for all $m,n \geqslant M$.

Such a value of M exists by convergence of (x_{n_i}) and the Cauchy property of (x_n).

Fix $n \geqslant M$, and let $i \in \mathbb{N}$ be arbitrary such that $n_i \geqslant M$. Then we have

$$\begin{aligned}
&|x_n - x^\star| \\
&= |(x_n - x_M) + (x_M - x_{n_i}) + (x_{n_i} - x^\star)| && \text{rearranging} \\
&\leqslant |x_n - x_M| + |x_M - x_{n_i}| + |x_{n_i} - x^\star| && \text{by the triangle inequality} \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} && \text{by the above properties} \\
&= \varepsilon
\end{aligned}$$

Hence $(x_n) \to x^\star$, as required. □

Section 9.3

Series and sums

A *series* can be thought of as the result of adding up all of the terms in a sequence. The uses of series inside and outside of mathematics is widespread, particularly in analysis and statistics. In fact, we will use series repeatedly when we study probability theory in Chapter 11!

Unfortunately the definition of a series is not quite as simple as 'the result of adding up all of the terms in a sequence'. For a start, we haven't defined what means to add up infinitely many numbers, and sometimes this might not even be possible—for example, you might encounter problems if you try adding up all of the terms in the constant sequence $(1,1,1,\dots)$.

The definition of a series, then, is that of a *formal* sum (see Definition 9.3.1). The word 'formal' here means that it is an expression that *represents* a sum, but is not actually evaluated. So for example

$$1+1+1+\cdots$$

is a series.

We will then separately define what it means for it to be possible to evaluate an infinite sum represented by a series (Definition 9.3.3); this definition implies that the series $1+1+1+\cdots$ is not summable, for example.

Definition 9.3.1

A (**real**) **series** is a formal sum of a sequence $(a_n)_{n\geqslant 0}$, denoted by

$$\sum_{n\geqslant 0} a_n \quad (\text{LaTeX code: } \texttt{\textbackslash sum\_\{n \textbackslash ge 0\}})$$

or by $\sum_{n=0}^{\infty} a_n$ (LaTeX code: `\sum_{n=0}^{\infty}`), or even by $a_0+a_1+a_2+\cdots$.

As with sequences, it is possible for a series to be indexed from a different starting number, like in the next example.

Example 9.3.2

The sequence $(\frac{1}{k})_{k\geqslant 1}$ defines the series

$$\sum_{k\geqslant 1} \frac{1}{k} = \frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots$$

We will soon see that this series *diverges* (Theorem 9.3.28), since adding these terms together one by one yields unboundedly larger and larger real numbers.

Series in isolation are not particularly useful or interesting. They become so by defining what it means to evaluate them—at least, when it is possible to do so.

Definition 9.3.3

Let $N \in \mathbb{N}$. The N^{th} **partial sum** of a series $\sum_{n\geqslant 0} a_n$ is the real number $s_N = \sum_{n=0}^{N} a_n$.

We say that the series **converges** if the sequence of partial sums $(s_N)_{N\geqslant 0}$ converges; in this case, the **sum** of the series is the real number $\lim_{N\to\infty}(s_N)$, also written $\sum_{n\geqslant 0} a_n$.

If the sequence of partial sums $(s_N)_{N\geqslant 0}$ diverges, then we say the series **diverges**.

Example 9.3.4

Consider the series

$$S = \sum_{n\geqslant 2} \binom{n}{2}^{-1}$$

We prove that S converges and its sum is 2.

To see this, note that for all $n \geqslant 2$, we have by Theorem 4.2.14 that

$$\binom{n}{2}^{-1} = \left(\frac{n(n-1)}{2}\right)^{-1} = \frac{2}{n(n-1)} = \frac{2}{n-1} - \frac{2}{n}$$

Therefore, for all $N \geqslant 2$, the N^{th} partial sum of S is given by

$$s_N = \sum_{n=2}^{N} \binom{n}{2}^{-1} = \left(\frac{2}{1} - \frac{2}{2}\right) + \left(\frac{2}{2} - \frac{2}{3}\right) + \cdots + \left(\frac{2}{N-1} - \frac{2}{N}\right) = 2 - \frac{2}{N}$$

It follows that $S = \lim_{N\to\infty}(s_N) = 2$, as required.

Exercise 9.3.5

Prove that the series $\sum_{n\geqslant 3} \binom{n}{3}^{-1}$ converges, and find its sum.

Example 9.3.6

We prove that the series $\sum_{n\geqslant 0} 1$ diverges. Indeed, for all $N \in \mathbb{N}$, we have

$$\sum_{n=0}^{N} = \underbrace{1+1+\cdots+1}_{N+1 \text{ times}} = N+1$$

Thus the sequence of partial sums is unbounded, so does not converge to a real number.

Exercise 9.3.7

Prove that the series $\sum_{n\geqslant 0} (-1)^n$ diverges.

The underlying reason why the series in Example 9.3.6 and Exercise 9.3.7 diverge is that their terms do not get smaller and smaller. We will prove in Theorem 9.3.23 that in order for a series to converge, its terms must tend to zero.

Theorem 9.3.8 · *Sum of geometric series*
Let $r \in (-1,1)$. Then $\displaystyle\sum_{n\geqslant 0} r^n = \frac{1}{1-r}$.

Proof
Given $N \in \mathbb{N}$, the N^{th} partial sum s_N of the series is given by by

$$s_N = \sum_{n=0}^{N} r^n = 1 + r + r^2 + \cdots + r^N$$

Note that

$$rs_N = \sum_{n=0}^{n} r^{n+1} = r + r^2 + \cdots + r^{N+1} = s_{N+1} - 1$$

and hence

$$(1-r)s_N = s_N - rs_N = s_N - (s_{N+1} - 1) = 1 - (s_{N+1} - s_N) = 1 - r^{N+1}$$

and hence dividing by $1-r$, which is permissible since $r \neq 1$, yields

$$s_N = \frac{1 - r^{N+1}}{1-r}$$

Since $|r| < 1$, we have $(r^{N+1}) \to 0$ by Proposition 9.2.21, and so

$$\sum_{n\geqslant 0} r^n = \lim_{N\to\infty} \frac{1-r^{N+1}}{1-r} = \frac{1-0}{1-r} = \frac{1}{1-r}$$

as claimed. □

Exercise 9.3.9
Prove that the series $\displaystyle\sum_{n\geqslant 0} r^n$ diverges for all $r \in \mathbb{R} \setminus (-1,1)$.

The next result allows us to add two series together by adding their terms, and to multiply a series by a constant by multiplying their terms by the constant.

Theorem 9.3.10 · *Linearity of summation*
Let $\displaystyle\sum_{n\geqslant 0} a_n$ and $\displaystyle\sum_{n\geqslant 0} b_n$ be convergent series. Then

(a) The series $\displaystyle\sum_{n\geqslant 0} (a_n + b_n)$ is convergent, and its sum is $\displaystyle\sum_{n\geqslant 0} a_n + \sum_{n\geqslant 0} b_n$;

(b) For all $c \in \mathbb{R}$, the series $\displaystyle\sum_{n\geqslant 0} ca_n$ is convergent, and its sum is $\displaystyle c\sum_{n\geqslant 0} a_n$.

Proof of (a)
For (a), note that the partial sums of $\sum_{n\geqslant 0}(a_n+b_n)$ are given by

$$\sum_{n=0}^{N}(a_n+b_n)=\sum_{n=0}^{N}a_n+\sum_{n=0}^{N}b_n$$

so we may apply Theorem 9.2.34(a) to obtain

$$\begin{aligned}
&\sum_{n\geqslant 0}(a_n+b_n)\\
&=\lim_{N\to\infty}\left(\sum_{n=0}^{N}(a_n+b_n)\right)\\
&=\lim_{N\to\infty}\left(\sum_{n=0}^{N}a_n+\sum_{n=0}^{N}b_n\right)\\
&=\lim_{N\to\infty}\left(\sum_{n=0}^{N}a_n\right)+\lim_{N\to\infty}\left(\sum_{n=0}^{N}b_n\right)\\
&=\sum_{n\geqslant 0}a_n+\sum_{n\geqslant 0}b_n
\end{aligned}$$

as required. □

Exercise 9.3.11
Prove part (b) of Theorem 9.3.10, and deduce that if $\sum_{n\geqslant 0}a_n$ and $\sum_{n\geqslant 0}b_n$ are convergent series, then $\sum_{n\geqslant 0}(a_n-b_n)$ converges, and its sum is equal to $\sum_{n\geqslant 0}a_n-\sum_{n\geqslant 0}b_n$.

Expansions of real numbers in number bases

We now take a brief detour away from the general theory of series to discuss an application, namely expansions of real numbers in number bases.

You are likely familiar with decimal expansions of real numbers, for example

$$\frac{1}{2}=0.5,\quad \frac{1}{7}=0.142857142857\ldots,\quad \sqrt{2}=1.414213562373\ldots$$

A decimal expansion is really a series in disguise. For example

$$0.142857\ldots=1\cdot\frac{1}{10}+4\cdot\frac{1}{10^2}+2\cdot\frac{1}{10^3}+8\cdot\frac{1}{10^4}+5\cdot\frac{1}{10^5}+7\cdot\frac{1}{10^6}+\cdots$$

Thus when we write out a decimal expansion of a (non-negative, say) real number x as $x=x_0.x_1x_2x_3\ldots$, what we are really saying is that

$$x=x_0+\sum_{i\geqslant 1}x_i\cdot 10^{-i}$$

where $x_0 \in \mathbb{N}$ and $x_i \in \{0,1,2,3,4,5,6,7,8,9\}$ for all $i \geqslant 1$.

We can apply this to other number bases. For example, the binary (base-2) expansion of $\frac{1}{2}$ is 0.1 since $\frac{1}{2} = 1 \cdot 2^{-1}$, and the binary expansion of $\frac{1}{3}$ is $0.010101\ldots$ since

$$0.010101\ldots_{(2)} = \sum_{n \geqslant 1} \frac{1}{2^{2n}} = \frac{1}{4} \sum_{k \geqslant 0} \frac{1}{4^k} = \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}$$

Our goal is to give a precise definition of the base-b expansion of a real number for an arbitrary number base $b > 1$.

In order to do this, we must prove that they are well-defined—that is, every real number has a unique base-b expansion. But wait! It's not true! Expansions are not *quite* unique—for example, we have

$$0.999\ldots = \sum_{i \geqslant 1} 9 \cdot 10^{-i} = \frac{9}{10} \sum_{k \geqslant 0} \frac{1}{10^k} = \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = 1 = 1.000\ldots$$

Conveniently, the only problem with uniqueness occurs with recurring 0s and recurring 9s: every number that has a decimal expansion with recurring 9s also has one with recurring 0s. More generally, in a base $b > 0$, recurring '$b-1$'s may be replaced by recurring 0s.

We are now ready to proceed.

Theorem 9.3.12

Let $b \in \mathbb{N}$ with $b > 1$. For all $x \in [0, \infty)$, there is a unique series $\sum_{i \geqslant 0} \frac{x_i}{b^i}$, such that:

(i) $x_0 \in \mathbb{N}$, and $x_i \in \{0, 1, \ldots, b-1\}$ for all $i \geqslant 1$;

(ii) The series $\sum_{i \geqslant 0} \frac{x_i}{b^i}$ converges, and its sum is equal to x; and

(iii) The sequence (x_i) is not eventually equal to $b-1$.

Proof of existence

Fix $x \geqslant 0$. Define integers x_i and real numbers $y_i \in [0,1)$ for $i \in \mathbb{N}$ recursively as follows:

- Let $x_0 \in \mathbb{N}$ be such that $x_0 \leqslant x < x_0 + 1$, and let $y_0 = x - x_0$—note that $0 \leqslant y_0 < 1$, as required.
- Given $i \in \mathbb{N}$, let $x_{i+1} \in \mathbb{Z}$ such that $x_{i+1} \leqslant b y_i < x_{i+1} + 1$, and let $y_{i+1} = b y_i - x_{i+1}$—note $0 \leqslant y_{i+1} < 1$ as required.

For all $n \in \mathbb{N}$ we have

$$x - \sum_{i=0}^{n} \frac{x_i}{b^i} = \frac{y_n}{b^n}$$

We can prove this by induction:

- (**Base case**) We have $x - \frac{x_0}{b^0} = x - x_0 = y_0 = \frac{y_0}{b^0}$ by construction.

- (**Induction step**) Fix $n \geqslant 0$ and suppose that $x - \sum_{i=0}^{n} \frac{x_i}{b^i} = \frac{y_n}{b^n}$. Then

$$x - \sum_{i=0}^{n+1} \frac{x_i}{b^i} = \frac{y_n}{b^n} - \frac{x_{n+1}}{b^{n+1}} = \frac{by_n - x_{n+1}}{b^{n+1}} = \frac{y_{n+1}}{b^{n+1}}$$

  as required.

We now verify the conditions in the statement of the theorem.

- Condition (i) is satisfied by construction of the sequence (x_i): indeed, we defined x_i to be integers for all $i \in \mathbb{N}$, and if $i \geqslant 1$ then the fact that $x_i \in \{0, 1, \ldots, b-1\}$ follows from the facts that $x_i \leqslant by_{i-1} < x_i + 1$ and $y_i \in [0, 1)$.
- To see that (ii) holds, note that for all $n \in \mathbb{N}$ we have

$$0 \leqslant x - \sum_{i=0}^{n} \frac{x_i}{b^i} = \frac{y_n}{b^n} < \frac{1}{b^n}$$

  But $(\frac{1}{b^n}) \to 0$ by Proposition 9.2.21, and so $\sum_{i=0}^{n} \frac{x_i}{b^i}$ converges to x.
- To prove (iii), suppose there is some $n \in \mathbb{N}$ such that $x_i = b - 1$ for all $i > n$. Then

$$\begin{aligned}
y_n &= b^n \left(x - \sum_{i=0}^{n} \frac{x_i}{b^i} \right) && \text{as we proved above} \\
&= b^n \left(\sum_{i \geqslant 0} \frac{x_i}{b^i} - \sum_{i=0}^{n} \frac{x_i}{b^i} \right) && \text{by condition (ii)} \\
&= b^n \sum_{i \geqslant n+1} \frac{x_i}{b^i} && \text{simplifying} \\
&= b^n \sum_{i \geqslant n+1} \frac{b-1}{b^i} && \text{since } x_i = b-1 \text{ for all } i > n \\
&= b^n \sum_{j \geqslant 0} \frac{b-1}{b^{j+n+1}} && \text{substituting } j = i - n - 1 \\
&= b^n \cdot \frac{b-1}{b^{n+1}} \sum_{j \geqslant 0} \frac{1}{b^j} && \text{rearranging} \\
&= b^n \cdot \frac{b-1}{b^{n+1}} \cdot \frac{1}{1 - \frac{1}{b}} && \text{by Theorem 9.3.8} \\
&= 1 && \text{simplifying}
\end{aligned}$$

  But this contradicts the fact that $y_n \in [0, 1)$. So we do indeed have that (x_i) is not eventually equal to $b - 1$.

This completes the proof of existence. □

Exercise 9.3.13

Prove the 'uniqueness' part of Theorem 9.3.12—that is, prove that for all $x \in [0, \infty)$, if

$$\sum_{i \geqslant 0} \frac{u_i}{b^i} \quad \text{and} \quad \sum_{i \geqslant 0} \frac{v_i}{b^i}$$

are two series satisfying conditions (i)–(iii) of Theorem 9.3.12, then $u_i = v_i$ for all $i \in \mathbb{N}$.

Theorem 9.3.12 justifies the following definition.

Definition 9.3.14

Let $b > 1$. The **base-b expansion** of a real number x is the unique signed series $\pm \sum_{i \geqslant 0} \frac{x_i}{b^i}$ such that:

(i) $x_0 \in \mathbb{N}$, and $x_i \in \{0, 1, \ldots, b-1\}$ for all $i \geqslant 1$;

(ii) The series $\sum_{i \geqslant 0} \frac{x_i}{b^i}$ converges, and its sum is equal to $|x|$; and

(iii) The sequence (x_i) is not eventually equal to $b-1$.

To denote the fact that this is the base-b expansion of x, we may also write

$$x = \pm x_0 . x_1 x_2 x_3 \ldots_{(b)} \quad \text{or} \quad x = \pm d_r d_{r-1} \ldots d_1 d_0 . x_1 x_2 \ldots_{(b)}$$

where $d_r d_{r-1} \ldots d_1 d_0$ is the base-b expansion of $|x_0|$ (as in Definition 0.31), and $\pm$ is the sign of x (positive or negative).

The recursive definition of the sequence (x_i) in the proof of Theorem 9.3.12 yields an algorithm for computing base-b expansions.

Strategy 9.3.15 · *Finding base-b expansions of real numbers*

In order to find the base-b expansion of a real number x:

- Let $x_0 \in \mathbb{N}$ be such that $x_0 \leqslant |x| < x_0 + 1$, and define $y_0 = |x| - x_0$.
- For $n \in \mathbb{N}$, given x_n and y_n, let $x_{n+1} \in \mathbb{N}$ be such that $x_{n+1} \leqslant b y_n < x_{n+1} + 1$, and let $y_{n+1} = b y_n - x_{n+1}$. [Note that the value of x_{n+1} depends only on the value of y_n, not on x_n.]

Then $x = \pm x_0 . x_1 x_2 x_3 \ldots_{(b)}$, where $\pm$ is '+' if $x \geqslant 0$, and $-$ if $x < 0$.

Example 9.3.16

Let's find the decimal expansion of $\frac{1}{3}$.

- $0 \leqslant \frac{1}{3} < 1$, so $x_0 = 0$ and $y_0 = \frac{1}{3} - 0 = \frac{1}{3}$.
- $3 \leqslant 10 \cdot \frac{1}{3} < 4$, so $x_1 = 3$ and $y_1 = \frac{10}{3} - 3 = \frac{1}{3}$.
- $3 \leqslant 10 \cdot \frac{1}{3} < 4$, so $x_2 = 3$ and $y_2 = \frac{10}{3} - 3 = \frac{1}{3}$.

- ... evidently, this pattern repeats. (In fact, this can be proved by induction!)

So $\frac{1}{3} = 0.33333\ldots$.

Example 9.3.17

Now let's find the decimal expansion of $\frac{1}{7}$.

- $0 \leqslant \frac{1}{7} < 1$, so $x_0 = 0$ and $y_0 = \frac{1}{7} - 0 = \frac{1}{7}$.
- $1 \leqslant 10 \cdot \frac{1}{7} < 2$, so $x_1 = 1$ and $y_1 = \frac{10}{7} - 1 = \frac{3}{7}$.
- $4 \leqslant 10 \cdot \frac{3}{7} < 5$, so $x_2 = 4$ and $y_2 = \frac{30}{7} - 4 = \frac{2}{7}$.
- $2 \leqslant 10 \cdot \frac{2}{7} < 3$, so $x_3 = 2$ and $y_3 = \frac{20}{7} - 2 = \frac{6}{7}$.
- $8 \leqslant 10 \cdot \frac{6}{7} < 9$, so $x_4 = 8$ and $y_4 = \frac{60}{7} - 8 = \frac{4}{7}$.
- $5 \leqslant 10 \cdot \frac{4}{7} < 6$, so $x_5 = 5$ and $y_5 = \frac{40}{7} - 5 = \frac{5}{7}$.
- $7 \leqslant 10 \cdot \frac{5}{7} < 8$, so $x_6 = 7$ and $y_6 = \frac{50}{7} - 7 = \frac{1}{7}$.
- ... and now it repeats with the same pattern, since $y_6 = y_0$.

So $\frac{1}{7} = 0.142857142857\ldots$.

Exercise 9.3.18

Use Strategy 9.3.15 to find the decimal expansion of $\frac{1}{6}$.

Exercise 9.3.19

Use Strategy 9.3.15 to find the *binary* expansion of $\frac{1}{7}$.

Exercise 9.3.20

Prove that between any two distinct real numbers, there is a rational number.

Decimal expansions can be used to prove that $\mathbb{R}$ is uncountable using Cantor's diagonal argument (Strategy 6.2.19).

Theorem 9.3.21

$\mathbb{R}$ is uncountable.

Proof

Let $f : \mathbb{N} \to \mathbb{R}$ be an arbitrary function. We prove that f is not surjective.

For each $n \in \mathbb{N}$, let the decimal expansion of $f(n) \in \mathbb{R}$ be $x_0^n . x_1^n x_2^n x_3^n \ldots$—that is

$$f(n) = \sum_{i \geqslant 0} x_i^n \cdot 10^{-i}$$

[The superscript 'n' here is a label, not an exponent.]

We define $b \in \mathbb{R}$ by forcing the decimal expansions of b and $f(n)$ to differ in the n^{th} place for each $n \in \mathbb{N}$. Specifically, for each $n \in \mathbb{N}$ let

$$b_n = \begin{cases} 0 & \text{if } x_n^n \neq 0 \\ 1 & \text{if } x_n^n = 0 \end{cases}$$

and let $b = b_0.b_1b_2b_3\ldots = \sum_{i \geqslant 0} b_i \cdot 10^{-i}$.

Note that b does not have recurring 9s, so it is a valid decimal expansion of a real number. Moreover $b \neq f(n)$ for any $n \in \mathbb{N}$ by uniqueness of decimal expansions of real numbers (Theorem 9.3.12), since $b_n \neq x_n^n$.

Hence f is not surjective, so that $\mathbb{R}$ is uncountable. □

Now let's return to learning about series in the abstract.

Tests for convergence and divergence

Sometimes all we need to know about a series is whether it converges or diverges. In such cases, it can be very tricky to find an exact value for the sum of the series. We now develop some techniques for determining whether or not a series converges.

Theorem 9.3.22 · *Cauchy's convergence test*
A series $\sum_{n \geqslant 0} a_n$ converges if and only if, for all $\varepsilon > 0$, there is some $K \in \mathbb{N}$ such that

$$\left| \sum_{n=n_0}^{n_1} a_n \right| < \varepsilon \text{ for all } n_1 \geqslant n_0 \geqslant K$$

Proof
Let $(s_N)_{N \geqslant 0}$ be the sequence of partial sums of the series. By Theorem 9.2.62, we know that $\sum_{n \geqslant 0} a_n$ is convergent if and only if $(s_N)_{N \geqslant 0}$ is a Cauchy sequence.

But the assertion that $(s_N)_{N \geqslant 0}$ is Cauchy is equivalent to the condition in the statement of this theorem: note that we may replace K by any larger value (in particular, we may assume $K \geqslant 1$), and so

$$\sum_{n=n_0}^{n_1} a_n = s_{n_1} - s_{n_0 - 1}$$

as required. □

We will use Cauchy's convergence test frequently in our proofs. One example of how Cauchy's

convergence test can be put to work is the *term test*, which is very useful for proving that a series *diverges*—in fact, it instantly implies that the series in Example 9.3.6 and Exercise 9.3.7 diverge.

Theorem 9.3.23 · *Term test*
Let $S = \sum_{n\geqslant 0} a_n$ be a series. If S converges, then $(a_n) \to 0$.

Proof
Fix $\varepsilon > 0$. Since S converges, by Cauchy's convergence test (Theorem 9.3.22) there exists $K \in \mathbb{N}$ such that $\left|\sum_{n=n_0}^{n_1} a_n\right| < \varepsilon$ for all $n_1 \geqslant n_0 \geqslant K$.

But then for all $k \geqslant N$ we have $k+1 \geqslant k \geqslant K$, and so

$$|a_k| = \left|\sum_{n=k}^{k+1} a_n\right| < \varepsilon$$

so that $(a_k) \to 0$, as required. □

Example 9.3.24
The series $\sum_{n\geqslant 0} n$ diverges since the sequence (n) does not tend to zero.

Exercise 9.3.25
Let $a \in \mathbb{R}$. Prove that the series $\sum_{n\geqslant 0} a$ converges if and only if $a = 0$.

Theorem 9.3.26 · *Comparison test*
Let $\sum_{n\geqslant 0} a_n$ and $\sum_{n\geqslant 0} b_n$ be series.

(a) If $\sum_{n\geqslant 0} a_n$ converges and eventually $0 \leqslant b_n \leqslant a_n$, then $\sum_{n\geqslant 0} b_n$ converges; and

(b) If $\sum_{n\geqslant 0} a_n$ diverges and eventually $0 \leqslant a_n \leqslant b_n$, then $\sum_{n\geqslant 0} b_n$ diverges.

Proof of (a)
Suppose $\sum_{n\geqslant 0} a_n$ converges and its sum is equal to A. Let $N \in \mathbb{N}$ be sufficiently large that $0 \leqslant b_n \leqslant a_n$ for all $n \geqslant N$.

For all $L \geqslant K \geqslant N$ we have

$$\sum_{n=0}^{L} b_n = \sum_{n=0}^{K} b_n + \underbrace{\sum_{n=K+1}^{L} b_n}_{\geqslant 0} \geqslant \sum_{n=0}^{K} b_n$$

and so the sequence of partial sums of $\sum_{n\geqslant 0} b_n$ is eventually increasing.

Also for all $M \geqslant N$ we have

$$\sum_{n=0}^{M} b_n = \sum_{n=0}^{N} b_n + \sum_{n=N+1}^{M} b_n \leqslant \sum_{n=0}^{N} b_n + \sum_{n=N+1}^{M} a_n = \sum_{n=0}^{N} (b_n - a_n) + \sum_{n=0}^{M} a_n$$

Moreover $a_n \geqslant 0$ for all $n > M$, and so we have

$$\sum_{n=0}^{M} b_n \leqslant \sum_{n=0}^{N} (b_n - a_n) + \sum_{n\geqslant 0} a_n$$

Thus the sequence of partial sums of $\sum_{n\geqslant 0} b_n$ is eventually bounded above.

By the monotone convergence theorem (Theorem 9.2.48), the sequence of partial sums of $\sum_{n\geqslant 0} b_n$ converges, hence so does the series. □

Exercise 9.3.27
Prove part (b) of Theorem 9.3.26.

The next result is a nice example of an *indirect* use of the term test (Theorem 9.3.23): although the terms in the series $\sum_{n\geqslant 1} \frac{1}{n}$ converge to zero, we can manipulate it to bound it below by a series whose terms are unbounded.

Theorem 9.3.28 · *Divergence of the harmonic series*
The series $\sum_{n\geqslant 1} \frac{1}{n}$ diverges.

Proof
By rounding up denominators to the next power of 2, we get

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \cdots \geqslant \frac{1}{1} + \frac{1}{2} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{=1/2} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{=1/2} + \cdots$$

This diverges since we're adding $\frac{1}{2}$ infinitely many times.

More precisely, define a sequence $(a_n)_{n\geqslant 1}$ by letting $a_n = \frac{1}{2^k}$ for the least $k \in \mathbb{N}$ with $\frac{1}{2^k} \leqslant \frac{1}{n}$. Then $0 \leqslant a_n \leqslant \frac{1}{n}$ for each $n \in \mathbb{N}$.

Now note that

$$\sum_{n\geqslant 1} a_n = 1 + \sum_{k\geqslant 0} \sum_{n=2^k+1}^{2^{k+1}} \frac{1}{2^{k+1}} = 1 + \sum_{k\geqslant 0} (2^{k+1} - 2^k) \cdot \frac{1}{2^{k+1}} = 1 + \sum_{k\geqslant 0} \frac{1}{2}$$

which diverges by the term test since $(\frac{1}{2}) \nrightarrow 0$.

Thus $\sum_{r\geqslant 0} \frac{1}{r}$ diverges by the comparison test. □

Exercise 9.3.29
Prove that $\sum_{n\geqslant 1} n^{-r}$ diverges for all real $r \geqslant 1$.

Theorem 9.3.30 · *Alternating series test*
Let (a_n) be a sequence such that $(a_n) \to 0$ and $a_n \geqslant 0$ for all $n \in \mathbb{N}$. If (a_n) is decreasing—that is, if $a_m \geqslant a_n$ for all $m, n \in \mathbb{N}$ with $m \leqslant n$—then the series $\sum_{n\geqslant 0} (-1)^n a_n$ converges.

Proof
Define sequences (e_N) and (o_N) by $e_N = s_{2N}$ and $o_N = s_{2N+1}$ for all $n \in \mathbb{N}$. That is, (e_N) is the sequence of *even* partial sums, and (o_N) is the sequence of *odd* partial sums.

Then for all $N \in \mathbb{N}$, we have

$$e_{N+1} = s_{2N+2} = s_{2N} - a_{2N+1} + a_{2N+2} = e_N - \underbrace{(a_{2N+1} - a_{2N+2})}_{\geqslant 0} \leqslant e_N$$

so that (e_N) is a decreasing sequence, and

$$e_N = a_0 - a_1 + \sum_{k=1}^{N-1} \underbrace{(a_{2k} - a_{2k+1})}_{\geqslant 0} + \underbrace{a_{2N}}_{\geqslant 0} \geqslant a_0 - a_1$$

so that (e_N) is bounded below. By the monotone convergence theorem (Theorem 9.2.48), the sequence (e_N) converges.

Likewise, for all $N \in \mathbb{N}$, we have

$$o_{N+1} = s_{2N+3} = s_{2N+1} + a_{2N+2} - a_{2N+1} = o_N + \underbrace{(a_{2N+2} - a_{2N+1})}_{\geqslant 0} \geqslant o_N$$

so that (o_N) is an increasing sequence, and

$$o_N = a_0 + \sum_{k=0}^{N} \underbrace{(a_{2k+2} - a_{2k+1})}_{\leqslant 0} - \underbrace{a_{2N+2}}_{\geqslant 0} \leqslant a_0$$

so that (o_N) is bounded above. By the monotone convergence theorem again, the sequence (o_N) converges.

Moreover for all $N \in \mathbb{N}$ we have

$$o_N - e_N = a_{2N+1} \geqslant 0 \quad \text{and} \quad e_{N+1} - o_N = a_{2N+2} \geqslant 0$$

so that $e_N \leqslant o_N \leqslant e_{N+1}$ for all $N \in \mathbb{N}$.

But then by the squeeze theorem (Theorem 9.2.38), (e_N) and (o_N) converge to the same limit $A \in \mathbb{R}$.

Finally, let $\varepsilon > 0$ and let $K \in \mathbb{N}$ be sufficiently large that $|e_M - A| < \varepsilon$ and $|o_M - A| < \varepsilon$ for all $M \geqslant K$. Then given $N \geqslant 2K$, if N is even then

$$|s_N - A| = |e_{\frac{N}{2}} - A| < \varepsilon$$

and if N is odd then

$$|s_N - A| = |o_{\frac{N-1}{2}} - A| < \varepsilon$$

as required. So $(s_N) \to A$, and so the series converges. □

Example 9.3.31

The sequence $(\frac{1}{n})$ is positive and decreasing, so the series

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

converges.

Exercise 9.3.32

Prove that if (a_n) is a sequence such that $(a_n) \to 0$ and $a_n \leqslant 0$ for all $n \in \mathbb{N}$. Prove that if (a_n) is an increasing sequence, then the series $\sum_{n \geqslant 0} (-1)^n a_n$ converges.

Exercise 9.3.33

Find a decreasing sequence (a_n) of non-negative real numbers such that $\sum_{n \geqslant 0} (-1)^n a_n$ diverges.

Absolute convergence

Definition 9.3.34

A series $\sum_{n \geqslant 0} a_n$ **converges absolutely** if the series $\sum_{n \geqslant 0} |a_n|$ converges.

Example 9.3.35

For all $r \in (-1, 1)$, the geometric series $\sum_{n \geqslant 0} r^n$ converges absolutely. Indeed, for all $r \in (-1, 1)$ we have $|r| \in (-1, 1)$ as well, and so $\sum_{n \geqslant 0} |r|^n$ converges by Theorem 9.3.8.

Example 9.3.36
Let $\sum_{n\geqslant 0} a_n$ be a convergent series.

If $a_n \geqslant 0$ for all $n \in \mathbb{N}$, then the series absolutely, since $|a_n| = a_n$ for all $n \in \mathbb{N}$.

Likewise, if $a_n \leqslant 0$ for all $n \in \mathbb{N}$, then

$$\sum_{n\geqslant 0} |a_n| \;=\; \sum_{n\geqslant 0} (-a_n) \;=\; -\sum_{n\geqslant 0} a_n$$

by linearity of summation, and so again the series converges absolutely.

Exercise 9.3.37
Find a series that converges, but does not converge absolutely.

Exercise 9.3.38
Prove Theorem 9.3.10 with 'convergent' replaced by 'absolutely convergent' throughout.

Absolutely convergent series enjoy some properties that are not enjoyed by series that converge but not absolutely—for example, they do not depend on what order you choose to add up their terms. We will prove this in Theorem 9.3.43.

The *ratio test* is useful for proving that a series converges absolutely.

Theorem 9.3.39 · *Ratio test*
Let (a_n) be a sequence of real numbers, and suppose that $\left(\left|\frac{a_{n+1}}{a_n}\right|\right) \to \ell \geqslant 0$.

(a) If $\ell < 1$, then $\sum_{n\geqslant 0} a_n$ converges absolutely.

(b) If $\ell > 1$, then $\sum_{n\geqslant 0} a_n$ diverges.

Proof of (a)
Assume $\ell < 1$, and pick ε with $0 < \varepsilon < 1 - \ell$. Define $r = \ell + \varepsilon$ and note that $0 < r < 1$.

Since $\left(\left|\frac{a_{n+1}}{a_n}\right|\right) \to \ell$, there exists $N \in \mathbb{N}$ such that $\left|\frac{a_{n+1}}{a_n} - \ell\right| < \varepsilon$ for all $n \geqslant N$. But then

$$0 \leqslant \left|\frac{a_{n+1}}{a_n}\right| < \ell + \varepsilon = r$$

Note that for all $n \geqslant N$ we have

$$|a_n| = |a_N| \times \left|\frac{a_{N+1}}{a_N}\right| \times \cdots \times \frac{a_n}{a_{n-1}} < |a_N| r^{n-N}$$

The series $\sum_{n=N}^{\infty} r^{n-N}$ converges by Theorem 9.3.8 since $r \in (-1,1)$, and so the series $\sum_{n=0}^{M} |a_n|$ converges by the comparison test, as required. □

Exercise 9.3.40

Prove part (b) of Theorem 9.3.39.

Example 9.3.41

Let $r \in \mathbb{R}$ and consider the series $\sum_{n\geqslant 1} \frac{r^n}{n}$. Then

$$\left| \frac{\frac{r^{n+1}}{n+1}}{\frac{r^n}{n}} \right| = \frac{n}{n+1}|r| \to |r|$$

By the ratio test, if $|r| < 1$ then the series converges absolutely, and if $|r| > 1$ then the series diverges.

The ratio test tells us notihng about what happens when $|r| = 1$. However, we already know: we proved in Theorem 9.3.28 that this series diverges when $r = 1$, and in Example 9.3.31 that it converges when $r = -1$.

Exercise 9.3.42

Use the ratio test to prove that the series $\sum_{n\geqslant 0} \frac{x^n}{n!}$ converges for all $x \in \mathbb{R}$.

Theorem 9.3.43 · *Absolutely convergent series can be reordered*

Let $\sum_{n\geqslant 0} a_n$ be an absolutely convergent series and let $\sigma : \mathbb{N} \to \mathbb{N}$ be a bijection. Then the series $\sum_{n\geqslant 0} a_{\sigma(n)}$ converges absolutely, and $\sum_{n\geqslant 0} a_{\sigma(n)} = \sum_{n\geqslant 0} a_n$.

Proof

Write $A = \sum_{n\geqslant 0} a_n$. In order to prove $\sum_{n\geqslant 0} a_{\sigma(n)} = A$, we need to prove that for all $\varepsilon > 0$, there is some $N \in \mathbb{N}$ such that $\left| \sum_{n=0}^{K} a_{\sigma(n)} - A \right| < \varepsilon$ for all $K \geqslant N$.

So let $\varepsilon > 0$. Then:

(1) Since $\sum_{n\geqslant 0} a_n = A$, there is some $M_1 \in \mathbb{N}$ such that

$$\left| \sum_{n=0}^{L} a_n - A \right| < \frac{\varepsilon}{2}$$

for all $L \geqslant M_1$.

(2) Since $\sum_{n\geqslant 0} a_n$ converges absolutely, it follows from Cauchy's convergence test (Theorem 9.3.22) there is some $M_2 \in \mathbb{N}$ such that

$$\sum_{n=n_0}^{n_1} |a_n| < \frac{\varepsilon}{2}$$

for all $n_1 \geqslant n_0 \geqslant M_2$.

Let M be the greater of M_1 and M_2.

Now let $N \in \mathbb{N}$ be such that $\{0,1,\ldots,M\} \subseteq \{\sigma(0),\sigma(1),\ldots,\sigma(N)\}$. This ensures that the terms $a_0,a_1,\ldots,a_M$ appear amongst the terms $a_{\sigma(0)},a_{\sigma(1)},\ldots,a_{\sigma(N)}$. Such a value exists since σ is a bijection; for example, we can take N to be the greatest value of $\sigma^{-1}(i)$ for $i \leqslant M$. Note that $N \geqslant M$.

It remains to prove that $\left|\sum_{n=0}^{K} a_{\sigma(n)} - A\right| < \varepsilon$ for all $K \geqslant N$.

So let $K \geqslant N$, let $L = \max\{\sigma(0),\sigma(1),\ldots,\sigma(K)\}$, and let $m_0,m_1,\ldots,m_r \in \mathbb{N}$ be such that the terms in the list $a_0,a_1,\ldots,a_L$ that remain after the terms $a_{\sigma(0)},\ldots,a_{\sigma(K)}$ have been deleted are precisely the terms a_{m_0}, a_{m_1}, …, a_{m_r}. Thus

$$\sum_{n=0}^{K} a_{\sigma(n)} = \sum_{n=0}^{L} a_n - \sum_{i=0}^{r} a_{m_i}$$

Note in particular that $L \geqslant K$. Additionally, $m_i \geqslant N$ for all $i \leqslant r$, since we ensured that the terms $a_0,a_1,\ldots,a_M$ appear amongst the terms $a_{\sigma(0)},a_{\sigma(1)},\ldots,a_{\sigma(N)}$.

By the triangle inequality (Theorem 9.1.9) we have

$$\left|\sum_{n=0}^{K} a_{\sigma(n)} - A\right| = \left|\sum_{n=0}^{L} a_n - A - \sum_{i=0}^{r} a_{m_i}\right| \leqslant \left|\sum_{n=0}^{L} a_n - A\right| + \left|\sum_{i=0}^{r} a_{m_i}\right|$$

Now conditions (1) and (2) above give the following:

(1) $L \geqslant K \geqslant N \geqslant M \geqslant M_1$, and so $\left|\sum_{n=0}^{L} a_n - A\right| < \frac{\varepsilon}{2}$.

(2) By the triangle inequality again, we have

$$\left|\sum_{i=0}^{r} a_{m_i}\right| \leqslant \sum_{i=0}^{r} |a_{m_i}| \leqslant \sum_{n=N}^{L} |a_{m_i}| < \frac{\varepsilon}{2}$$

since $L \geqslant N \geqslant M \geqslant M_2$.

Putting all of this together, we have

$$\left|\sum_{n=0}^{K} a_{\sigma(n)} - A\right| \leqslant \left|\sum_{n=0}^{L} a_n - A\right| + \left|\sum_{i=0}^{r} a_{m_i}\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as required.

By replacing a_n with $|a_n|$ in the above proof, it follows that $\sum_{n\geqslant 0} |a_{\sigma(n)}|$ converges to the same limit as $\sum_{n\geqslant 0} |a_n|$. In particular, it converges, and so the series $\sum_{n\geqslant 0} a_{\sigma(n)}$ converges absolutely. □

Example 9.3.44

Consider the following geometric series

$$\sum_{n\geqslant 0} \frac{(-1)^n}{2^n} = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \frac{1}{32} + \cdots = \frac{1}{1-(-\frac{1}{2})} = \frac{2}{3}$$

As noted in Example 9.3.35, this series converges absolutely. So if we were to add the terms in any other way, then we'd obtain the same result. For example

$$1 + \frac{1}{4} - \frac{1}{2} - \frac{1}{8} + \frac{1}{16} + \frac{1}{64} - \frac{1}{32} - \frac{1}{128} + \cdots = \frac{2}{3}$$

Absolute convergence is what allows us to do this.

Theorem 9.3.43 allows us to index absolutely convergent series over sets other than $\mathbb{N}$. This will be useful in Chapter 11, when the numbers we add are probabilities that are more naturally indexed by the outcomes of a random process than by natural numbers.

Definition 9.3.45

Let $I = \{i_n \mid n \in \mathbb{N}\}$ be a set with $i_m \neq i_n$ for all $m, n \in \mathbb{N}$ with $m \neq n$, and let $(a_i)_{i\in I}$ be an I-indexed sequence of real numbers—formally $(a_i)_{i\in I}$ is a function $a : I \to \mathbb{R}$, like in Definition 9.2.9. The **I-indexed series** $\sum_{i\in I} a_i$ is defined by

$$\sum_{i\in I} a_i = \sum_{n\geqslant 0} a_{i_n}$$

Note that, for general sums, the value of $\sum_{i\in I} a_i$ might depend on how the set I is enumerated. However, in practice, we will only use the notation $\sum_{i\in I} a_i$ when either (i) the series $\sum_{n\geqslant 0} a_{i_n}$ converges absolutely, in which case the terms can be reordered however we like; or (ii) the terms a_{i_n} are all non-negative, in which case the series either converges absolutely, or diverges no matter how the elements of I are ordered.

Example 9.3.46

Suppose $(a_k)_{k\in\mathbb{Z}}$ is a sequence of non-negative real numbers. Given $n \in \mathbb{N}$, define

$$i_n = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$

Then $\mathbb{Z} = \{i_n \mid n \in \mathbb{N}\} = \{0, -1, 1, -2, 2, \ldots\}$, and so

$$\sum_{k\in\mathbb{Z}} a_k = a_0 + a_{-1} + a_1 + a_{-2} + a_3 + a_{-3} + \cdots \overset{\star}{=} \sum_{i\geqslant 0} a_i + \sum_{i\geqslant 1} a_{-i}$$

as expected. The step $(\star)$ implicitly used non-negativity of the terms in the sequence: either the sequence diverges, or it converges absolutely, in which case we can shuffle the order of the terms.

Exercise 9.3.47

Let $J = \{j_n \mid n \in \mathbb{N}\}$ be a set such that, for all $m, n \in \mathbb{N}$, we have $j_m \neq j_n$. Let $(a_j)_{j\in J}$ be a J-indexed sequence such that either (i) $\sum_{j\in J} a_j$ is absolutely convergent, or (ii) $a_j \geqslant 0$ for all $j \in J$.

Prove that if there is a bijection $\sigma : I \to J$, then $\sum_{i\in I} a_{\sigma(i)} = \sum_{j\in J} a_j$.

Absolutely convergent series can be multiplied using the so-called *Cauchy product*, which is a kind of convolution operation on series.

Theorem 9.3.48 · *The Cauchy product*

Let $\sum_{n\geqslant 0} a_n$ and $\sum_{n\geqslant 0} b_n$ be absolutely convergent series. Then the series

$$\sum_{n\geqslant 0}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) = \left(\sum_{i\geqslant 0} a_i\right)\left(\sum_{j\geqslant 0} b_j\right)$$

Proof

By linearity of summation we have

$$\left(\sum_{n\geqslant 0} a_n\right)\left(\sum_{n\geqslant 0} b_n\right) = \sum_{i\geqslant 0}\left(\sum_{j\geqslant 0} a_i b_j\right) = \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} a_i b_j$$

Note that this series converges absolutely. Indeed, given $\varepsilon > 0$, since $\sum_{n\geqslant 0} a_n$ converges absolutely, there is some $N \in \mathbb{N}$ such that $\sum_{i>N} |a_n| < \dfrac{\varepsilon}{\sum_{j\in\mathbb{N}} |b_j|}$, and then

$$\sum_{i>N}\sum_{j=0}^{\infty} |a_i b_j| = \sum_{i>N}\left(|a_i|\sum_{j=0}^{\infty} |b_j|\right) < \varepsilon$$

as required.

Now define $P = \{(n,k) \in \mathbb{N}\times\mathbb{N} \mid k \leqslant n\}$, and note that

$$\sum_{n\geqslant 0}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) = \sum_{(n,k)\in P} a_k b_{n-k}$$

The function $\sigma : \mathbb{N}\times\mathbb{N} \to P$ defined by $\sigma(i,j) = (i+j, i)$ is a bijection. Since the series converges absolutely, we may apply Exercise 9.3.47 to obtain

$$\sum_{(n,k)\in P} a_k b_{n-k} = \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} a_i b_{(i+j)-i} = \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} a_i b_j$$

as required. □

An example of how the Cauchy product can come in handy is to prove the multiplicativity of the *exponential function*, defined in Definition 9.3.49.

Definition 9.3.49

The **exponential function** is the function $\exp : \mathbb{R} \to \mathbb{R}$ defined by

$$\exp(x) = \sum_{n \geqslant 0} \frac{x^n}{n!}$$

for all $x \in \mathbb{R}$.

Note that the exponential function is well-defined by Exercise 9.3.42.

Theorem 9.3.50 · *Multiplicativity of the exponential function*

Let $x, y \in \mathbb{R}$. Then $\exp(x+y) = \exp(x)\exp(y)$.

Proof

First note that the function

$$R : \{(n,k) \in \mathbb{N} \times \mathbb{N} \mid k \leqslant n\} \to \mathbb{N} \times \mathbb{N}$$

defined by $R(n,k) = (k, n-k)$ is a bijection. Indeed:

- Let $(n,k), (m,\ell) \in \mathbb{N} \times \mathbb{N}$ with $k \leqslant n$ and $\ell \leqslant m$. If $(k, n-k) = (\ell, m-\ell)$ then $k = \ell$, and so $n - k = m - \ell$, so that $m = n$. So R is injective.
- Let $(a,b) \in \mathbb{N} \times \mathbb{N}$. Then $b \leqslant a+b$ and $b = (a+b) - b$, and so $(a,b) = R(a+b, a)$. So R is surjective.

Now we proceed by computation:

$$\begin{aligned}
&\exp(x+y) \\
&= \sum_{n \geqslant 0} \frac{(x+y)^n}{n!} && \text{by definition of exp} \\
&= \sum_{n \geqslant 0} \sum_{k=0}^{n} \frac{1}{n!} \binom{n}{k} x^k y^{n-k} && \text{by the binomial theorem (Theorem 4.2.17)} \\
&= \sum_{n \geqslant 0} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} x^k y^{n-k} && \text{by Theorem 4.2.14} \\
&= \sum_{a \geqslant 0} \sum_{b \geqslant 0} \frac{1}{a!b!} x^a y^b && \text{using the reindexing bijection } R \\
&= \left(\sum_{a \geqslant 0} \frac{x^a}{a!} \right) \left(\sum_{b \geqslant 0} \frac{y^b}{b!} \right) && \text{by linearity of summation} \\
&= \exp(x)\exp(y) && \text{by definition of exp}
\end{aligned}$$

This is as required. □

Exercise 9.3.51
Let $x \in (-1,1)$. Prove that $\sum_{n\geqslant 1} nx^{n-1} = \dfrac{1}{(1-x)^2}$.

The constant e

We conclude this section by defining e, a mathematical constant used heavily in calculus and analysis. We will prove that e is irrational.

Lemma 9.3.52
The series $\sum_{n\geqslant 0} \dfrac{1}{n!}$ converges.

Proof
This follows from the ratio test. Indeed

$$\frac{1/(n+1)!}{1/n!} = \frac{n!}{(n+1)!} = \frac{1}{n+1} \to 0$$

so by the ratio test (Theorem 9.3.39), the series converges. □

Definition 9.3.53
The real number e, also known as **Euler's constant**, is defined by $e = \exp(1) = \sum_{n\geqslant 0} \dfrac{1}{n!}$.

The number e, like its more famous cousin π, is one of the fundamental constants of mathematics. It might seem arbitrary now, but it has remarkable analytic properties that are beyond the scope of this book.

Example 9.3.54
We can prove that $2 < e < 3$. Indeed $e > 2$ since

$$e > \frac{1}{0!} + \frac{1}{1!} = 1 + 1 = 2$$

To see that $e < 3$, note that $n! \geqslant 2^{n-1}$ for all $n \geqslant 1$, with strict inequality for $n \geqslant 2$, and so

$$e < \frac{1}{0!} + \sum_{n\geqslant 1} \frac{1}{2^{n-1}} = 1 + \frac{1}{1-\frac{1}{2}} = 3$$

Exercise 9.3.55
Prove that $\exp(x) = e^x$ for all $x \in \mathbb{Q}$.

Theorem 9.3.56
e is irrational.

Proof
Towards a contradiction, suppose that $e \in \mathbb{Q}$.

Then $k!e \in \mathbb{Z}$ for some natural number $k \geqslant 2$. Indeed, letting $e = \dfrac{a}{b}$ for some $a, b \in \mathbb{Z}$ with $b \neq 0$, we obtain $|b|e = \pm a \in \mathbb{N}$, and so we can take k to be the greatest of 2 and $|b|$.

Now observe that

$$k!e \;=\; \sum_{n \geqslant 0} \frac{k!}{n!} \;=\; \sum_{n=0}^{k} \frac{k!}{n!} \;+\; \sum_{n \geqslant k+1} \frac{k!}{n!}$$

Define $c = \displaystyle\sum_{n \geqslant k+1} \frac{k!}{n!}$. We will prove that $c \in \mathbb{Z}$ and $0 < c < 1$, which is a contradiction.

Note that for all $n \leqslant k$ we have $n!$ divides $k!$, and so $\displaystyle\sum_{n=0}^{k} \frac{k!}{n!} \in \mathbb{Z}$. Therefore

$$c \;=\; k!e - \sum_{n=0}^{k} \frac{k!}{n!} \;\in \mathbb{Z}$$

since this is the difference of two integers.

Now all the terms in the sum $\displaystyle\sum_{n \geqslant k+1} \frac{k!}{n!}$ are positive, and so

$$c \;=\; \sum_{n=k+1} \frac{k!}{n!} \;>\; \frac{k!}{(k+1)!} \;=\; \frac{1}{k+1} \;>\; 0$$

Furthermore, for all $n \geqslant k+1$, we have

$$\frac{k!}{n!} \;=\; \frac{1 \times 2 \times \cdots \times k}{1 \times 2 \times \cdots \times k \times (k+1) \times \cdots \times n} \;=\; \frac{1}{(k+1) \times \cdots \times n} \;\leqslant\; \frac{1}{(k+1)^{n-k}}$$

It follows that

$$\begin{aligned}
c = \sum_{n \geqslant k+1} \frac{k!}{n!} &\leqslant \sum_{n \geqslant k+1} \frac{1}{(k+1)^{n-k}} && \text{as we just observed} \\
&= \sum_{r \geqslant 0} \frac{1}{(k+1)^{r+1}} && \text{substituting } r = n-k-1 \\
&= \frac{1}{k+1} \sum_{r \geqslant 0} \frac{1}{(k+1)^r} && \text{by linearity} \\
&= \frac{1}{k+1} \cdot \frac{1}{1-\frac{1}{k+1}} && \text{by Theorem 9.3.8} \\
&= \frac{1}{k} && \text{rearranging} \\
&< 1 && \text{since } k \geqslant 2
\end{aligned}$$

But this implies that $0 < c < 1$, which is nonsense since $c \in \mathbb{Z}$.

We have arrived at a contradiction, so it follows that e is irrational. □

Section 9.E

Chapter 9 exercises

Inequalities and means

9.E.1. Compute $\|2(-1,-5,1,5,1)-(2,-1,1,-2,8)\|$.

9.E.2. Let $a,b\in\mathbb{R}$ and let $a_0,a_1,a_2,\ldots,a_n\in\mathbb{R}$ with $a_0=a$ and $a_n=b$. Prove that $|a-b|\leqslant\sum_{k=1}^{n}|a_{k-1}-a_k|$.

9.E.3. Let $n\in\mathbb{N}$. Prove that $\vec{x}\cdot\vec{y}=\vec{y}\cdot\vec{x}$ for all $\vec{x},\vec{y}\in\mathbb{R}^n$.

9.E.4. Let $n\in\mathbb{N}$, let $\vec{x},\vec{y}\in\mathbb{R}^n$ and let $a,b,c,d\in\mathbb{R}$. Prove that $(a\vec{x}+b\vec{y})\cdot(c\vec{x}+d\vec{y})=ac\|\vec{x}\|^2+bd\|\vec{y}\|^2+(ad+bc)(\vec{x}\cdot\vec{y})$.

9.E.5. Let $a,b,c,d\in\mathbb{R}$. Assume that $a^2+b^2=9$, $c^2+d^2=25$ and $ac+bd=15$. Find the value of $\dfrac{a+2b}{c+2d}$.

9.E.6. Let $\vec{a}\in\mathbb{R}^3$ and $r>0$, and define

$$B(\vec{a};r)=\{\vec{x}\in\mathbb{R}^3\mid\|\vec{x}-\vec{a}\|<r\}$$

Prove that $\|\vec{x}-\vec{y}\|<2r$ for all $\vec{x},\vec{y}\in B(\vec{a};r)$.

9.E.7. Prove that $a^3+b^3+c^3\geqslant abc$ for all positive real numbers a,b,c. When does equality occur?

9.E.8. Let $a,b,c>0$. Prove that $\dfrac{b+c}{a}+\dfrac{c+a}{b}+\dfrac{a+b}{c}\geqslant 6$, with equality if and only if $a=b=c$.

9.E.9. Let $a,b,c\in\mathbb{R}$. Prove that $a^4+b^4+c^4\geqslant\dfrac{(ab+bc+ca)^2}{3}$. When does equality occur?

9.E.10. *Pound cost averaging* is an investment strategy in which an investor purchases asset in equal instalments over a fixed period of time.

Let $a,b\in\mathbb{R}$ with $a<b$, let $n\geqslant 1$, and let $f:[a,b]\to(0,\infty)$ be such that at time $t\in[a,b]$, a stock is trading at £$f(t)$ per share. Assume that you purchase a total of £M of the stock in equal instalments of £$\dfrac{M}{n}$ at times $t_0,t_1,\ldots,t_{n-1}$, where $M\in[0,\infty)$, $n\geqslant 1$, and $t_i=a+i\cdot\dfrac{b-a}{n}$ for all $0\leqslant i<n$.

(a) Prove that the value of the shares that you hold at time b is equal to

$$£\,\frac{f(b)}{\mathrm{HM}\big(f(t_0),\,f(t_1),\,\ldots,\,f(t_{n-1})\big)}\cdot M$$

where $\mathrm{HM}(a_1,a_2,\ldots,a_n)$ denotes the harmonic mean of $a_1,a_2,\ldots,a_n\in\mathbb{R}$.

(b) Under what condition have you made a profit?

Sequences

9.E.11. Does there exist a sequence (x_n) such that $(x_{n+1}-x_n)\to 0$ but (x_n) diverges?

True–False questions

In Questions 9.E.12 to 9.E.18, determine (with proof) whether the statement is true or false.

9.E.12. For all $x, y \in \mathbb{R}$ we have $|x-y| \leqslant |x| - |y|$.

9.E.13. The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\vec{x}) = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$ is injective.

9.E.14. The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\vec{x}) = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$ is surjective.

9.E.15. For all $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\|2\vec{x} - 3\vec{y}\| \leqslant 2\|\vec{x}\| + 3\|\vec{y}\|$.

9.E.16. Every convergent sequence is bounded.

9.E.17. Every convergent sequence is eventually monotone.

9.E.18. Every subsequence of a divergent sequence diverges.

Always–Sometimes–Never questions

In Questions 9.E.19 to 9.E.26, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

9.E.19. Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and suppose that $\|\vec{x} - \vec{y}\| = 0$. Then $\vec{x} = \vec{y}$.

9.E.20. Let $x_1, x_2, \ldots, x_n \geqslant 0$. Then the harmonic mean of the x_is is greater than their quadratic mean.

9.E.21. Let $y_1, y_2, \ldots, y_m \geqslant 0$. Then the geometric mean of the y_js is less than or equal to their quadratic mean.

9.E.22. Let (x_n) be a sequence of real numbers and suppose that the subsequences (x_{2n}) and (x_{2n+1}) both converge. Then (x_n) converges.

9.E.23. Let (x_n) be a sequence of real numbers and suppose that the subsequences (x_{2n}), (x_{2n+1}) and (x_{3n}) all converge. Then (x_n) converges.

9.E.24. Let $f : \mathbb{R} \to \mathbb{R}$, let (x_n) be a sequence of real numbers, and suppose that $(x_n) \to a \in \mathbb{R}$. Then $(x_n^2) \to a^2$.

9.E.25. Let $f : \mathbb{R} \to \mathbb{R}$, let (x_n) be a sequence of real numbers, and suppose that $(x_n) \to a \in \mathbb{R}$. Then $(f(x_n)) \to f(a)$.

9.E.26. Let (a_n) be a sequence of real numbers and suppose that $(a_n) \to 0$. Then $\sum_{n \geqslant 0} a_n$ converges.

Chapter 10

Infinity

Section 10.1
Cardinality

Section 6.1 was all about defining a notion of *size* for finite sets, and using this definition to compare and contrast the sizes of finite sets by constructing injections, surjections and bijections between them.

We made some progress in comparing the sizes of infinite sets in Section 6.2, but only to a certain extent: at this point, we can only distinguish between two sizes of infinity, namely 'countable' and 'uncountable'. While this is interesting, we can do much better—that is where this section comes in.

The *cardinality* (defined in Definition 10.1.1) of a set can be understood as a measure of its size, generalising the notion of size for finite sets. In particular:

- Whereas the size of a *finite* set X is a natural number $|X| \in \mathbb{N}$, the cardinality of an arbitrary set is a *cardinal number*—but if the set happens to be finite, then this cardinal number is equal to the natural number given by its size.
- We proved in Theorem 6.1.13 that two finite sets X and Y have equal size if and only if there is a bijection $X \to Y$. We generalise this fact to arbitrary sets by building it into the definition of cardinality: that is, two sets X and Y will have equal cardinality if and only if there is a bijection $X \to Y$.

Without further ado, behold the definition of cardinality.

Definition 10.1.1

The **cardinality** of a set X is an element $|X|$ of the collection Card (LaTeX code: `\mathsf{Card}`) of all **cardinal numbers**, defined so that the following properties hold:

(i) For every set X, there is a unique cardinal number $\kappa \in \mathsf{Card}$ such that $|X| = \kappa$;

(ii) For all sets X and Y, we have $|X| = |Y|$ if and only if there exists a bijection $X \to Y$;

(iii) $\mathbb{N} \subseteq \mathsf{Card}$, and if X is finite, then its cardinality $|X|$ is equal to its size; and

(iv) For each cardinal number κ, there exists a set $[\kappa]$ with $|[\kappa]| = \kappa$, with $[n]$ defined as in Definition 0.25 for all $n \in \mathbb{N} \subseteq \mathsf{Card}$.

A cardinal number $\kappa \in \mathsf{Card} \setminus \mathbb{N}$ is called an **infinite cardinal number**.

Condition (iii) ensures that cardinality generalises the notion of size: whereas size is defined only for finite sets, cardinality is defined for both finite and infinite sets.

Definition 10.1.1 tells us only what properties cardinal numbers satisfy; it doesn't tell us what they actually *are*, how they are defined, or even whether they exist. This is on purpose: the construction of the cardinal numbers is very intricate, and far beyond the scope of this introductory textbook.

One cardinal number of interest to us is the cardinality of the natural numbers. In a sense that we will

make precise later (Theorem 10.1.13), we can think of this cardinal number as being the *smallest* infinite cardinal number.

Definition 10.1.2

The cardinal number $\aleph_0$ (LaTeX code: `\aleph_0`), called **aleph naught** (or **aleph null**), is defined by $\aleph_0 = |\mathbb{N}|$.

The symbol $\aleph$ is the Hebrew letter *aleph*. The cardinal number $\aleph_0$ is the first infinite cardinal number in a hierarchy of so-called *well-orderable cardinals*.

Example 10.1.3

A set X is countably infinite if and only if $|X| = \aleph_0$. Indeed, to say that X is countably infinite is to say that there is a bijection $\mathbb{N} \to X$, which by Definition 10.1.1(ii) is equivalent to saying that $\aleph_0 = |\mathbb{N}| = |X|$.

In light of Example 10.1.3, we have

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = \aleph_0$$

So what about $\mathbb{R}$? We know that $|\mathbb{R}| \neq \aleph_0$ by Theorem 9.3.21. The cardinality of the real numbers has its own notation, given next.

Definition 10.1.4

The **cardinality of the continuum** is the cardinal number $\mathfrak{c}$ (LaTeX code: `\mathfrak{c}`) defined by $\mathfrak{c} = |\mathbb{R}|$.

Exercise 10.1.5

Consider the function $f : \mathbb{R} \to (-1,1)$ defined by

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

for all $x \in \mathbb{R}$. Prove that f is a bijection, and deduce that $|(-1,1)| = \mathfrak{c}$.

Exercise 10.1.6

Let $a, b \in \mathbb{R}$ with $a < b$. Prove that each of the sets (a,b), $[a,b)$, $(a,b]$ and $[a,b]$ has cardinality $\mathfrak{c}$.

Ordering the cardinal numbers

In Theorem 6.1.13 we proved that, given finite sets X and Y, we can prove that $|X| \leqslant |Y|$ by constructing an injection $X \to Y$. This made intuitive sense: for an injection $X \to Y$ to exist, there must be sufficiently many elements in Y to be able to spread out all of the elements of X.

The intuition bestowed upon us by the finite case yields the following definition.

Definition 10.1.7
Given cardinal numbers κ and λ, we say that κ is **less than or equal to** λ, and write $\kappa \leqslant \lambda$, if there exists an injection $[\kappa] \to [\lambda]$. We say κ is **less than** λ, and write $\kappa < \lambda$, if $\kappa \leqslant \lambda$ and $\kappa \neq \lambda$.

A word of warning is pertinent at this point. Given $m, n \in \mathbb{N}$, but regarding m and n as cardinal numbers, Definition 10.1.7 defines the expression '$m \leqslant n$' to mean that there exists an injection $[m] \to [n]$. This is not how '$m \leqslant n$' is typically defined for natural numbers m and n. Fortunately for us, we proved in Theorem 6.1.6 that $m \leqslant n$ (in the usual sense) if and only if there is an injection $[m] \to [n]$. This ensures that these two notions of '$\leqslant$'—for cardinal numbers and for natural numbers—are consistent with one another.

This is typical for the definitions we will make involving cardinal numbers, particularly in Section 10.2: results that we *proved* for natural numbers in Chapter 8 will be generalised to form *definitions* for cardinal numbers, but then these generalised definitions are consistent with the usual definitions for natural numbers. While it is not worth losing sleep over these matters, it is good practice to check that the definitions we make are not mutually contradictory!

We may at times write $\lambda \geqslant \kappa$ to mean $\kappa \leqslant \lambda$, and $\lambda > \kappa$ to mean $\kappa < \lambda$. Note that $\lambda \geqslant \kappa$ should *not* be interpreted to mean 'there exists a surjection $[\lambda] \to [\kappa]$'—for example, that would imply that $1 \geqslant 0$ is false, which is nonsense.

In any case Definition 10.1.7 provides the following proof strategy.

Strategy 10.1.8
Let X and Y be sets.

- In order to prove $|X| \leqslant |Y|$, it suffices to find an injection $X \to Y$.
- In order to prove $|X| < |Y|$, it suffices to find an injection $X \to Y$, and prove that there does not exist a surjection $X \to Y$.

Example 10.1.9
$\aleph_0 \leqslant \mathfrak{c}$. To see this, note that the function $i : \mathbb{N} \to \mathbb{R}$ given by $i(n) = n$ for all $n \in \mathbb{N}$ is an injection.

Exercise 10.1.10
Prove that $n < \aleph_0$ for all $n \in \mathbb{N}$.

Exercise 10.1.11
Prove that $\leqslant$ is a reflexive, transitive relation on Card. That is, prove that:

(a) $\kappa \leqslant \kappa$ for all cardinal numbers κ; and

(b) For all cardinal numbers κ, λ, μ, if $\kappa \leqslant \lambda$ and $\lambda \leqslant \mu$, then $\kappa \leqslant \mu$.

Theorem 10.1.12 · *Cantor's theorem*
Let X be a set. Then $|X| < |\mathscr{P}(X)|$.

Proof
The function $x \mapsto \{x\}$ evidently defines an injection $X \to \mathscr{P}(X)$, so $|X| \leqslant |\mathscr{P}(X)|$. The fact that $|X| \neq |\mathscr{P}(X)|$ is then immediate from Exercise 3.2.16, which proves that no function $X \to \mathscr{P}(X)$ is surjective. □

Cantor's theorem implies that there is no bound on how large a cardinal number can be. Indeed, if κ is any cardinal number, then Cantor's theorem implies that

$$\kappa = |[\kappa]| < |\mathscr{P}([\kappa])|$$

and so $|\mathscr{P}([\kappa])|$ is a larger cardinal number yet.

Earlier in the section we claimed that $\aleph_0$ is, in a suitable sense, the *smallest* infinite cardinal number. Theorem 10.1.13 makes it clear what we mean by this.

Theorem 10.1.13
The cardinal number $\aleph_0$ is the smallest infinite cardinal in the following sense:

(a) $n \leqslant \aleph_0$ for all $n \in \mathbb{N}$; and

(b) For all cardinal numbers κ, if $n \leqslant \kappa$ for all $n \in \mathbb{N}$, then $\aleph_0 \leqslant \kappa$.

Thus $\aleph_0$ can be thought of as the cardinal supremum of $\mathbb{N} \subseteq \mathsf{Card}$.

Proof
For part (a), note that for each $n \in \mathbb{N}$, we have $[n] \subseteq \mathbb{N}$, and so the function $i : [n] \to \mathbb{N}$ given by $i(k) = k$ for all $k \in [n]$ is an injection. It follows that

$$n = |[n]| \leqslant |\mathbb{N}| = \aleph_0$$

as required.

For part (b), fix a cardinal number κ and assume that $n \leqslant \kappa$ for all $n \in \mathbb{N}$. Then there exist injections $j_n : [n] \to [\kappa]$ for each $n \in \mathbb{N}$.

We will use these injections to construct an injection $f : \mathbb{N} \to [\kappa]$ in the following way: for each $n \in \mathbb{N}$, the function $j_{n+1} : [n+1] \to [\kappa]$ takes exactly $n+1$ values since it is injective. This means that for all $n \in \mathbb{N}$, the function j_{n+2} takes at least one value in $[\kappa]$ that the function j_{n+1} does not take—we will define $f(n+1)$ to be one such value.

Now let's define f properly: define $f(n) \in [\kappa]$ for $n \in \mathbb{N}$ recursively as follows:

- $f(0) = j_1(0) \in [\kappa]$.
- Fix $n \in \mathbb{N}$, and suppose $f(k)$ has been defined for all $k \leqslant n$ and satisfies $f(k) \neq f(\ell)$ for $\ell < k$. Let
$$A = \{a \in [n+2] \mid j_{n+2}(a) \neq f(k) \text{ for all } k \leqslant n\}$$
Note that A is inhabited: the set $\{j_{n+2}(a) \mid a \in [n+2]\}$ has size $n+2$ since j_{n+2} is injective, and the set $\{f(k) \mid k \leqslant n\}$ has size $n+1$ by construction, so at least one $a \in [n+2]$ must satisfy the requirement that $j_{n+2}(a) \neq f(k)$ for all $k \leqslant n$.

Let a be the least element of A, and define $f(n+1) = j_{n+2}(a)$. Then by construction we have $f(n+1) \neq f(k)$ for any $k < n$, as required.

By construction, the function f is injective, since for all $n \in \mathbb{N}$, the value $f(n) \in [\kappa]$ was defined so that $f(n) \neq f(k)$ for any $k < n$.

Hence $\aleph_0 = |\mathbb{N}| \leqslant |[\kappa]| = \kappa$, as required. □

The Cantor–Schröder–Bernstein theorem

Even if two sets X and Y have the same cardinality, it is not always easy to find a bijection $X \to Y$. This is problematic if we want to prove that they *do* have the same cardinality! The Cantor–Schröder–Bernstein theorem greatly simplifies this process, allowing us to deduce that $|X| = |Y|$ from the existence of injections $X \to Y$ and $Y \to X$.

Theorem 10.1.14 · *Cantor–Schröder–Bernstein theorem*
The relation $\leqslant$ on Card is antisymmetric. That is, for all cardinal numbers κ and λ, if $\kappa \leqslant \lambda$ and $\lambda \leqslant \kappa$, then $\kappa = \lambda$.

Proof
This is one of the most involved proofs that we will see in this book, and so we break it into steps. Some details are left as exercises, in part because the details cloud the bigger picture of the proof, and in part because they provide good practice with working with all the definitions.

Let κ and λ be cardinal numbers and assume that $\kappa \leqslant \lambda$ and $\lambda \leqslant \kappa$. Then there exist injections $f : [\kappa] \to [\lambda]$ and $g : [\lambda] \to [\kappa]$.

The steps we will follow are these:

- **Step 1.** We will use the injections f and g to partition $[\kappa]$ and $[\lambda]$ into equivalence classes. Intuitively, two elements of $[\kappa]$ will be 'equivalent' if one can be obtained from the other by successively applying $g \circ f$, and likewise for $[\lambda]$ with $f \circ g$.
- **Step 2.** We will prove that f and g induce a bijection $[\kappa]/{\sim} \to [\lambda]/{\approx}$—that is, they pair up the $\sim$-equivalence classes with the $\approx$-equivalence classes.
- **Step 3.** We will prove that there is a bijection between each pair of the paired-up equivalence classes.
- **Step 4.** We will piece together the bijections between equivalence classes to obtain a bijection $[\kappa] \to [\lambda]$.

So here we go.

Step 1. Define a relation $\sim$ on $[\kappa]$ by letting

$$a \sim b \quad \Leftrightarrow \quad (g \circ f)^n(a) = b \text{ or } (g \circ f)^n(b) = a \text{ for some } n \in \mathbb{N}$$

for all $a, b \in [\kappa]$. Here $(g \circ f)^n$ refers to the n-fold composite of $g \circ f$, that is

$$(g \circ f)^n(x) = \underbrace{((g \circ f) \circ (g \circ f) \circ \cdots \circ (g \circ f))}_{n \text{ copies of } (g \circ f)}(x)$$

In other words, $a \sim b$ means that we can get from a to b, or from b to a, by applying the function $g \circ f$ some number of times.

Exercise 10.1.15
Prove that $\sim$ is an equivalence relation on $[\kappa]$.

Likewise the relation $\approx$ on $[\lambda]$, defined by letting

$$c \approx d \quad \Leftrightarrow \quad (f \circ g)^n(c) = d \text{ or } (f \circ g)^n(d) = c \text{ for some } n \in \mathbb{N}$$

for all $c, d \in [\lambda]$, is an equivalence relation.

Step 2. Define functions $p : [\kappa]/\sim \to [\lambda]/\approx$ and $q : [\lambda]/\approx \to [\kappa]/\sim$ by

$$p([a]_\sim) = [f(a)]_\approx \quad \text{and} \quad q([b]_\approx) = [g(b)]_\sim$$

for all $[a]_\sim \in [\kappa]/\sim$ and $[b]_\approx \in [\lambda]$.

Exercise 10.1.16
Prove that p and q are well-defined, and that q is an inverse for p.

In particular, p defines a bijection $[\kappa]/\sim \to [\lambda]/\approx$.

Step 3. Fix $a \in [\kappa]$ and let $b = f(a) \in [\lambda]$. We prove that there is a bijection $[a]_\sim \to [b]_\approx$. There are three possible cases:

- **Case 1.** Suppose $[a]_\sim$ contains an element a_0 such that $a_0 \neq g(y)$ for any $y \in [\lambda]$. Define a function $h_a : [a]_\sim \to [b]_\approx$ by letting $h_a(x) = f(x)$ for all $x \in [a]_\sim$.
- **Case 2.** Suppose $[b]_\approx$ contains an element b_0 such that $b_0 \neq f(x)$ for any $x \in [\kappa]$. Define a function $k_b : [b]_\approx \to [a]_\sim$ by letting $k_b(y) = g(x)$ for all $y \in [b]_\approx$.
- **Case 3.** Otherwise, define $h_a : [a]_\sim \to [b]_\approx$ by $h_a(x) = f(x)$ for all $x \in [a]_\sim$.

Exercise 10.1.17
Prove that each of the functions defined in the above three cases is a bijection.

Step 4. By taking $h_a = k_b^{-1} : [a]_\sim \to [b]_\approx$ in Case 2 above, we have proved that:

- There is a bijection $p : [\kappa]/\sim \to [\lambda]/\approx$; and
- For each $E = [a]_\sim \in [\kappa]/\sim$, there is a bijection $h_a : E \to p(E)$.

By Exercise 5.2.30, it follows that there is a bijection $[\kappa] \to [\lambda]$, as required. This completes the proof. □

The Cantor–Schröder–Bernstein theorem is not just an interesting fact: it is useful for proving that two sets have the same cardinality without having to explicitly construct a bijection between them.

❖ **Strategy 10.1.18** · *Proving that two sets have equal cardinality*
Let X and Y be sets. In order to prove that $|X| = |Y|$, it suffices to show that there exists an injection $X \to Y$ and an injection $Y \to X$.

✎ **Example 10.1.19**
Let $a, b \in \mathbb{R}$ with $a < b$, and consider the chain of functions

$$(0,1) \xrightarrow{f_1} (a,b) \xrightarrow{f_2} [a,b] \xrightarrow{f_3} [0,1] \xrightarrow{f_4} (0,1)$$

defined by:

- $f_1(t) = a + t(b-a)$ for all $t \in (0,1)$.
- $f_2(t) = t$ for all $t \in (a,b)$.
- $f_3(t) = \dfrac{t-a}{b-a}$ for all $t \in [a,b]$.
- $f_4(t) = \frac{1}{4} + \frac{1}{2}t$ for all $t \in [0,1]$.

Each of these functions is injective by Exercise 3.2.8. Hence we can compose these functions to obtain injections from any of these sets to any other—for example, $f_1 \circ f_4 \circ f_3 : [0,1] \to (a,b)$ is an injection.

By the Cantor–Schröder–Bernstein theorem, it follows that

$$|(0,1)| = |(a,b)| = |[a,b]| = |[0,1]|$$

✎ **Example 10.1.20**
Here is a proof that $\mathbb{N} \times \mathbb{N}$ is countably infinite using the Cantor–Schröder–Bernstein theorem.

Define $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ by $f(n) = (n,0)$ for all $n \in \mathbb{N}$. Given $m, n \in \mathbb{N}$ we have

$$f(m) = f(n) \quad \Rightarrow \quad (m,0) = (n,0) \quad \Rightarrow \quad m = n$$

So f is injective.

Next, define $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $g(m,n) = 2^m \cdot 3^n$ for all $m, n \in \mathbb{N}$. Then given $(m,n), (p,q) \in \mathbb{N} \times \mathbb{N}$, if $g(m,n) = g(p,q)$, then $2^m \cdot 3^n = 2^p \cdot 3^q$. It follows from the fundamental theorem of arithmetic (Theorem 7.2.12) that $m = n$ and $p = q$, so that g is injective.

Since $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ and $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ are injective, it follows from the Cantor–Schröder–Bernstein theorem that

$$|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}| = \aleph_0$$

so that $\mathbb{N} \times \mathbb{N}$ is countably infinite.

We can use the Cantor–Schröder–Bernstein theorem to prove that $\mathscr{P}(\mathbb{N})$ has the cardinality of the coninuum.

Theorem 10.1.21
$|\mathscr{P}(\mathbb{N})| = \mathfrak{c}$

Proof
Define a function $f : \mathscr{P}(\mathbb{N}) \to \mathbb{R}$ as follows. Given $U \subseteq \mathbb{N}$, define

$$U_n = \begin{cases} 1 & \text{if } n \in U \\ 0 & \text{if } n \notin U \end{cases}$$

Let $f(U)$ be the real number whose decimal expansion is $U_0.U_1U_2\ldots$. Then f is an injection: given $U, V \subseteq \mathbb{N}$, if $f(U) = f(V)$, then $f(U)$ and $f(V)$ have the same decimal expansion, so that $U_n = V_n$ for all $n \in \mathbb{N}$. But that says precisely that $n \in U$ if and only if $n \in V$ for all $n \in \mathbb{N}$, so that $U = V$.

Next, define a function $g : [0,1) \to \mathscr{P}(\mathbb{N})$ as follows: given $x \in \mathbb{R}$, let the binary expansion of x be

$$x = 0.x_1x_2x_3x_4\ldots{}_{(2)}$$

Define $g(x) = \{i \in \mathbb{N} \mid x_i = 1\}$. Then g is injective by uniqueness of binary expansions again.

Then:

- Since f is injective we have $|\mathscr{P}(\mathbb{N})| \leqslant |\mathbb{R}| = \mathfrak{c}$.
- Since g is injective we have $\mathfrak{c} = |[0,1)| \leqslant |\mathscr{P}(\mathbb{N})|$.

By the Cantor–Schröder–Bernstein theorem (Theorem 10.1.14), it follows that $|\mathscr{P}(\mathbb{N})| = \mathfrak{c}$. □

Exercise 10.1.22
Let $\mathscr{F}(\mathbb{N})$ be the set of all finite subsets of $\mathbb{N}$, and define $f : \mathscr{F}(\mathbb{N}) \to \mathbb{N}$ by

$$f(U) = \sum_{a \in U} 10^a = \sum_{k=1}^{n} 10^{a_k}$$

for all $U = \{a_1, a_2, \ldots, a_n\} \in \mathscr{F}(\mathbb{N})$. Put another way, $f(U)$ is the natural number whose decimal expansion has a 1 in the r^{th} position (counting from $r = 0$) if $r \in U$, and a 0 otherwise. For example

$$f(\{0,1,3,4,8\}) = 100011011 \quad \text{and} \quad f(\varnothing) = 0$$

Prove that f is injective, and use the Cantor–Schröder–Bernstein theorem to deduce that $\mathscr{F}(\mathbb{N})$ is countably infinite.

Exercise 10.1.23
Let X, Y and Z be sets. Prove that if $|X| = |Z|$ and $X \subseteq Y \subseteq Z$, then $|X| = |Y| = |Z|$.

Section 10.2

Cardinal arithmetic

In this section we will define arithmetic operations for cardinal numbers, and then derive *infintary* counting principles by analogy with Section 8.1. It is surprising how little additional work needs to be done for the results proved there to carry over.

We begin by defining the sum $\kappa+\lambda$, product $\kappa\cdot\lambda$ and power λ^κ, for cardinal numbers κ and λ.

Cardinal addition

Definition 10.2.1

The (**cardinal**) **sum** of cardinal numbers κ and λ is the cardinal number $\kappa+\lambda$ defined by

$$\kappa+\lambda = |[\kappa]\sqcup[\lambda]|$$

where for sets A and B, the notation $A\sqcup B$ denotes the **disjoint union**

$$A\sqcup B = (A\times\{0\})\cup(B\times\{1\})$$

as discussed in Exercise 6.1.19.

Note that Definition 10.2.1 is compatible with addition for natural numbers by Exercise 6.1.19(b), which implies that $|[m]\sqcup[n]| = m+n$ for all $m,n\in\mathbb{N}$.

The following lemma makes cardinal addition easier to work with.

Lemma 10.2.2

Let X and Y be sets. If $X\cap Y=\varnothing$, then $|X\cup Y| = |X|+|Y|$.

Proof

Let $\kappa=|X|$ and $\lambda=|Y|$. Note that $|[\kappa]\times\{0\}| = \kappa = |X|$ and $|[\lambda]\times\{1\}| = \lambda = |Y|$, so there are bijections $f:[\kappa]\times\{0\}\to X$ and $g:[\lambda]\times\{1\}\to Y$.

Define a function $h:[\kappa]\sqcup[\lambda]\to X\cup Y$ by

$$h(a,i) = \begin{cases} f(a) & \text{if } i=0\\ g(b) & \text{if } i=1\end{cases}$$

for all $(a,i)\in[\kappa]\sqcup[\lambda]$.

Then h is a bijection; to see this, define $k:X\cup Y\to[\kappa]\sqcup[\lambda]$ by

$$k(a) = \begin{cases} (f^{-1}(a),0) & \text{if } a\in X\\ (g^{-1}(a),1) & \text{if } a\in Y\end{cases}$$

for all $a \in X \cup Y$. Then k is well-defined since $X \cap Y = \varnothing$ and f and g are bijections. And k can readily be seen to be an inverse for h.

Since $h : [\kappa] \sqcup [\lambda] \to X \cup Y$ is a bijection, we have

$$|X \cup Y| = |[\kappa] \sqcup [\lambda]| = \kappa + \lambda$$

as required. □

Example 10.2.3

We show that $\aleph_0 + \aleph_0 = \aleph_0$. To see this, note that $\mathbb{N} = E \cup O$, where E is the set of all even natural numbers and O is the set of all odd natural numbers. But E and O are disjoint, so that by Lemma 10.2.2 we have

$$\aleph_0 + \aleph_0 \;=\; |E| + |O| \;=\; |E \cup O| \;=\; |\mathbb{N}| \;=\; \aleph_0$$

as claimed.

Many of the basic properties enjoyed by addition of natural numbers carry over to cardinal numbers.

Theorem 10.2.4 · *Properties of cardinal addition*

(a) $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$ for all cardinal numbers κ, λ, μ;

(b) $\kappa + \lambda = \lambda + \kappa$ for all cardinal numbers κ, λ;

(c) $0 + \kappa = \kappa = \kappa + 0$ for all cardinal numbers κ.

Proof of (a)

Let κ, λ and μ be cardinal numbers, and define

$$X = [\kappa] \times \{0\}, \quad Y = [\lambda] \times \{1\}, \quad Z = [\mu] \times \{2\}$$

Note that $|X| = \kappa$, $|Y| = \lambda$ and $|Z| = \mu$, and that X, Y and Z are pairwise disjoint. Therefore X and $Y \cup Z$ are disjoint, and $X \cup Y$ and Z are disjoint, so that

$$\kappa + (\lambda + \mu) = |X| + (|Y| + |Z|) = |X| + |Y \cup Z| = |X \cup (Y \cup Z)|$$

and likewise

$$(\kappa + \lambda) + \mu = (|X| + |Y|) + |Z| = |X \cup Y| + |Z| = |(X \cup Y) \cup Z|$$

But $X \cup (Y \cup Z) = (X \cup Y) \cup Z$, so that $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$, as required. □

Exercise 10.2.5

Prove parts (b) and (c) of Theorem 10.2.4.

We can generalise the argument in Example 10.2.3 to prove the following proposition.

Proposition 10.2.6
$\kappa + \aleph_0 = \kappa$ for all cardinal numbers $\kappa \geqslant \aleph_0$.

Proof
Let $\kappa \geqslant \aleph_0$ and let $A = [\kappa]$. Note that by Definition 10.1.7 there is an injection $i : \mathbb{N} \to A$. Write $a_n = i(n)$ for all $n \in \mathbb{N}$, so that $i[\mathbb{N}] = \{a_0, a_1, a_2, \dots\} \subseteq A$.

Partition A as $A = B \cup U \cup V$, where:

- $B = A \setminus i[\mathbb{N}] = A \setminus \{a_0, a_1, a_2, a_3, \dots\}$;
- $U = i[E] = \{a_0, a_2, a_4, \dots\}$; and
- $V = i[O] = \{a_1, a_3, a_5, \dots\}$.

Here E and O are the sets of even and odd natural numbers, respectfully.

The sets U, V and $i[\mathbb{N}]$ are all countably infinite, since E, O and $\mathbb{N}$ are countably infinite and i is an injection.

Since U and $i[\mathbb{N}]$ are countably infinite, there is a bijection $f : U \to i[\mathbb{N}]$, which in turn yields a bijection $g : B \cup U \to B \cup i[\mathbb{N}]$ defined by

$$g(a) = \begin{cases} a & \text{if } a \in B \\ f(a) & \text{if } a \in U \end{cases}$$

Note that g is well-defined since $A \cap U = \varnothing$ and $B \cap i[\mathbb{N}] = \varnothing$. And g is bijective: it has an inverse function, defined similarly but with f replaced by f^{-1}.

Thus

$$|A| \;=\; |B \cup i[\mathbb{N}]| \;=\; |B \cup U|$$

Since U and V are countably infinite, we have $|U| = |V| = \aleph_0$. It follows that

$$\kappa + \aleph_0 \;=\; |A| + |V| \;=\; |B \cup U| + |V| \;=\; |B \cup U \cup V| \;=\; |A| \;=\; \kappa$$

as required. □

Cardinal multiplication

Products of cardinal numbers are defined by analogy with the result of Exercise 6.1.21, which says that $|[m] \times [n]| = mn$ for all $m, n \in \mathbb{N}$.

Definition 10.2.7
The (**cardinal**) **product** of cardinal numbers κ and λ is the cardinal number $\kappa \cdot \lambda$ defined by

$$\kappa \cdot \lambda = |[\kappa] \times [\lambda]|$$

That is, the cardinal product is the cardinality of the cartesian product.

Like Lemma 10.2.2, the following lemma allows us to replace $[\kappa]$ and $[\lambda]$ in Definition 10.2.7 by arbitrary sets of size κ and λ, respectively.

Lemma 10.2.8
Let X and Y be sets. Then $|X \times Y| = |X| \times |Y|$.

Proof
Let $\kappa = |X|$ and $\lambda = |Y|$, and fix bijections $f : [\kappa] \to X$ and $g : [\lambda] \to Y$. By Exercise 3.2.47, there is a bijection $[\kappa] \times [\lambda] \to X \times Y$, and so

$$|X \times Y| = |[\kappa] \times [\lambda]| = \kappa \cdot \lambda$$

as required. □

Example 10.2.9
We will prove that $\aleph_0 \cdot \aleph_0 = \aleph_0$. Indeed, we proved in Example 10.1.20 that $|\mathbb{N} \times \mathbb{N}| = \aleph_0$, and so by Lemma 10.2.8 we have

$$\aleph_0 \cdot \aleph_0 \;=\; |\mathbb{N} \times \mathbb{N}| \;=\; \aleph_0$$

as required.

Example 10.2.10
We will prove that $\aleph_0 \cdot \mathfrak{c} = \mathfrak{c}$.

Define $f : \mathbb{Z} \times [0,1) \to \mathbb{R}$ by $f(n,r) = n + r$ for all $n \in \mathbb{Z}$ and all $0 \leqslant r < 1$. Then:

- f **is injective.** To see this, let $(n,r),(m,s) \in \mathbb{Z} \times [0,1)$ and assume that $n + r = m + s$. Then $n - m = s - r$. Since $s, r \in [0,1)$, we have $s - r \in (-1,1)$, and so $n - m \in (-1,1)$. Since $n - m \in \mathbb{Z}$, we must have $n - m = 0$, and so $m = n$. But then $n + r = n + s$, and so $r = s$. Thus $(n,r) = (m,s)$, as required.
- f **is surjective.** To see this, let $x \in \mathbb{R}$. Let $n \in \mathbb{Z}$ be such that $n \leqslant x < n + 1$, and let $r = x - n$. Note that $0 \leqslant r < 1$, so that $(n,r) \in \mathbb{Z} \times [0,1)$, and then $x = n + r = f(n,r)$, as required.

Since f is a bijection, we have

$$\aleph_0 \cdot \mathfrak{c} \;=\; |\mathbb{Z}| \times |[0,1)| \;=\; |\mathbb{Z} \times [0,1)| \;=\; |\mathbb{R}| \;=\; \mathfrak{c}$$

as required.

Exercise 10.2.11
Prove that $\mathfrak{c} \cdot \mathfrak{c} = \mathfrak{c}$.

Exercise 10.2.12
Let X be a set and let $\sim$ be an equivalence relation on X. Prove that if κ is a cardinal number such that $|[a]_\sim| = \kappa$ for all $a \in X$, then $|X| = |X/\sim| \cdot \kappa$.

Theorem 10.2.13 · *Properties of cardinal multiplication*

(a) $\kappa \cdot (\lambda \cdot \mu) = (\kappa \cdot \lambda) \cdot \mu$ for all cardinal numbers κ, λ, μ;

(b) $\kappa \cdot \lambda = \lambda \cdot \kappa$ for all cardinal numbers κ, λ;

(c) $1 \cdot \kappa = \kappa = \kappa \cdot 1$ for all cardinal numbers κ;

(d) $\kappa \cdot (\lambda + \mu) = (\kappa \cdot \lambda) + (\kappa \cdot \mu)$ for all cardinal numbers κ, λ, μ.

Proof of (a)
Define $f : [\kappa] \times ([\lambda] \times [\mu]) \to ([\kappa] \times [\lambda]) \times [\mu]$ by $f(a,(b,c)) = ((a,b),c)$ for all $a \in [\kappa]$, $b \in [\lambda]$ and $c \in [\mu]$. Then f is a bijection, since it has an inverse $g : ([\kappa] \times [\lambda]) \times [\mu] \to [\kappa] \times ([\lambda] \times [\mu])$ defined by $g((a,b),c) = (a,(b,c))$ for all $a \in [\kappa]$, $b \in [\lambda]$ and $c \in [\mu]$.

But then by Lemma 10.2.8 we have

$$\kappa \cdot (\lambda \cdot \mu) \;=\; |[\kappa] \times ([\lambda] \times [\mu])| \;=\; |([\kappa] \times [\lambda]) \times [\mu]| \;=\; (\kappa \cdot \lambda) \cdot \mu$$

as required. □

Exercise 10.2.14
Prove parts (b), (c) and (d) of Theorem 10.2.13.

Cardinal exponentiation

Definition 10.2.15
Let κ and λ be cardinal numbers. The κ^{th} **(cardinal) power** of λ is the cardinal number λ^κ defined by

$$\lambda^\kappa = |[\lambda]^{[\kappa]}|$$

where for sets A and B, the notation B^A refers to the set of functions $A \to B$.

Again, exponentiation of cardinal numbers agrees with that of natural numbers, since we proved in Exercise 8.1.19 that $|[n]^{[m]}| = n^m$ for all $m, n \in \mathbb{N}$.

Like with cardinal multiplication, the next lemma proves that we can replace the sets $[\kappa]$ and $[\lambda]$ in Definition 10.2.15 with arbitrary sets of cardinality κ and λ, respectively.

Lemma 10.2.16
Let X and Y be sets. Then $|Y^X| = |Y|^{|X|}$.

Proof
Let $\kappa = |X|$ and $\lambda = |Y|$, and fix bijections $f : [\kappa] \to X$ and $g : [\lambda] \to Y$.

Define $H : [\lambda]^{[\kappa]} \to Y^X$ as follows. Given $\theta \in [\lambda]^\kappa$, that is $\theta : [\kappa] \to [\lambda]$, define $h_\theta = g \circ \theta \circ f^{-1} : X \to Y$, and let $H(\theta) = h_\theta \in Y^X$.

To see that H is a bijection, note that the function $K : Y^X \to [\lambda]^{[\kappa]}$ defined by $K(\varphi) = k_\varphi = g^{-1} \circ \varphi \circ f$ is a bijection, since for all $\theta : [\kappa] \to [\lambda]$ we have

$$K(H(\theta)) = K(g \circ \theta \circ f^{-1} = g^{-1} \circ g \circ \theta \circ f^{-1} \circ f = \theta$$

and for all $\varphi : X \to Y$, we have

$$H(K(\varphi)) = H(g^{-1} \circ \varphi \circ f) = g \circ g^{-1} \circ \varphi \circ f \circ f^{-1} = \varphi$$

Since $H : [\lambda]^{[\kappa]} \to Y^X$ is a bijection, we have

$$|Y^X| \;=\; |[\lambda]^{[\kappa]}| \;=\; \lambda^\kappa$$

as required. □

Example 10.2.17
We prove that $\kappa^2 = \kappa \cdot \kappa$ for all cardinal numbers κ.

To see this, define $f : [\kappa]^{[2]} \to [\kappa] \times [\kappa]$ by letting $f(\theta) = (\theta(0), \theta(1))$ for all $\theta : [2] \to [\kappa]$. Then

- f is injective. To see this, let $\theta, \varphi : [2] \to [\kappa]$ and assume $f(\theta) = f(\varphi)$. Then $(\theta(0), \theta(1)) = (\varphi(0), \varphi(1))$, so that $\theta(0) = \varphi(0)$ and $\theta(1) = \varphi(1)$. But then $\theta = \varphi$ by function extensionality, as required.
- f is surjective. To see this, let $(a, b) \in [\kappa] \times [\kappa]$, and define $\theta : [2] \to [\kappa]$ by letting $\theta(0) = a$ and $\theta(1) = b$. Then we have $f(\theta) = (\theta(0), \theta(1)) = (a, b)$, as required.

Since f is a bijection, it follows that

$$\kappa^2 \;=\; |[\kappa]|^{|[2]|} \;=\; |[\kappa]^{[2]}| \;=\; |[\kappa] \times [\kappa]| \;=\; \kappa \cdot \kappa$$

as required.

Cardinal exponentiation gives us a convenient way of expressing the cardinalities of power sets.

Theorem 10.2.18
Let X be a set. Then $|\mathscr{P}(X)| = 2^{|X|}$.

Proof
There is a bijection $i : \mathscr{P}(X) \to [2]^X$ defined for all $U \subseteq X$ by letting $i(U) = \chi_U \in [2]^X$, where $\chi_U : X \to [2]$ is the characteristic function of U (see Definition 3.1.23). Note that i is a bijection by Theorem 3.1.25.

It follows by Lemma 10.2.16 that

$$|\mathscr{P}(X)| = |[2]^X| = |[2]|^{|X|} = 2^{|X|}$$

as required. □

In light of Theorem 10.2.18, we can interpret Cantor's theorem (Theorem 10.1.12) as saying that $\kappa < 2^\kappa$ for all cardinal numbers κ. Thus the cardinal numbers are unbounded.

Furthermore, Theorem 10.2.18 allows us to express the cardinality of the continuum $\mathfrak{c}$ in terms of $\aleph_0$.

Corollary 10.2.19
$\mathfrak{c} = 2^{\aleph_0}$

Proof
We proved in Theorem 10.1.21 that $|\mathscr{P}(\mathbb{N})| = \mathfrak{c}$, and so by Theorem 10.2.18 we have

$$\mathfrak{c} = |\mathscr{P}(\mathbb{N})| = 2^{|\mathbb{N}|} = 2^{\aleph_0}$$

as claimed. □

Exercise 10.2.20
Prove that for all cardinal numbers κ, λ, μ, if $\lambda \leqslant \mu$, then $\lambda^\kappa \leqslant \mu^\kappa$.

Many of the properties satisfied by exponentiation of natural numbers generalise to cardinal numbers.

Theorem 10.2.21 · *Properties of cardinal exponentiation*

(a) $\mu^{\kappa+\lambda} = \mu^\kappa \cdot \mu^\lambda$ for all cardinal numbers κ, λ, μ;

(b) $(\mu \cdot \lambda)^\kappa = \mu^\kappa \cdot \lambda^\kappa$ for all cardinal numbers κ, λ, μ;

(c) $(\mu^\lambda)^\kappa = \mu^{\kappa \cdot \lambda}$ for all cardinal numbers κ, λ, μ.

Proof of (a)
Let κ, λ, μ be cardinal numbers and let X, Y and Z be sets with $|X| = \kappa$, $|Y| = \lambda$ and $|Z| = \mu$. Assume furthermore that Y and Z are disjoint.

Given a function $f : X \cup Y \to Z$, define $f_X : X \to Z$ and $f_Y : Y \to Z$ by $f_X(a) = f(a)$ for all $a \in X$, and $f_Y(a) = f(a)$ for all $y \in X$.

Define $H : Z^{X \cup Y} \to Z^X \times Z^Y$ by $H(f) = (f_X, f_Y)$. Then:

- H is injective. To see this, let $f, g : X \cup Y \to Z$ and suppose that $H(f) = H(g)$. Then $f_X = g_X$ and $f_Y = g_Y$. Now let $a \in X \cup Y$. Then:

- ◇ If $a \in X$, then $f(a) = f_X(a) = g_X(a) = g(a)$;
- ◇ If $a \in Y$, then $f(a) = f_Y(a) = g_Y(a) = g(a)$.

In both cases we have $f(a) = g(a)$, so that $f = g$ by function extensionality.

- H is surjective. To see this, let $(p,q) \in Z^X \times Z^Y$, so that we have $p : X \to Z$ and $q : Y \to Z$. Define $f : X \cup Y \to Z$ by

$$f(a) = \begin{cases} p(a) & \text{if } a \in X \\ q(a) & \text{if } a \in Y \end{cases}$$

Then f is well-defined since X and Y are disjoint. Moreover for all $a \in X$ we have $f_X(a) = p(a)$ and for all $a \in Y$ we have $f_Y(a) = q(a)$, so that $(p,q) = (f_X, f_Y) = H(f)$, as required.

Since H is a bijection, it follows that

$$\mu^{\kappa+\lambda} = |Z|^{|X|+|Y|} = |Z|^{|X \cup Y|} = |Z^{X \cup Y}| = |Z^X \times Z^Y| = |Z^X| \cdot |Z^Y| = \mu^\kappa \cdot \mu^\lambda$$

as required. □

Exercise 10.2.22
Prove parts (b) and (c) of Theorem 10.2.21.

Example 10.2.23
We prove that $\aleph_0^{\aleph_0} = 2^{\aleph_0}$. Indeed:

- We know that $\aleph_0 < 2^{\aleph_0}$ by Cantor's theorem (Theorem 10.1.12), so that:

$$\begin{aligned} \aleph_0^{\aleph_0} &\leqslant (2^{\aleph_0})^{\aleph_0} && \text{by Exercise 10.2.20} \\ &= 2^{\aleph_0 \cdot \aleph_0} && \text{by Theorem 10.2.21(c)} \\ &= 2^{\aleph_0} && \text{by Example 10.2.9} \end{aligned}$$

- Since $\aleph_0 > 2$, we have $2^{\aleph_0} \leqslant \aleph_0^{\aleph_0}$ by Exercise 10.2.20.

It follows from the Cantor–Schröder–Bernstein theorem (Theorem 10.1.14) that $\aleph_0^{\aleph_0} = 2^{\aleph_0}$.

Exercise 10.2.24
Prove that $\mathfrak{c}^{\mathfrak{c}} = 2^{\mathfrak{c}}$.

Exercise 10.2.25
Prove that $\aleph_0^{\aleph_0^{\aleph_0}} = 2^{\mathfrak{c}}$.

Indexed cardinal sums and products

Definition 10.2.26

Let $\{A_i \mid i \in I\}$ be a family of sets indexed by a set I. The (**indexed**) **disjoint union** of $\{A_i \mid i \in I\}$ is the set $\bigsqcup_{i \in I} A_i$ (LaTeX code: `\bigsqcup_{i \in I}`) defined by

$$\bigsqcup_{i \in I} A_i \;=\; \bigcup_{i \in I} (\{i\} \times A_i) \;=\; \{(i,a) \mid i \in I,\ a \in A_i\}$$

Note that for all $i, j \in I$ with $i \neq j$, the sets $\{i\} \times A_i$ and $\{j\} \times A_j$ are disjoint even if the sets A_i and A_j are not—hence the term *disjoint* union!

An element $(i,a) \in \bigsqcup_{i \in I} A_i$ can be thought of as simply being an element $a \in A_i$, but keeping track of the label i of the set A_i.

Example 10.2.27

Given a set I and a set A, we have

$$\bigsqcup_{i \in I} A \;=\; \bigcup_{i \in I} (\{i\} \times A) \;=\; I \times A$$

Thus the disjoint union of 'I-many' copies of a set A is simply $I \times A$.

Exercise 10.2.28

Let $\{X_i \mid i \in I\}$ be a family of sets indexed by a set I, and define a function

$$q : \bigsqcup_{i \in I} X_i \to \bigcup_{i \in I} X_i$$

by $q(i,a) = a$ for all $i \in I$ and $a \in X_i$. Prove that if the sets X_i for $i \in I$ are pairwise disjoint, then q is a bijection.

Definition 10.2.29

Let $\{\kappa_i \mid i \in I\}$ be an indexed family of cardinal numbers. The (**indexed**) **cardinal sum** of κ_i for $i \in I$ is defined by

$$\sum_{i \in I} \kappa_i = \left| \bigsqcup_{i \in I} [\kappa_i] \right|$$

That is, the indexed cardinal sum is the cardinality of the indexed disjoint union.

As with the finitary operations, we should check that this agrees with the definition of addition for natural numbers. And indeed it does—given a finite set I and natural numbers n_i for each $i \in I$, the fact that $\left| \bigsqcup_{i \in I} [n_i] \right| = \sum_{i \in I} n_i$ is an immediate consequence of the addition principle (Theorem 8.1.26).

Example 10.2.30
By Example 10.2.27 we have

$$\sum_{i\in I}\kappa = |I\times[\kappa]| = |I|\cdot\kappa$$

for all sets I and all cardinal numbers κ.

Example 10.2.31
We prove that $\sum_{n\in\mathbb{N}} n = \aleph_0$.

Define a function $f : \bigsqcup_{n\in\mathbb{N}}[n] \to \mathbb{N}\times\mathbb{N}$ by $f(n,k) = (n,k)$ for all $n\in\mathbb{N}$ and $k\in[n]$. Evidently f is injective, since it is the inclusion function of a subset. Therefore

$$\sum_{n\in\mathbb{N}} n \;\leqslant\; |\mathbb{N}\times\mathbb{N}| \;=\; \aleph_0\cdot\aleph_0 \;=\; \aleph_0$$

Define a function $g : \mathbb{N} \to \bigsqcup_{n\in\mathbb{N}}[n]$ by $g(n) = (n+1,0)$ for all $n\in\mathbb{N}$. Then g is an injection, since given $m,n\in\mathbb{N}$, if $g(m) = g(n)$, then $(m+1,0) = (n+1,0)$, and so $m+1 = n+1$. Thus $m = n$, as required. So we have

$$\aleph_0 \;=\; |\mathbb{N}| \;\leqslant\; \left|\bigsqcup_{n\in\mathbb{N}}[n]\right| \;=\; \sum_{n\in\mathbb{N}} n$$

By the Cantor–Schröder–Bernstein theorem, we have $\sum_{n\in\mathbb{N}} n = \aleph_0$.

Exercise 10.2.32
Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of natural numbers, and let $I = \{n\in\mathbb{N} \mid a_n > 0\}$. Prove that

$$\sum_{n\in\mathbb{N}} a_n = \begin{cases} \aleph_0 & \text{if } I \text{ is infinite} \\ \sum_{k=1}^{n} a_{n_k} & \text{if } I = \{n_k \mid k\in[n]\} \text{ is finite} \end{cases}$$

Lemma 10.2.33
Let $\{X_i \mid i\in I\}$ be a family of pairwise disjoint sets, indexed by a set I. Then

$$\left|\bigcup_{i\in I} X_i\right| \;=\; \sum_{i\in I}|X_i|$$

Proof
For each $i\in I$ let $\kappa_i = |X_i|$, and let $f_i : [\kappa_i] \to X_i$ be a bijection. Then the function

$$f : \bigsqcup_{i\in I}[\kappa_i] \to \bigsqcup_{i\in I} X_i$$

defined by $f(i,a) = f_i(a)$ for all $i \in I$ and $a \in [\kappa_i]$ is a bijection, since it has an inverse g given by $g(i,a) = f_i^{-1}(a)$ for all $i \in I$ and $a \in X_i$.

Since the sets X_i are pairwise disjoint, we have by Exercise 10.2.28 that there is a bijection $\bigsqcup_{i\in I} X_i \to \bigcup_{i\in I} X_i$. Hence

$$\left|\bigcup_{i\in I} X_i\right| = \left|\bigsqcup_{i\in I} X_i\right| = \left|\bigsqcup_{i\in I} [\kappa_i]\right| = \sum_{i\in I} \kappa_i = \sum_{i\in I} |X_i|$$

as required. □

Definition 10.2.34

Let $\{X_i \mid i \in I\}$ be family of sets indexed by a set I. The **(indexed) cartesian product** of the sets X_i for $i \in I$ is defined by

$$\prod_{i\in I} X_i = \left\{ f : I \to \bigcup_{i\in I} X_i \;\middle|\; f(i) \in X_i \text{ for all } i \in I \right\}$$

An element $f \in \prod_{i\in I} X_i$ is called a **choice function** for the family $\{X_i \mid i \in I\}$.

It is worth pointing out that the axiom of choice (Axiom 3.2.33) says precisely that the cartesian product of every family of inhabited sets is inhabited.

Example 10.2.35

Given a set I and a set X, a choice function for $\{X \mid i \in I\}$ is a function $f : I \to \bigcup_{i\in I} X = X$ such that $f(i) \in X$ for all $i \in I$. But then every function $f : I \to X$ is a choice function, and so

$$\prod_{i\in I} X = X^I$$

which is the set of functions $I \to X$.

Definition 10.2.36

Let $\{\kappa_i \mid i \in I\}$ be an indexed family of cardinal numbers. The **(indexed) cardinal product** of κ_i for $i \in I$ is defined by

$$\prod_{i\in I} \kappa_i = \left|\prod_{i\in I} [\kappa_i]\right|$$

That is, the indexed cardinal product is the cardinality of the indexed cartesian product.

Example 10.2.37

By Example 10.2.35, we have

$$\prod_{i\in I} \kappa = |[\kappa]|^{|I|} = \kappa^{|I|}$$

for all sets I and all cardinal numbers κ.

Exercise 10.2.38

Prove that $1 \times 2 \times 3 \times 4 \times \cdots = 2^{\aleph_0}$.

Exercise 10.2.39

Prove that there do not exist cardinal numbers $\{\kappa_n \mid n \in \mathbb{N}\}$ such that $\kappa_n \neq 1$ for all $n \in \mathbb{N}$ and $\prod_{n\in\mathbb{N}} \kappa_n = \aleph_0$.

Section 10.E

Chapter 10 exercises

Countability

In Questions 10.E.1 to 10.E.4, prove that the set is countable.

10.E.1. The set $D = \left\{ \frac{a}{2^n} \mid a \in \mathbb{Z}, n \in \mathbb{N} \right\}$ of *dyadic* rational numbers.

10.E.2. The set of all functions $[n] \to \mathbb{Z}$, where $n \in \mathbb{N}$.

10.E.3. The set of all real numbers whose square is rational.

10.E.4. The following set:

$$[(\mathbb{Z} \times \mathbb{Q}) \setminus (\mathbb{N} \times \mathbb{Z})] \cup \{n \in \mathbb{N} \mid \exists u, v \in \mathbb{N}, n = 5u + 6v\} \cup \{x \in \mathbb{R} \mid x - \sqrt{2} \in \mathbb{Q}\}$$

In Questions 10.E.5 to 10.E.7, prove that the set is uncountable.

10.E.5. The set of all functions $\mathbb{Z} \to [2]$.

10.E.6. The set of all subsets $U \subseteq \mathbb{N}$ such that neither U nor $\mathbb{N} \setminus U$ is finite.

10.E.7. The set of all sequences of rational numbers that converge to 0.

In Questions 10.E.8 to 10.E.12, determine whether the set is countable or uncountable, and then prove it.

10.E.8. The set of all functions $f : \mathbb{N} \to \mathbb{N}$ that are weakly decreasing—that is, such that for all $m, n \in \mathbb{N}$, if $m \leqslant n$, then $f(m) \geqslant f(n)$.

10.E.9. The set of all functions $f : \mathbb{N} \to \mathbb{Z}$ that are weakly decreasing.

10.E.10. The set of all periodic functions $f : \mathbb{Z} \to \mathbb{Q}$—that is, such that there is some integer $p > 0$ such that $f(x + p) = f(x)$ for all $x \in \mathbb{Z}$.

10.E.11. The set of all periodic functions $f : \mathbb{Q} \to \mathbb{Z}$—that is, such that there is some rational number $p > 0$ such that $f(x + p) = f(x)$ for all $x \in \mathbb{Q}$.

10.E.12. The set of all real numbers x such that $a_0 + a_1 x + \cdots + a_d x^d = 0$ for some rational numbers $a_0, a_1, \ldots, a_d$ with $a_d \neq 0$.

10.E.13. A subset $D \subseteq \mathbb{R}$ is *dense* if $(a - \varepsilon, a + \varepsilon) \cap U$ is inhabited for all $a \in \mathbb{R}$ and all $\varepsilon > 0$—intuitively, this means that there are elements of U arbitrarily close to any real number. Must a dense subset of $\mathbb{R}$ be uncountable?

Cardinality

In Questions 10.E.14 to 10.E.20, determine whether the cardinality of the set is equal to $\aleph_0$, equal to $\mathfrak{c}$, or greater than $\mathfrak{c}$.

10.E.14. $\mathscr{P}(\mathbb{R})$.

10.E.15. The set of all finite subsets of $\mathbb{R}$.

10.E.16. The set of all countably infinite subsets of $\mathbb{R}$.

10.E.17. $\mathbb{R} \times \mathbb{R}$.

10.E.18. The set of all functions $f : \mathbb{N} \to \mathbb{R}$ such that $f(n) = 0$ for all but finitely many $n \in \mathbb{N}$.

10.E.19. The set of all functions $\mathbb{Q} \to \mathbb{R}$.

10.E.20. The set of all functions $\mathbb{R} \to \mathbb{R}$.

10.E.21. Prove that there is a set $\mathscr{C}$ of pairwise disjoint circles in $\mathbb{R}^2$ such that $|\mathscr{C}| = \mathfrak{c}$; formally, a circle is a subset $C \subseteq \mathbb{R}^2$ of the form

$$C = \{(x,y) \mid (x-a)^2 + (y-b)^2 = r^2\}$$

for some $a, b \in \mathbb{R}$ and some $r > 0$.

10.E.22. Prove that there does not exist a set $\mathscr{D}$ of pairwise disjoint discs in $\mathbb{R}^2$ such that $|\mathscr{D}| = \mathfrak{c}$; formally, a disc is a subset $D \subseteq \mathbb{R}^2$ of the form

$$D = \{(x,y) \mid (x-a)^2 + (y-b)^2 \leqslant r^2\}$$

for some $a, b \in \mathbb{R}$ and some $r > 0$.

Cardinal arithmetic

10.E.23. Prove that $\kappa + 0 = \kappa \cdot 1 = \kappa^1 = \kappa$ for all cardinal numbers κ.

10.E.24. Prove that $\kappa^0 = 1^\kappa = 1$ for all cardinal numbers κ.

10.E.25. Let κ be an infinite cardinal number such that $\kappa \cdot \kappa = \kappa$. Prove that $\kappa^\kappa = 2^\kappa$.

10.E.26. Prove that $\mathfrak{c}^\mathfrak{c} = 2^{2^{\aleph_0}}$.

10.E.27. Let κ, λ and μ be cardinal numbers.

(a) Prove that if $\kappa \leqslant \lambda$, then $\kappa + \mu \leqslant \lambda + \mu$.

(b) Suppose that $\kappa + \mu \leqslant \lambda + \mu$. Must it be the case that $\kappa \leqslant \lambda$?

10.E.28. Let κ, λ and μ be cardinal numbers.

(a) Prove that if $\kappa \leqslant \lambda$, then $\kappa \cdot \mu \leqslant \lambda \cdot \mu$.

(b) Suppose that $\kappa \cdot \mu \leqslant \lambda \cdot \mu$ and $\mu > 0$. Must it be the case that $\kappa \leqslant \lambda$?

10.E.29. Let κ, λ and μ be cardinal numbers.

(a) Prove that if $\kappa \leqslant \lambda$, then $\kappa^\mu \leqslant \lambda^\mu$.

(b) Suppose that $\kappa^\mu \leqslant \lambda^\mu$ and $\mu > 0$. Must it be the case that $\kappa \leqslant \lambda$?

10.E.30. Let κ, λ and μ be cardinal numbers.

(a) Prove that if $\kappa \leqslant \lambda$, then $\mu^\kappa \leqslant \mu^\lambda$.

(b) Suppose that $\mu^\kappa \leqslant \mu^\lambda$ and $\mu > 1$. Must it be the case that $\kappa \leqslant \lambda$?

Definition 10.E.31
Given cardinal numbers κ and λ, define the **binomial coefficient** $\binom{\kappa}{\lambda}$ by

$$\binom{\kappa}{\lambda} = |\{U \subseteq [\kappa] \mid |U| = \lambda\}|$$

10.E.32. Let κ be a cardinal number. Prove that

$$\binom{\aleph_0}{\kappa} = \begin{cases} 1 & \text{if } \kappa = 0 \\ \aleph_0 & \text{if } \kappa \in \mathbb{N} \text{ and } \kappa > 0 \\ 2^{\aleph_0} & \text{if } \kappa = \aleph_0 \\ 0 & \text{if } \kappa \notin \mathbb{N} \cup \{\aleph_0\} \end{cases}$$

10.E.33. Find the values of $\binom{\mathfrak{c}}{\kappa}$ for $\kappa \in \mathbb{N} \cup \{\aleph_0, \mathfrak{c}\}$.

10.E.34. Prove that $\sum_{\lambda \leqslant \kappa} \binom{\kappa}{\lambda} = 2^\kappa$ for all cardinal numbers κ.

10.E.35. Define the factorial $\kappa!$ of a cardinal number κ in a way that generalises the notion of factorial $n!$ for a natural number n. Find expressions for $\aleph_0!$ and $\mathfrak{c}!$.

True–False questions

In Questions 10.E.36 to 10.E.46, determine (with proof) whether the statement is true or false.

10.E.36. For any two countably infinite sets X and Y, there exists a bijection $X \to Y$.

10.E.37. For any two uncountable sets X and Y, there exists a bijection $X \to Y$.

10.E.38. The product of countably many countable sets is countable.

10.E.39. Every countably infinite set can be partitioned into infinitely many infinite subsets.

10.E.40. There are countably many infinite subsets of $\mathbb{N}$.

10.E.41. The set $\mathbb{N} \times \mathbb{Z} \times \mathbb{Q}$ is countably infinite.

10.E.42. $\aleph_0$ is the smallest cardinal number.

10.E.43. There exist sets X and Y of different cardinalities that have the same number of finite subsets.

10.E.44. There exist sets X and Y of different cardinalities that have the same number of countably infinite subsets.

10.E.45. For all sets X, if $U \subsetneqq X$, then $|U| < |X|$.

10.E.46. $2^\kappa < 3^\kappa$ for all cardinal numbers $\kappa > 0$.

Always–Sometimes–Never questions

In Questions 10.E.47 to 10.E.52, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

10.E.47. Let X be a set and suppose that there exists an injection $\mathbb{N} \to X$. Then there exists an injection $\mathbb{Q} \to X$.

10.E.48. Let X be a set. Then $\mathscr{P}(X)$ is countably infinite.

10.E.49. Let X be a set, let κ and λ be cardinal numbers, and suppose that $\lambda \leqslant \kappa = |X|$. Then X has a subset of size λ.

10.E.50. Let X and Y be sets with $|X| \leqslant |Y|$. Then $|\mathscr{P}(X)| \leqslant |\mathscr{P}(Y)|$.

10.E.51. Let $n \in \mathbb{N}$ with $n \geqslant 2$. Then $\aleph_0^n > \aleph_0$.

10.E.52. Let κ be a cardinal number with $\kappa \geqslant \aleph_0$. Then $\aleph_0^\kappa > \aleph_0$.

Trick question

10.E.53. Does there exist a cardinal number κ such that $\aleph_0 < \kappa < 2^{\aleph_0}$?

Chapter 11

Discrete probability theory

Section 11.1

Discrete probability spaces

Probability theory is a field of mathematics which attempts to model randomness and uncertainty in the 'real world'. The mathematical machinery it develops allows us to understand how this randomness behaves and to extract information which is useful for making predictions.

Discrete probability theory, in particular, concerns situations in which the possible outcomes form a *countable* set. This simplifies matters considerably: if there are only countably many outcomes, then the probability that any event occurs is determined entirely by the probabilities that the individual outcomes comprised by the event occur.

For example, the number N of words spoken by a child over the course of a year takes values in $\mathbb{N}$, so is discrete. To each $n \in \mathbb{N}$, we may assign a probability that $N = n$, which can take positive values in a meaningful way, and from these probabilities we can compute the probabilities of more general events occurring (e.g. the probability that the child says under a million words). However, the height H grown by the child over the same period takes values in $[0,\infty)$, which is uncountable; for each $h \in [0,\infty)$, the probability that $H = h$ is zero, so these probabilities give us no information. We must study the behaviour of H through some other means.

In this chapter, we will concern ourselves only with the discrete setting.

It is important to understand from the outset that, although we use language like *outcome*, *event*, *probability* and *random*, and although we use real-world examples, everything we do concerns mathematical objects: sets, elements of sets, and functions. If we say, for example, "the probability that a roll of a fair six-sided die shows 3 or 4 is $\frac{1}{3}$," we are actually interpreting the situation mathematically—the *outcomes* of the die rolls are interpreted as the elements of the set $\{1,2,3,4,5,6\}$; the *event* that the die shows 3 or 4 is interpreted as the subset $\{3,4\} \subseteq \{1,2,3,4,5,6\}$; and the *probability* that this event occurs is the value of a particular function $\mathbb{P} : \mathscr{P}(\{1,2,3,4,5,6\}) \to [0,1]$ on input $\{3,4\}$. The mathematical interpretation is called a **model** of the real-world situation.

Definition 11.1.1

A **discrete probability space** is a pair $(\Omega, \mathbb{P})$ (LaTeX code: `(\Omega, \mathbb{P})`), consisting of a countable set Ω and a function $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$, such that

(i) $\mathbb{P}(\Omega) = 1$; and

(ii) (**Countable additivity**) If $\{A_i \mid i \in I\}$ is any family of pairwise disjoint subsets of Ω, indexed by a countable set I, then

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbb{P}(A_i)$$

The set Ω is called the **sample space**; the elements $\omega \in \Omega$ are called **outcomes**;[a] the subsets $A \subseteq \Omega$ are called **events**; and the function $\mathbb{P}$ is called the **probability measure**. Given an event A, the value $\mathbb{P}(A)$ is called the **probability of** A.

There is a general notion of a probability space, which does not require the sample space Ω to

be countable. This definition is significantly more technical, so we restrict our attention in this section to *discrete* probability spaces. Thus, whenever we say 'probability space' in this section, the probability space can be assumed to be discrete. However, when our proofs do not specifically use countability of Ω, they typically are true of arbitrary probability spaces. As such, we will specify discreteness in the statement of results only when countability of the sample space is required.

Example 11.1.2

We model the roll of a fair six-sided die.

The possible **outcomes** of the roll are 1, 2, 3, 4, 5 and 6, so we can take $\Omega = \{1,2,3,4,5,6\}$ to be the sample space.

The **events** correspond with subsets of Ω. For example:

- $\{4\}$ is the event that the die roll shows 4. This event occurs with probability $\frac{1}{6}$.
- $\{1,3,5\}$ is the event that the die roll is odd. This event occurs with probability $\frac{1}{2}$
- $\{1,4,6\}$ is the event that the die roll is not prime. This event occurs with probability $\frac{1}{2}$.
- $\{3,4,5,6\}$ is the event that the die roll shows a number greater than 2. This event occurs with probability $\frac{2}{3}$.
- $\{1,2,3,4,5,6\}$ is the event that anything happens. This event occurs with probability 1.
- $\varnothing$ is the event that nothing happens. This event occurs with probability 0.

More generally, since each outcome occurs with equal probability $\frac{1}{6}$, we can define

$$\mathbb{P}(A) = \frac{|A|}{6} \text{ for all events } A$$

We will verify that $\mathbb{P}$ defines a probability measure on Ω in Example 11.1.6.

Example 11.1.3

Let $(\Omega, \mathbb{P})$ be a probability space. We prove that $\mathbb{P}(\varnothing) = 0$.

Note that Ω and $\varnothing$ are disjoint, so by countable additivity, we have

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \varnothing) = \mathbb{P}(\Omega) + \mathbb{P}(\varnothing) = 1 + \mathbb{P}(\varnothing)$$

Subtracting 1 throughout yields $\mathbb{P}(\varnothing) = 0$, as required.

Exercise 11.1.4

Let $(\Omega, \mathbb{P})$ be a probability space. Prove that

$$\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$$

for all events A.

Countable additivity of probability measures—that is, condition (ii) in Definition 11.1.1—implies

that probabilities of events are determined by probabilities of individual outcomes. This is made precise in Proposition 11.1.5.

Proposition 11.1.5

Let Ω be a countable set and let $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$ be a function such that $\mathbb{P}(\Omega) = 1$. The following are equivalent:

(i) $\mathbb{P}$ is a probability measure on Ω;

(ii) $\sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \mathbb{P}(A)$ for all $A \subseteq \Omega$.

Proof

Since $\mathbb{P}(\Omega) = 1$, it suffices to prove that condition (ii) of Proposition 11.1.5 is equivalent to countable additivity of $\mathbb{P}$.

- (i)⇒(ii). Suppose $\mathbb{P}$ is a probability measure on Ω. Let $A \subseteq \Omega$.

  Note that since $A \subseteq \Omega$ and Ω is countably infinite, it follows that $\{\{\omega\} \mid \omega \in A\}$ is a countable family of pairwise disjoint sets. By countable additivity, we have

  $$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})$$

  as required. Hence condition (ii) of the proposition is satisfied.

- (ii)⇒(i). Suppose that $\sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \mathbb{P}(A)$ for all $A \subseteq \Omega$. We prove that $\mathbb{P}$ is a probability measure on Ω.

  So let $\{A_i \mid i \in I\}$ be a family of pairwise disjoint events, indexed by a countable set I. Define $A = \bigcup_{i \in I} A_i$. Since the sets A_i partition A, summing over elements of A is the same as summing over each of the sets A_i individually, and then adding those results together; specifically, for each A-tuple $(p_\omega)_{\omega \in A}$, we have

  $$\sum_{\omega \in A} p_\omega = \sum_{i \in I} \sum_{\omega \in A_i} p_\omega$$

  Hence

  $$\begin{aligned} \mathbb{P}(A) &= \sum_{\omega \in A} \mathbb{P}(\{\omega\}) && \text{by condition (ii) of the proposition} \\ &= \sum_{i \in I} \sum_{\omega \in A_i} \mathbb{P}(\{\omega\}) && \text{by the above observation} \\ &= \sum_{i \in I} \mathbb{P}(A_i) && \text{by condition (ii) of the proposition} \end{aligned}$$

  So $\mathbb{P}$ satisfies the countable additivity condition. Thus $\mathbb{P}$ is a probability measure on Ω.

Hence the two conditions are equivalent. □

Example 11.1.6

We prove that the function $\mathbb{P}$ from Example 11.1.2 truly does define a probability measure. Indeed, let $\Omega = \{1,2,3,4,5,6\}$ and let $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$ be defined by

$$\mathbb{P}(A) = \frac{|A|}{6} \text{ for all events } A$$

Then $\mathbb{P}(\Omega) = \frac{6}{6} = 1$, so condition (i) in Definition 11.1.1 is satisfied. Moreover, for each $A \subseteq \Omega$ we have

$$\sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \sum_{\omega \in A} \frac{1}{6} = \frac{|A|}{6} = \mathbb{P}(A)$$

so, by Proposition 11.1.5, $\mathbb{P}$ defines a probability measure on Ω.

Proposition 11.1.5 makes defining probability measures much easier, since it implies that probability measures are determined entirely by their values on individual outcomes. This means that, in order to define a probability measure, we only need to specify its values on individual outcomes and check that the sum of these probabilities is equal to 1. This is significantly easier than defining $\mathbb{P}(A)$ on *all* events $A \subseteq \Omega$ and checking the two conditions of Definition 11.1.1.

This is made precise in Proposition 11.1.7 below.

Proposition 11.1.7

Let Ω be a countable set and, for each $\omega \in \Omega$, let $p_\omega \in [0,1]$. If $\sum_{\omega \in \Omega} p_\omega = 1$, then there is a unique probability measure $\mathbb{P}$ on Ω such that $\mathbb{P}(\{\omega\}) = p_\omega$ for each $\omega \in \Omega$.

Proof

We prove existence and uniqueness of $\mathbb{P}$ separately.

- **Existence.** Define $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$ be defined by

  $$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$$

  for all events $A \subseteq \Omega$. Then condition (ii) of Proposition 11.1.5 is automatically satisfied, and indeed $\mathbb{P}(\{\omega\}) = p_\omega$ for each $\omega \in \Omega$. Moreover

  $$\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} p_\omega = 1$$

  and so condition (i) of Definition 11.1.1 is satisfied. Hence $\mathbb{P}$ defines a probability measure on Ω.

- **Uniqueness.** Suppose that $\mathbb{P}' : \mathscr{P}(\Omega) \to [0,1]$ is another probability measure such that $\mathbb{P}'(\{\omega\}) = p_\omega$ for all $\omega \in \Omega$. For each event $A \subseteq \Omega$, condition (ii) of Proposition 11.1.5 implies that

  $$\mathbb{P}'(A) = \sum_{\omega \in A} \mathbb{P}'(\{\omega\}) = \sum_{\omega \in A} p_\omega = \mathbb{P}(A)$$

  hence $\mathbb{P}' = \mathbb{P}$.

So $\mathbb{P}$ is uniquely determined by the values p_ω. □

The assignments of $p_\omega \in [0,1]$ to each $\omega \in \Omega$ in fact defines something that we will later defined to be a *probability mass function* (Definition 11.2.5).

With Proposition 11.1.7 proved, we will henceforth specify probability measures $\mathbb{P}$ on sample spaces Ω by specifying only the values of $\mathbb{P}(\{\omega\})$ for $\omega \in \Omega$.

Example 11.1.8

Let $p \in [0,1]$. A coin, which shows heads with probability p, is repeatedly flipped until heads shows.

The outcomes of such a sequence of coin flips all take the form

$$(\underbrace{\text{tails}, \text{tails}, \cdots, \text{tails}}_{n \text{ 'tails'}}, \text{heads})$$

for some $n \in \mathbb{N}$. Identifying such a sequence with the number n of flips before heads shows, we can take $\Omega = \mathbb{N}$ to be the sample space.

For each $n \in \mathbb{N}$, we can define

$$\mathbb{P}(\{n\}) = (1-p)^n p$$

This will define a probability measure on $\mathbb{N}$, provided these probabilities all sum to 1; and indeed by Theorem 9.3.8, we have

$$\sum_{n \in \mathbb{N}} \mathbb{P}(\{n\}) = \sum_{n \in \mathbb{N}} (1-p)^n p = p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1$$

By Proposition 11.1.7, it follows that $(\Omega, \mathbb{P})$ is a probability space.

Exercise 11.1.9

A fair six-sided die is rolled twice. Define a probability space $(\Omega, \mathbb{P})$ that models this situation.

Exercise 11.1.10

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events with $A \subseteq B$. Prove that $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

Set operations on events

In the real world, we might want to talk about the probability that two events both happen, or the probability that an event doesn't happen, or the probability that at least one of some collection of events happens. This is interpreted mathematically in terms of set operations.

Example 11.1.11

Let $(\Omega, \mathbb{P})$ be the probability space modelling two rolls of a fair six-sided die—that is, the sample space $\Omega = \{1,2,3,4,5,6\}^2$ with probability measure $\mathbb{P}$ defined by $\mathbb{P}(\{(a,b)\}) = \frac{1}{36}$ for each $(a,b) \in \Omega$.

Let A be the event that the sum of the die rolls is even, that is

$$A = \left\{ \begin{matrix} (1,1), & (1,3), & (1,5), & (2,2), & (2,4), & (2,6), \\ (3,1), & (3,3), & (3,5), & (4,2), & (4,4), & (4,6), \\ (5,1), & (5,3), & (5,5), & (6,2), & (6,4), & (6,6) \end{matrix} \right\}$$

and let B be the event that the sum of the die rolls is greater than or equal to 9, that is

$$B = \{(3,6),(4,5),(4,6),(5,4),(5,5),(5,6),(6,3),(6,4),(6,5),(6,6)\}$$

Then

- Consider the event that the sum of the die rolls is even **or** greater than or equal to 9. An outcome ω gives rise to this event precisely when either $\omega \in A$ or $\omega \in B$; so the event in question is $A \cup B$;
- Consider the event that the sum of the die rolls is even **and** greater than or equal to 9. An outcome ω gives rise to this event precisely when both $\omega \in A$ and $\omega \in B$; so the event in question is $A \cap B$;
- Consider the event that the sum of the die rolls is **not** even. An outcome ω gives rise to this event precisely when $\omega \notin A$; so the event in question is $\Omega \setminus A$.

Thus we can interpret 'or' as union, 'and' as intersection, and 'not' as relative complement in the sample space.

The intuition provided by Example 11.1.11 is formalised in Exercise 11.1.13. Before we do this, we adopt a convention that simplifies notation when discussing events in probability spaces.

Definition 11.1.12

Let $(\Omega, \mathbb{P})$ be a probability space. The **complement** of an event $A \subseteq \Omega$ is the event $\Omega \setminus A \subseteq \Omega$. We write A^c (LaTeX code: `A^c`) for $\Omega \setminus A$.

That is, when we talk about the complement *of an event*, we really mean their relative complement inside the sample space.

Exercise 11.1.13

Let $(\Omega, \mathbb{P})$ be a probability space, and let $p(\omega), q(\omega)$ be logical formulae with free variable ω ranging over Ω. Let

$$A = \{\omega \in \Omega \mid p(\omega)\} \quad \text{and} \quad B = \{\omega \in \Omega \mid q(\omega)\}$$

Prove that

- $\{\omega \in \Omega \mid p(\omega) \wedge q(\omega)\} = A \cap B$;
- $\{\omega \in \Omega \mid p(\omega) \vee q(\omega)\} = A \cup B$;
- $\{\omega \in \Omega \mid \neg p(\omega)\} = A^c$.

For reference, in Example 11.1.11, we had $\Omega = \{1,2,3,4,5,6\}^2$ and we defined $p(a,b)$ to be '$a+b$ is even' and $q(a,b)$ to be '$a+b \geqslant 7$'.

With this in mind, it will be useful to know how set operations on events interact with probabilities. Characteristic functions (Definition 3.1.23) are very useful for this purpose. Translating Theorem 3.1.27 to the context of probability, if $(\Omega, \mathbb{P})$ is a probability space, $A, B \subseteq \Omega$ are events and $\omega \in \Omega$ is an outcome, then:

(i) $\chi_{A\cap B}(\omega) = \chi_A(\omega)\chi_B(\omega)$;

(ii) $\chi_{A\cup B}(\omega) = \chi_A(\omega) + \chi_B(\omega) - \chi_{A\cap B}(\omega)$; and

(iii) $\chi_{A^c}(\omega) = 1 - \chi_A(\omega)$.

The formula in the result of the following exercise allows us to express probabilities of events in terms of probabilities of individual outcomes and characteristic functions.

Exercise 11.1.14

Let $(\Omega, \mathbb{P})$ be a discrete probability space, and for each $\omega \in \Omega$ let $p_\omega = \mathbb{P}(\{\omega\})$. Prove that, for each event A, we have

$$\mathbb{P}(A) = \sum_{\omega\in\Omega} p_\omega \chi_A(\omega)$$

Theorem 11.1.15

Let $(\Omega, \mathbb{P})$ be a probability space and let $A, B \subseteq \Omega$. Then

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Proof

For each $\omega \in \Omega$, let $p_\omega = \mathbb{P}(\{\omega\})$. Then

$$\begin{aligned}
\mathbb{P}(A \cup B) &= \sum_{\omega\in\Omega} p_\omega \chi_{A\cup B}(\omega) && \text{by Exercise 11.1.14} \\
&= \sum_{\omega\in\Omega} p_\omega (\chi_A(\omega) + \chi_B(\omega) - \chi_{A\cap B}(\omega)) && \text{by Theorem 3.1.27(ii)} \\
&= \sum_{\omega\in\Omega} p_\omega \chi_A(\omega) + \sum_{\omega\in\Omega} p_\omega \chi_B(\omega) + \sum_{\omega\in\Omega} p_\omega \chi_{A\cap B}(\omega) && \text{rearranging} \\
&= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) && \text{by Exercise 11.1.14}
\end{aligned}$$

as required. □

Although there are nice expressions for unions and complements of events, it is not always the case that intersection of events corresponds with multiplication of probabilities.

Example 11.1.16

Let $\Omega = [3]$ and define a probability measure $\mathbb{P}$ on Ω by letting

$$\mathbb{P}(\{0\}) = \frac{1}{4}, \qquad \mathbb{P}(\{1\}) = \frac{1}{2} \qquad \text{and} \qquad \mathbb{P}(\{2\}) = \frac{1}{4}$$

Then we have

$$\mathbb{P}(\{0,1\}\cap\{1,2\}) = \mathbb{P}(\{1\}) = \frac{1}{2} \neq \frac{9}{16} = \frac{3}{4}\cdot\frac{3}{4} = \mathbb{P}(\{0,1\})\cdot\mathbb{P}(\{1,2\})$$

This demonstrates that it is not always the case that $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for events A, B in a given probability space. Pairs of events A, B for which this equation *is* true are said to be *independent*.

Definition 11.1.17

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events. We say that A and B are **independent** if $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$; otherwise, we say they are **dependent**. More generally, events $A_1, A_2, \ldots, A_n$ are **pairwise independent** if $\mathbb{P}(A_i\cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for all $i \neq j$.

Example 11.1.18

A fair six-sided die is rolled twice. Let A be the event that the first roll shows 4, and let B be the event that the second roll is even. Then

$$A = \{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\} = \{4\}\times\{1,2,3,4,5,6\}$$

so $\mathbb{P}(A) = \frac{6}{36} = \frac{1}{6}$; and

$$B = \{1,2,3,4,5,6\}\times\{2,4,6\}$$

which has size $6\times 3 = 18$, so $\mathbb{P}(B) = \frac{18}{36} = \frac{1}{2}$.

Moreover $A\cap B = \{(4,2),(4,4),(4,6)\}$, so it follows that

$$\mathbb{P}(A\cap B) = \frac{3}{36} = \frac{1}{12} = \frac{1}{6}\cdot\frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B)$$

so the events A and B are independent.

Let C be the event that the sum of the two dice rolls is equal to 5. Then

$$C = \{(1,4),(2,3),(3,2),(4,1)\}$$

so $\mathbb{P}(C) = \frac{4}{36} = \frac{1}{9}$. Moreover $A\cap C = \{(4,1)\}$, so it follows that

$$\mathbb{P}(A\cap C) = \frac{1}{36} \neq \frac{1}{54} = \frac{1}{6}\cdot\frac{1}{9} = \mathbb{P}(A)\mathbb{P}(C)$$

so the events A and C are dependent.

Exercise 11.1.19

Let $(\Omega, \mathbb{P})$ be a probability space. Under what conditions is an event A independent from itself?

Conditional probability

Suppose we model a random process, such as the roll of a die or the flip of a coin, using a probability space $(\Omega, \mathbb{P})$. When we receive new information, the situation might change, and we might want

to model this new situation by updating our probabilities to reflect the fact that we know that B has occurred. This is done by defining a new probability measure $\widetilde{\mathbb{P}}$ on Ω. What follows is an example of this.

Example 11.1.20

Two cards are drawn at random, in order, without replacement, from a 52-card deck. We can model this situation by letting the sample space Ω be the set of ordered pairs of distinct cards, and letting $\mathbb{P}$ assign an equal probability (of $\frac{1}{|\Omega|}$) to each outcome. Note that $|\Omega| = 52 \cdot 51$, and so

$$\mathbb{P}(\{\omega\}) = \frac{1}{52 \cdot 51}$$

for each outcome ω.

We will compute two probabilities:

- The probability that the second card drawn is a heart.
- The probability that the second card drawn is a heart *given that* the first card drawn is a diamond.

Let $A \subseteq \Omega$ be the event that the second card drawn is a heart, and let $B \subseteq \Omega$ be the event that the first card drawn is a diamond.

To compute $\mathbb{P}(A)$, note first that $A = A' \cup A''$, where

- A' is the event that both cards are hearts, so that $|A'| = 13 \cdot 12$; and
- A'' is the event that only the second card is a heart, so that $|A''| = 39 \cdot 13$.

Since $A' \cap A'' = \varnothing$, it follows from countable additivity that

$$\mathbb{P}(A) = \mathbb{P}(A') + \mathbb{P}(A'') = \frac{13 \cdot 12 + 39 \cdot 13}{52 \cdot 51} = \frac{13 \cdot (12 + 39)}{52 \cdot 51} = \frac{1}{4}$$

Now suppose we know that first card drawn is a diamond—that is, event B has occurred—and we wish to update our probability that A occurs. We do this by defining a new probability measure

$$\widetilde{\mathbb{P}} : \mathscr{P}(\Omega) \to [0,1]$$

such that:

(a) The outcomes that do not give rise to the event B are assigned probability zero; that is, $\widetilde{\mathbb{P}}(\{\omega\}) = 0$ for all $\omega \notin B$; and

(b) The outcomes that give rise to the event B are assigned probabilities proportional to their old probability; that is, there is some $k \in \mathbb{R}$ such that $\widetilde{\mathbb{P}}(\omega) = k\mathbb{P}(\omega)$ for all $\omega \in B$.

In order for $\widetilde{\mathbb{P}}$ to be a probability measure on Ω, we need condition (i) of Definition 11.1.1 to

occur.

$$\begin{aligned}
\widetilde{\mathbb{P}}(\Omega) &= \sum_{\omega\in\Omega} \widetilde{\mathbb{P}}(\{\omega\}) && \text{by condition (ii) of Proposition 11.1.5}\\
&= \sum_{\omega\in B} \widetilde{\mathbb{P}}(\{\omega\}) && \text{since } \widetilde{\mathbb{P}}(\{\omega\}) = 0 \text{ for } \omega \notin B\\
&= \sum_{\omega\in B} k\mathbb{P}(\{\omega\}) && \text{since } \widetilde{\mathbb{P}}(\{\omega\}) = k\mathbb{P}(\{\omega\} \text{ for } \omega \in B\\
&= k\mathbb{P}(B) && \text{by condition (ii) of Proposition 11.1.5}
\end{aligned}$$

Since we need $\widetilde{\mathbb{P}}(\Omega) = 1$, we must therefore take $k = \frac{1}{\mathbb{P}(B)}$. (In particular, we need $\mathbb{P}(B) > 0$ for this notion to be well-defined.)

Recall that, before we knew that the first card was a diamond, the probability that the second card is a heart was $\frac{1}{4}$. We now calculate how this probability changes with the updated information that the first card was a diamond.

The event that the second card is a heart in the new probability space is precisely $A \cap B$, since it is the subset of B consisting of all the outcomes ω giving rise to the event A. As such, the new probability that the second card is a heart is given by

$$\widetilde{\mathbb{P}}(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Now:

- $A \cap B$ is the event that the first card is a diamond and the second is a heart. To specify such an event, we need only specify the ranks of the two cards, so $|A \cap B| = 13 \cdot 13$ and hence $\mathbb{P}(A \cap B) = \frac{13\cdot 13}{52\cdot 51}$.
- B is the event that the first card is a diamond. A similar procedure as with A yields $\mathbb{P}(B) = \frac{1}{4}$.

Hence

$$\widetilde{\mathbb{P}}(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{13 \cdot 13 \cdot 4}{52 \cdot 51} = \frac{13}{51}$$

Thus the knowledge that the first card drawn is a diamond very slightly increases the probability that the second card is a heart from $\frac{1}{4} = \frac{13}{52}$ to $\frac{13}{51}$.

Example 11.1.20 suggests the following schema: upon discovering that an event B occurs, the probability that event A occurs should change from $\mathbb{P}(A)$ to $\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$. This motivates the following definition of *conditional probability*.

Definition 11.1.21

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events such that $\mathbb{P}(B) > 0$. The **conditional probability of A given B** is the number $\mathbb{P}(A \mid B)$ (LaTeX code: `\mathbb{P}(A \mid B)`) defined by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Example 11.1.22

A fair six-sided die is rolled twice. We compute the probability that the first roll showed a 2 given that the sum of the die rolls is less than 5.

We can model this situation by taking the sample space to be $\{1,2,3,4,5,6\}^2$, with each outcome having an equal probability of $\frac{1}{36}$.

Let A be the event that the first die roll shows a 2, that is

$$A = \{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\}$$

and let B be the event that the sum of the die rolls is less than 5, that is

$$B = \{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\}$$

We need to compute $\mathbb{P}(A \mid B)$. Well,

$$A \cap B = \{(2,1),(2,2)\}$$

so $\mathbb{P}(A \cap B) = \frac{2}{36}$; and $\mathbb{P}(B) = \frac{6}{36}$. Hence

$$\mathbb{P}(A \mid B) = \frac{\frac{2}{36}}{\frac{6}{36}} = \frac{2}{6} = \frac{1}{3}$$

Exercise 11.1.23

A fair six-sided die is rolled three times. What is the probability that the sum of the die rolls is less than or equal to 12, given that each die roll shows a power of 2?

Exercise 11.1.24

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events with $\mathbb{P}(B) > 0$. Prove that

$$\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B \mid B)$$

Exercise 11.1.25

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events such that $\mathbb{P}(B) > 0$. Prove that $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ if and only if A and B are independent.

We will soon see some useful real-world applications of probability theory using *Bayes's theorem* (Theorem 11.1.30). Before we do so, some technical results will be useful in our proofs.

Proposition 11.1.26

Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events with $0 < \mathbb{P}(B) < 1$. Then

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)$$

Proof
Note first that we can write

$$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$$

and moreover the events $A \cap B$ and $A \cap B^c$ are mutually exclusive. Hence

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$$

by countable additivity. The definition of conditional probability (Definition 11.1.21) then gives

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)$$

as required. □

Example 11.1.27

An animal rescue centre houses a hundred animals, sixty of which are dogs and forty of which are cats. Ten of the dogs and ten of the cats hate humans. We compute the probability that a randomly selected animal hates humans.

Let A be the event that a randomly selected animal hates humans, and let B be the event that the animal is a dog. Note that B^c is precisely the event that the animal is a cat. The information we are given says that:

- $\mathbb{P}(B) = \frac{60}{100}$, since 60 of the 100 animals are dogs;
- $\mathbb{P}(B^c) = \frac{40}{100}$, since 40 of the 100 animals are cats;
- $\mathbb{P}(A \mid B) = \frac{10}{60}$, since 10 of the 60 dogs hate humans;
- $\mathbb{P}(A \mid B^c) = \frac{10}{40}$, since 10 of the 40 cats hate humans.

By Proposition 11.1.26, it follows that the probability that a randomly selected animal hates humans is

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c) = \frac{60}{100} \cdot \frac{10}{60} + \frac{40}{100} \cdot \frac{10}{40} = \frac{20}{100} = \frac{1}{5}$$

The following example generalises Proposition 11.1.26 to arbitrary partitions of a sample space into events with positive probabilities.

Example 11.1.28

The animal rescue centre from Example 11.1.27 acquires twenty additional rabbits, of whom sixteen hate humans. We compute the probability that a randomly selected animal hates humans, given the new arrivals.

A randomly selected animal must be either a dog, a cat or a rabbit, and each of these occurs with positive probability. Thus, letting D be the event that the selected animal is a dog, C be the event that the animal is a cat, and R be the event that the animal is a rabbit, we see that the sets D, C, R form a partition of the sample space.

Letting A be the event that the selected animal hates humans. Then

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A \mid D)\mathbb{P}(D) + \mathbb{P}(A \mid C)\mathbb{P}(C) + \mathbb{P}(A \mid R)\mathbb{P}(R) \\ &= \frac{10}{60} \cdot \frac{60}{120} + \frac{10}{40} \cdot \frac{40}{120} + \frac{16}{20} \cdot \frac{20}{120} \\ &= \frac{3}{10}\end{aligned}$$

Proposition 11.1.29 below is a technical result which proves that conditional probability truly does yield a new probability measure on a given sample space.

Proposition 11.1.29

Let $(\Omega, \mathbb{P})$ be a probability space and let B be an event such that $\mathbb{P}(B) > 0$. The function $\widetilde{\mathbb{P}} : \mathscr{P}(\Omega) \to [0,1]$ defined by

$$\widetilde{\mathbb{P}}(A) = \mathbb{P}(A \mid B) \text{ for all } A \subseteq \Omega$$

defines a probability measure on Ω.

Proof

First note that

$$\widetilde{\mathbb{P}}(\Omega) = \mathbb{P}(\Omega \mid B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

so condition (i) of Definition 11.1.1 is satisfied.

Moreover, for each $A \subseteq \Omega$ we have

$$\begin{aligned}\widetilde{\mathbb{P}}(A) &= \mathbb{P}(A \mid B) && \text{by definition of } \widetilde{\mathbb{P}} \\ &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} && \text{by Definition 11.1.21} \\ &= \frac{1}{\mathbb{P}(B)} \sum_{\omega \in A \cap B} \mathbb{P}(\{\omega\}) && \text{by Proposition 11.1.5} \\ &= \sum_{\omega \in A \cap B} \mathbb{P}(\{\omega\} \mid B) && \text{by Definition 11.1.21} \\ &= \sum_{\omega \in A} \mathbb{P}(\{\omega\} \mid B) && \text{since } \mathbb{P}(\{\omega\} \mid B) = 0 \text{ for } \omega \in A \setminus B \\ &= \sum_{\omega \in A} \widetilde{\mathbb{P}}(\{\omega\}) && \text{by definition of } \widetilde{\mathbb{P}}\end{aligned}$$

so condition (ii) of Proposition 11.1.5 is satisfied. Hence $\widetilde{\mathbb{P}}$ defines a probability measure on Ω. □

Proposition 11.1.29 implies that we can use all the results we've proved about probability measures to conditional probability given a fixed event B. For example, Theorem 11.1.15 implies that

$$\mathbb{P}(A \cup A' \mid B) = \mathbb{P}(A \mid B) + \mathbb{P}(A' \mid B) - \mathbb{P}(A \cap A' \mid B)$$

for all events A, A', B in a probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(B) > 0$.

The next theorem we prove has a very short proof, but is extremely important in applied probability theory.

Theorem 11.1.30 · *Bayes's theorem*
Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events with positive probabilities. Then

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A)}$$

Proof
Definition 11.1.21 gives

$$\mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(B \cap A) = \mathbb{P}(B \mid A)\mathbb{P}(A)$$

Dividing through by $\mathbb{P}(A)$ yields the desired equation. □

As stated, Bayes's theorem is not necessarily particularly enlightening, but its usefulness increases sharply when combined with Proposition 11.1.26 to express the denominator of the fraction in another way.

Corollary 11.1.31
Let $(\Omega, \mathbb{P})$ be a probability space and let A, B be events such that $\mathbb{P}(A) > 0$ and $0 < \mathbb{P}(B) < 1$. Then

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)}$$

Proof
Bayes's theorem tells us that

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A)}$$

By Proposition 11.1.26 we have

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)$$

Substituting for $\mathbb{P}(A)$ therefore yields

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)}$$

as required. □

The following example is particularly counterintuitive.

Example 11.1.32

A town has 10000 people, 30 of whom are infected with Disease X. Medical scientists develop a test for Disease X, which is accurate 99% of the time. A person takes the test, which comes back positive. We compute the probability that the person truly is infected with Disease X.

Let A be the event that the person tests positive for Disease X, and let B be the event that the person is infected with Disease X. We need to compute $\mathbb{P}(B \mid A)$.

By Corollary 11.1.31, we have

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)}$$

It remains to compute the individual probabilities on the right-hand side of this equation. Well,

- $\mathbb{P}(A \mid B)$ is the probability that the person tests positive for Disease X, given that they are infected. This is equal to $\frac{99}{100}$, since the test is accurate with probability 99%.
- $\mathbb{P}(A \mid B^c)$ is the probability that the person tests positive for Disease X, given that they are *not* infected. This is equal to $\frac{1}{100}$, since the test is *inaccurate* with probability 1%.
- $\mathbb{P}(B) = \frac{30}{10000}$, since 30 of the 10000 inhabitants are infected with Disease X.
- $\mathbb{P}(B^c) = \frac{9970}{10000}$, since 9970 of the 10000 inhabitants are *not* infected with Disease X.

Piecing this together gives

$$\mathbb{P}(B \mid A) = \frac{\frac{99}{100} \cdot \frac{30}{10000}}{\frac{99}{100} \cdot \frac{30}{10000} + \frac{1}{100} \cdot \frac{9970}{10000}} = \frac{297}{1294} \approx 0.23$$

Remarkably, the probability that the person is infected with Disease X given that the test is positive is only 23%, even though the test is accurate 99% of the time!

The following result generalises Corollary 11.1.31 to arbitrary partitions of the sample space into sets with positive probabilities.

Corollary 11.1.33

Let $(\Omega, \mathbb{P})$ be a probability space, let A be an event with $\mathbb{P}(A) > 0$, and let $\{B_i \mid i \in I\}$ be a family of mutually exclusive events indexed by a countable set I such that

$$\mathbb{P}(B_i) > 0 \text{ for all } i \in I \qquad \text{and} \qquad \bigcup_{i \in I} B_i = \Omega$$

Then

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}{\sum_{i \in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}$$

for each $i \in I$.

Proof
Bayes's theorem tells us that

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}{\mathbb{P}(A)}$$

By countable additivity, we have

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i\in I} A \cap B_i\right) = \sum_{i\in I} \mathbb{P}(A \cap B_i) = \sum_{i\in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)$$

Substituting for $\mathbb{P}(A)$ therefore yields

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}{\sum_{i\in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)}$$

as required. □

Example 11.1.34
A small car manufacturer, *Cars N'At*, makes three models of car: the *Allegheny*, the *Monongahela* and the *Ohio*. It made 3000 Alleghenys, 6500 Monongahelas and 500 Ohios. In a given day, an Allegheny breaks down with probability $\frac{1}{100}$, a Monongahela breaks down with probability $\frac{1}{200}$, and the notoriously unreliable Ohio breaks down with probability $\frac{1}{20}$. An angry driver calls Cars N'At to complain that their car has broken down. We compute the probability that the driver was driving an Ohio.

Let A be the event that the car is an Allegheny, let B be the event that the car is a Monongahela, and let C be the event that the car is an Ohio. Then

$$\mathbb{P}(A) = \frac{3000}{10000}, \quad \mathbb{P}(B) = \frac{6500}{10000}, \quad \mathbb{P}(C) = \frac{500}{10000}$$

Let D be the event that the car broke down. Then

$$\mathbb{P}(D \mid A) = \frac{1}{100}, \quad \mathbb{P}(D \mid B) = \frac{1}{200}, \quad \mathbb{P}(D \mid C) = \frac{1}{20}$$

We need to compute $\mathbb{P}(C \mid D)$. Since the events A, B, C partition the sample space and have positive probabilities, we can use Corollary 11.1.33, which tells us that

$$\mathbb{P}(C \mid D) = \frac{\mathbb{P}(D \mid C)\mathbb{P}(C)}{\mathbb{P}(D \mid A)\mathbb{P}(A) + \mathbb{P}(D \mid B)\mathbb{P}(B) + \mathbb{P}(D \mid C)\mathbb{P}(C)}$$

Substituting the probabilities that we computed above, it follows that

$$\mathbb{P}(C \mid D) = \frac{\frac{1}{20} \cdot \frac{500}{10000}}{\frac{1}{100} \cdot \frac{3000}{10000} + \frac{1}{200} \cdot \frac{6500}{10000} + \frac{1}{20} \cdot \frac{500}{10000}} = \frac{2}{7} \approx 0.29$$

Exercise 11.1.35

In Example 11.1.34, find the probabilities that the car was an Allegheny and that the car was a Monongahela.

Section 11.2

Discrete random variables

Events in a probability space are sometimes unenlightening when looked at in isolation. For example, suppose we roll a fair six-sided die twice. The outcomes are elements of the set $\Omega = \{1,2,3,4,5,6\}^2$, each occurring with equal probability $\frac{1}{36}$. The event that the die rolls sum to 7 is precisely the subset

$$\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \subseteq \Omega$$

and so we can say that the probability that the two rolls sum to 7 is

$$\mathbb{P}(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \frac{1}{6}$$

However, it is not at all clear from the expression $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ that, when we wrote it down, what we had in mind was the event that the sum of the die rolls is 7. Moreover, the expression of the event in this way does not make it clear how to generalise to other possible sums of die rolls.

Note that the sum of the die rolls defines a function $S : \Omega \to \mathbb{N}$, defined by

$$S(a,b) = a+b \text{ for all } (a,b) \in \Omega$$

The function S allows us to express our event in a more enlightening way: indeed,

$$\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} = \{(a,b) \in \Omega \mid a+b = 7\} = S^{-1}[\{7\}]$$

(Recall the definition of *preimage* in Definition 3.1.37.) Thus the probability that the sum of the two die rolls is 7 is equal to $\mathbb{P}(S^{-1}[\{7\}])$.

If we think of S not as a function $\Omega \to \mathbb{N}$, but as a $\mathbb{N}$-valued *random variable*, which varies according to a random outcome in Ω, then we can informally say

$$\mathbb{P}\{S = 7\} = \frac{1}{6} \quad \text{which formally means} \quad \mathbb{P}(S^{-1}[\{7\}]) = \frac{1}{6}$$

This affords us much more generality. Indeed, we could ask what the probability is that the die rolls sum to a value greater than or equal to 7. In this case, note that the die rolls (a,b) sum to a number greater than or equal to 7 if and only if $a+b \in \{7,8,9,10,11,12\}$, which occurs if and only if $(a,b) \in S^{-1}[\{7,8,9,10,11,12\}]$. Thus, we might informally say

$$\mathbb{P}\{S \geqslant 7\} = \frac{7}{12} \quad \text{which formally means} \quad \mathbb{P}(S^{-1}[\{7,8,9,10,11,12\}]) = \frac{7}{12}$$

We might also ask what the probability is that the sum of the die rolls is prime. In this case, we might informally say

$$\mathbb{P}\{S \text{ is prime}\} = \frac{5}{12} \quad \text{which formally means} \quad \mathbb{P}(S^{-1}[\{2,3,5,7,11\}]) = \frac{5}{12}$$

and so on. In each of these cases, we're defining events—which are subsets of the sample space—in terms of conditions on the values of a random variable (which is, formally, a function).

We make the above intuition formal in Definition 11.2.1.

Definition 11.2.1

Let $(\Omega, \mathbb{P})$ be a probability space and let E be a set. An **E-valued random variable on** $(\Omega, \mathbb{P})$ is a function $X : \Omega \to E$ such that the image

$$X[\Omega] = \{X(\omega) \mid \omega \in \Omega\}$$

is countable. The set E is called the **state space** of X. A random variable with countable state space is called a **discrete random variable**.

Before we proceed with examples, some notation for events regarding values of random variables will be particularly useful.

Notation 11.2.2

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a set and let X be an E-valued random variable on $(\Omega, \mathbb{P})$. For each $e \in E$, write

$$\{X = e\} = \{\omega \in \Omega \mid X(\omega) = e\} = X^{-1}[\{e\}]$$

to denote the event that X takes the value e. More generally, for each logical formula $p(x)$ with free variable x ranging over E, we write

$$\{p(X)\} = \{\omega \in \Omega \mid p(X(\omega))\} = X^{-1}[\{e \in E \mid p(e)\}]$$

for the event that the value of X satisfies $p(x)$.

We will usually write $\mathbb{P}\{X = e\}$ instead of $\mathbb{P}(\{X = e\})$ for the probability that a random variable X takes a value e, and so on.

Example 11.2.3

We can model a sequence of three coin flips using the probability space $(\Omega, \mathbb{P})$, where $\Omega = \{\mathrm{H}, \mathrm{T}\}^3$ and $\mathbb{P}(\{\omega\}) = \frac{1}{8}$ for all $\omega \in \Omega$.

Let N be the real-valued random variable representing number of heads that show. This is formalised as a function

$$N : \Omega \to \mathbb{R} \quad \text{where} \quad N(i_1, i_2, i_3) = \text{the number of heads amongst } i_1, i_2, i_3$$

for example, $N(\mathrm{H}, \mathrm{T}, \mathrm{H}) = 2$. Now

- The probability that exactly two heads show is

$$\begin{aligned} \mathbb{P}\{N = 2\} &= \mathbb{P}(N^{-1}[\{2\}]) && \text{by Notation 11.2.2} \\ &= \mathbb{P}(\{(\mathrm{H}, \mathrm{H}, \mathrm{T}), (\mathrm{H}, \mathrm{T}, \mathrm{H}), (\mathrm{T}, \mathrm{H}, \mathrm{H})\}) && \text{evaluating event } N^{-1}[\{2\}] \\ &= \frac{3}{2^3} = \frac{3}{8} \end{aligned}$$

- The probability that at least two heads show is

$$\begin{aligned}\mathbb{P}\{N \geqslant 2\} &= \mathbb{P}(\{\omega \in \Omega \mid N(\omega) \geqslant 2\}) && \text{by Notation 11.2.2}\\ &= \mathbb{P}\left(\left\{\begin{matrix}(\mathrm{H},\mathrm{H},\mathrm{T}), & (\mathrm{H},\mathrm{T},\mathrm{H}),\\ (\mathrm{T},\mathrm{H},\mathrm{H}), & (\mathrm{H},\mathrm{H},\mathrm{H})\end{matrix}\right\}\right) && \text{evaluating event}\\ &= \frac{4}{2^3} = \frac{1}{2}\end{aligned}$$

Exercise 11.2.4

With probability space $(\Omega, \mathbb{P})$ and random variable N defined as in Example 11.2.3, compute $\mathbb{P}\{N \text{ is odd}\}$ and $\mathbb{P}\{N = 4\}$.

Each random variable comes with an associated *probability mass function*, which allows us to 'forget' the underlying probability space for the purposes of studying only the random variable.

Definition 11.2.5

Let $(\Omega, \mathbb{P})$ be a probability space, let $X : \Omega \to E$ be an E-valued random variable. The **probability mass function** of X is the function $f_X : E \to [0,1]$ defined by

$$f_X(e) = \mathbb{P}\{X = e\} \text{ for all } e \in E$$

Example 11.2.6

The probability mass function of the random variable N from Example 11.2.3 is the function $f_N : \mathbb{R} \to [0,1]$ defined by

$$f_N(e) = \mathbb{P}\{N = e\} = \frac{1}{8}\binom{3}{e}$$

for all $e \in \{0,1,2,3\}$, and $f_N(e) = 0$ otherwise. Indeed, there are $2^3 = 8$ possible outcomes, each equally likely, and $\binom{3}{e}$ of those outcomes show exactly e heads for $e \in \{0,1,2,3\}$.

Exercise 11.2.7

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a set, let X be an E-valued random variable and let $U \subseteq E$. Prove that the event $\{X \in U\}$ is equal to the preimage $X^{-1}[U]$. Deduce that

$$\mathbb{P}\{X \in U\} = \sum_{e \in U} f_X(e)$$

In Example 11.2.6, we could have just specified the value of f_N on $\{0,1,2,3\}$, with the understanding that N does not take values outside of this set and hence that $\mathbb{P}\{N = e\} = 0$ for all $e \notin \{0,1,2,3\}$. This issue arises frequently when dealing with real-valued discrete random variables, and it will be useful to ignore most (or all) of those real numbers which are not values of the random variable.

As such, for $E \subseteq \mathbb{R}$, we will from now on blur the distinction between the following concepts:

(i) E-valued random variables;

(ii) Real-valued random variables X such that $\mathbb{P}\{X = x\} = 0$ for all $x \notin E$.

Example 11.2.8

The probability mass function of the random variable N from Example 11.2.3 can be taken to be the function $f_X : \{0,1,2,3\} \to [0,1]$ defined by

$$f_X(k) = \frac{1}{8}\binom{3}{k} \text{ for all } k \in \{0,1,2,3\}$$

Lemma 11.2.9

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a set and let X be an E-valued random variable. The events $\{X = e\}$ for $e \in E$ are mutually exclusive, and their union is Ω.

Proof

If $e, e' \in E$, then for all $\omega \in \Omega$ we have

$$\begin{aligned} \omega \in \{X = e\} \cap \{X = e'\} &\Leftrightarrow \omega \in X^{-1}[\{e\}] \cap X^{-1}[\{e'\}] && \text{by Notation 11.2.2} \\ &\Leftrightarrow X(\omega) = e \text{ and } X(\omega) = e' && \text{by definition of preimage} \\ &\Rightarrow e = e' \end{aligned}$$

so if $e \neq e'$ then $\{X = e\} \cap \{X = e'\} = \varnothing$. So the events are mutually exclusive.

Moreover, if $\omega \in \Omega$, then $\omega \in \{X = X(\omega)\}$. Hence

$$\Omega = \bigcup_{e \in E} \{X = e\}$$

as required. □

Theorem 11.2.10

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a countable set, and let X be an E-valued random variable. Then

$$\sum_{e \in E} f_X(e) = 1$$

Proof

Since $f_X(e) = \mathbb{P}\{X = e\}$ for all $e \in E$, we need to check that

$$\sum_{e \in E} \mathbb{P}\{X = e\} = 1$$

By Lemma 11.2.9, we have

$$\sum_{e \in E} \mathbb{P}\{X = e\} = \mathbb{P}\left(\bigcup_{e \in E} \{X = e\}\right) = \mathbb{P}(\Omega) = 1$$

as required. □

The following corollary follows immediately.

Corollary 11.2.11

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a countable set, and let X be an E-valued random variable. The function $X_*\mathbb{P} : \mathscr{P}(E) \to [0,1]$ defined by

$$(X_*\mathbb{P})(A) = \sum_{e \in A} f_X(e) = \mathbb{P}\{X \in A\}$$

for all $A \subseteq E$ defines a probability measure on E. The space $(E, X_*\mathbb{P})$ is called the **pushforward probability measure** of X.

Corollary 11.2.11 implies that any statement about probability measures can be applied to the pushforward measure. For example,

$$\mathbb{P}\{X \in A \cup B\} = \mathbb{P}\{X \in A\} + \mathbb{P}\{X \in B\} - \mathbb{P}\{X \in A \cap B\}$$

for all subsets $A, B \subseteq E$.

As with events, there is a notion of independence for random variables.

Definition 11.2.12

Let $(\Omega, \mathbb{P})$ be a discrete probability space and let $X, Y : \Omega \to E$ be discrete random variables on $(\Omega, \mathbb{P})$. We say X and Y are **independent** if, for all $e, e' \in E$, the events $\{X = e\}$ and $\{Y = e'\}$ are independent. More generally, random variables $X_1, X_2, \ldots, X_n$ are **pairwise independent** if, for all $e_1, e_2, \ldots, e_n \in E$, the events $\{X_i = e_i\}$ are pairwise independent.

Example 11.2.13

A fair six-sided die is rolled twice. Let X be the value shown on the first roll and Y be the value shown on the second roll.

We can model this situation by letting $\Omega = \{1,2,3,4,5,6\}^2$ with $\mathbb{P}(\{(a,b)\}) = \frac{1}{36}$ for all $(a,b) \in \Omega$. The random variables X, Y can thus be taken to be functions $\Omega \to \{1,2,3,4,5,6\}$ defined by

$$X(a,b) = a \text{ and } Y(a,b) = b \text{ for all } (a,b) \in \Omega$$

So let $e, e' \in \{1,2,3,4,5,6\}$. Note first that

$$\begin{aligned} &\{X = e\} \cap \{Y = e'\} \\ &= \{(a,b) \in \Omega \mid a = e\} \cap \{(a,b) \in \Omega \mid b = e'\} && \text{by Notation 11.2.2} \\ &= \{(a,b) \in \Omega \mid a = e \text{ and } b = e'\} \\ &= \{(e,e')\} \end{aligned}$$

Hence

$$\mathbb{P}(\{X = e\} \cap \{Y = e'\}) = \mathbb{P}(\{(e,e')\}) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}\{X = e\}\mathbb{P}\{Y = e'\}$$

The events $\{X = e\}$ and $\{Y = e'\}$ are independent, and so X and Y are independent.

Exercise 11.2.14

A coin which shows heads with probability $p \in [0,1]$, and tails otherwise, is flipped five times. For each $1 \leqslant i \leqslant 5$, let

$$X_i = \begin{cases} 0 & \text{if the } i^{\text{th}} \text{ flip shows tails} \\ 1 & \text{if the } i^{\text{th}} \text{ flip shows heads} \end{cases}$$

Prove that the random variables X_1, X_2, X_3, X_4, X_5 are pairwise independent.

One final technicality that we mention before continuing concerns performing arithmetic with random variables which assume real values.

Notation 11.2.15

Let $(\Omega, \mathbb{P})$ be a probability space, and let X, Y be real-valued random variables on $(\Omega, \mathbb{P})$. Then we can define a new real-valued random variable $X+Y$ by

$$(X+Y)(\omega) = X(\omega) + Y(\omega) \text{ for all } \omega \in \Omega$$

Likewise for multipication, scalar multiplication and constants: for each $\omega \in \Omega$, define

$$(XY)(\omega) = X(\omega)Y(\omega), \quad (aX)(\omega) = a \cdot X(\omega), \quad a(\omega) = a$$

where $a \in \mathbb{R}$. Note that the random variables $X+Y, XY, aX, a$ are all supported on a countable set.

Example 11.2.16

A coin which shows heads with probability $p \in [0,1]$, and tails otherwise, is flipped n times. For each $1 \leqslant i \leqslant n$, let

$$X_i = \begin{cases} 0 & \text{if the } i^{\text{th}} \text{ flip shows tails} \\ 1 & \text{if the } i^{\text{th}} \text{ flip shows heads} \end{cases}$$

Then each X_i is a $[2]$-valued random variable.

Define $X = X_1 + X_2 + \cdots + X_n$. Then X is a $\mathbb{N}$-valued random variable representing the number of heads that show in total after the coin is flipped n times.

Probability distributions

Most of the random variables we are interested in are characterised by one of a few *probability distributions*. We won't define the term 'probability distribution' precisely—indeed, its use in the mathematical literature is often ambiguous and informal—instead, we will take it to mean any description of the random behaviour of a probability space or random variable.

The *uniform distribution* models the real-world situation in which any of a fixed number of outcomes occurs with equal probability.

Definition 11.2.17 · *Uniform distribution*

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a finite set, and let $X : \Omega \to E$ be a random variable. We say X follows the **uniform distribution on** E, or X is **uniformly distributed on** E, if f_X is

constant—that is, if

$$f_X(e) = \frac{1}{|E|} \text{ for all } e \in E$$

If X is uniformly distributed on E, we write $X \sim \text{Unif}(E)$ (LaTeX code: `\sim`).

Example 11.2.18

Let $(\Omega, \mathbb{P})$ be the probability space modelling the roll of a fair six-sided die, and let X be the $\{1,2,3,4,5,6\}$-valued random variable representing the number shown. Then for each $k \in \{1,2,3,4,5,6\}$ we have

$$f_X(k) = \mathbb{P}\{X = k\} = \mathbb{P}(\{k\}) = \frac{1}{6}$$

so X is uniformly distributed on $\{1,2,3,4,5,6\}$.

Exercise 11.2.19

Let $(\Omega, \mathbb{P})$ be the probability space modelling the roll of a fair six-sided die, and let X be the $\{0,1\}$-valued random variable which is equal to 0 if the die shows an even number and 1 if the die shows an odd number. Prove that $X \sim \text{Unif}(\{0,1\})$.

Before we continue, we prove that the notion of 'uniform distribution' does not make sense for countably infinite sets.

Theorem 11.2.20

Let $(\Omega, \mathbb{P})$ be a probability space and let E be a countably infinite set. There is no notion of a uniformly E-valued random variable X—that is, there is no $p \in [0,1]$ such that $f_X(e) = p$ for all $e \in E$.

Proof

We may assume $E = \mathbb{N}$; otherwise, re-index the sums accordingly.

Let $p \in [0,1]$. Note that

$$\sum_{n\in\mathbb{N}} f_X(n) = \sum_{n\in\mathbb{N}} p = \lim_{N\to\infty} \sum_{n=0}^{N} p = \lim_{N\to\infty} (N+1)p$$

If $p = 0$ then

$$\lim_{N\to\infty} (N+1)p = \lim_{N\to\infty} 0 = 0$$

If $p > 0$ then, for all $K > 0$, letting $N = \frac{K}{p}$ yields $(N+1)p = K + p > K$, and hence

$$\lim_{N\to\infty} (N+1)p = \infty$$

Thus $\sum_{n\in\mathbb{N}} p \neq 1$ for all $p \in [0,1]$.

In both cases, we have contradicted Theorem 11.2.10. As such, there can be no random variable $X : \Omega \to \mathbb{N}$ such that f_X is constant. □

The *Bernoulli distribution* models real-world situations in which one of two outcomes occurs, but not necessarily with the same probability.

Definition 11.2.21 · *Bernoulli distribution*

Let $(\Omega, \mathbb{P})$ be a probability space. A $\{0,1\}$-valued random variable X follows the **Bernoulli distribution with parameter** p if its probability mass function $f_X : \{0,1\} \to [0,1]$ satisfies

$$f_X(0) = 1 - p \quad \text{and} \quad f_X(1) = p$$

If X follows the Bernoulli distribution with parameter p, we write $X \sim \mathrm{B}(1,p)$.

The reason behind the notation $\mathrm{B}(1,p)$ will become clear soon—the Bernoulli distribution is a specific instance of a more general distribution, which we will see in Definition 11.2.24.

Example 11.2.22

A coin shows 'heads' with probability p and 'tails' with probability $1-p$. Let X be the random variable which takes the value 0 if the coin shows tails and 1 if the coin shows heads. Then $X \sim \mathrm{B}(1,p)$.

Exercise 11.2.23

Let X be a $\{0,1\}$-valued random variable. Prove that $X \sim \mathrm{U}(\{0,1\})$ if and only if $X \sim \mathrm{B}(1, \frac{1}{2})$.

Suppose that, instead of flipping a coin just once, as in Example 11.2.22, you flip it n times. The total number of heads that show must be an element of $\{0,1,\ldots,n\}$, and each such element occurs with some positive probability. The resulting probability distribution is called the *binomial distribution*.

Definition 11.2.24 · *Binomial distribution*

Let $(\Omega, \mathbb{P})$ be a probability space. A $\{0,1,\ldots,n\}$-valued random variable X follows the **binomial distribution with parameters** n, p if its probability mass function $f_X : \{0,1,\ldots,n\} \to [0,1]$ satisfies

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for all $k \in \{0,1,\ldots,n\}$. If X follows the binomial distribution with parameters n, p, we write $X \sim \mathrm{B}(n,p)$.

Example 11.2.25

A coin which shows heads with probability $p \in [0,1]$, and tails otherwise, is flipped n times. We will prove that the number of heads that show is binomially distributed.

We can model this situation with probability space $(\Omega, \mathbb{P})$ defined by taking $\Omega = \{\mathrm{H}, \mathrm{T}\}^n$, and letting $\mathbb{P}(\{\omega\}) = p^h (1-p)^t$ for all $\omega \in \Omega$, where h is the number of heads that show and t is the number of tails that show in outcome ω. For example, if $n = 5$ then

$$\mathbb{P}(\{\mathrm{HTHHT}\}) = p^3(1-p)^2 \quad \text{and} \quad \mathbb{P}(\{\mathrm{TTTTT}\}) = (1-p)^5$$

Note in particular that $h + t = n$.

Let X be the random variable which counts the number of heads that show. Formally, we can define $X : \{\mathrm{H},\mathrm{T}\}^n \to \{0,1,\ldots,n\}$ by letting $X(\omega)$ be the number of heads that show in outcome ω. For example if $n = 5$ then

$$X(\mathrm{HTHHT}) = 3 \quad \text{and} \quad X(\mathrm{TTTTT}) = 0$$

The event $\{X = k\}$ is the set of n-tuples of 'H's and 'T's which contain exactly k 'H'. Hence $|\{X = k\}| = \binom{n}{k}$, since such an n-tuple can be specified by choosing the k positions of the 'H's, and putting 'T's in the remaining positions. Since each outcome in this event occurs with equal probability $p^k(1-p)^{n-k}$, it follows that

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for all $k \in \{0,1,\ldots,n\}$. Hence $X \sim \mathrm{B}(n,p)$.

The following theorem proves that the sum of Bernoulli random variables follows the binomial distribution.

Theorem 11.2.26

Let $(\Omega, \mathbb{P})$ be a probability space, let $p \in [0,1]$ and let $X_1, X_2, \ldots, X_n : \Omega \to \{0,1\}$ be independent random variables such that $X_i \sim \mathrm{B}(1,p)$. Then

$$X_1 + X_2 + \cdots + X_n \sim \mathrm{B}(n,p)$$

Proof

Let $X = X_1 + X_2 + \cdots + X_n$. For each outcome ω and each $k \in \{0,1,\ldots,n\}$, we have $X(\omega) = k$ if and only if exactly k of the values $X_1(\omega), X_2(\omega), \ldots, X_n(\omega)$ are equal to 1.

For each specification S of *which* of the random variables X_i is equal to 1, let $A_S \subseteq \Omega$ be the event that this occurs. Formally, this is to say that, for each $S \subseteq [n]$, we define

$$A_S = \{\omega \in \Omega \mid X_i(\omega) = 0 \text{ for all } i \notin S \text{ and } X_i(\omega) = 1 \text{ for all } i \in S\}$$

Then $\mathbb{P}(A_S) = p^k(1-p)^{n-k}$, since the random variables $X_1, X_2, \ldots, X_n$ are mutually independent.

As argued above sets $\{A_S \mid U \subseteq [n],\ |S| = k\}$ form a partition of $\{X = k\}$, and hence

$$f_X(k) = \sum_{S \in \binom{[n]}{k}} \mathbb{P}(A_S) = \sum_{S \in \binom{[n]}{k}} p^k(1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

which is to say that $X \sim \mathrm{B}(n,p)$. □

We will make heavy use of Theorem 11.2.26 when we will study the *expectation* of binomially distributed random variables (Definition 11.2.34). First, let's will look at a couple of scenarios in which a binomially distributed random variable is expressed as a sum of independent Bernoulli random variables.

Example 11.2.27

In Example 11.2.25, we could have defined $\{0,1\}$-valued random variables $X_1, X_2, \ldots, X_n$ by letting

$$X_i(\omega) = \begin{cases} 0 & \text{if the } i^{\text{th}} \text{ coin flip shows tails} \\ 1 & \text{if the } i^{\text{th}} \text{ coin flip shows heads} \end{cases}$$

Then the number of heads shown in total is the random variable $X = X_1 + X_2 + \cdots + X_n$. Note that each random variable X_i follows the Bernoulli distribution with parameter p, and they are independent, so that $X \sim \mathrm{B}(n,p)$ by Theorem 11.2.26.

In Example 11.2.25, we flipped a coin a fixed number of times and counted how many heads showed. Now suppose that we flip a coin repeatedly until heads show, and then stop. The number of times the coin was flipped before heads shows could, theoretically, be any natural number. This situation is modelled by the *geometric distribution*.

Definition 11.2.28 · *Geometric distribution on* $\mathbb{N}$

Let $(\Omega, \mathbb{P})$ be a probability space. An $\mathbb{N}$-valued random variable X follows the **geometric distribution with parameter** p if its probability mass function $f_X : \mathbb{N} \to [0,1]$ satisfies

$$f_X(k) = (1-p)^k p \text{ for all } k \in \mathbb{N}$$

If X follows the geometric distribution with parameter p, we write $X \sim \mathrm{Geom}(p)$.

Example 11.2.29

A coin which shows heads with probability $p \in [0,1]$, and tails otherwise, is flipped repeatedly until heads shows.

Exercise 11.2.30

Let $p \in [0,1]$ and let $X \sim \mathrm{Geom}(p)$. Prove that

$$\mathbb{P}\{X \text{ is even}\} = \frac{1}{2-p}$$

What is the probability that X is odd?

Occasionally, it will be useful to consider geometrically distributed random variables which are valued in the set

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

of all *positive* natural numbers. The probability mass function of such a random variable is slightly different.

Definition 11.2.31 · *Geometric distribution on* $\mathbb{N}^+$

Let $(\Omega, \mathbb{P})$ be a probability space. An $\mathbb{N}^+$-valued random variable X follows the **geometric distribution with parameter** p if its probability mass function $f_X : \mathbb{N}^+ \to [0,1]$ satisfies

$$f_X(k) = (1-p)^{k-1} p \text{ for all } k \in \mathbb{N}^+$$

If X follows the geometric distribution with parameter p, we write $X \sim \mathrm{Geom}(p)$.

It is to be understood from context whether a given geometric random variable is $\mathbb{N}$-valued or $\mathbb{N}^+$-valued.

Example 11.2.32

An urn contains $n \geqslant 1$ distinct coupons. Each time you draw a coupon that you have not drawn before, you get a stamp. When you get all n stamps, you win. Let X be the number of coupons drawn up to, and including, a winning draw.

For each $k \in [n]$, let X_k be the random variable representing the number of draws required to draw the k^{th} new coupon, after $k-1$ coupons have been collected. Then the total number of times a coupon must be drawn is $X = X_1 + X_2 + \cdots + X_n$.

After $k-1$ coupons have been collected, there are $n-k+1$ uncollected coupons remaining in the urn, and hence on any given draw, an uncollected coupon is drawn with probability $\frac{n-k+1}{n}$, and a coupon that has already been collected is drawn with probability $\frac{k-1}{n}$. Hence for each $r \in \mathbb{N}^+$ we have

$$\mathbb{P}\{X_k = r\} = \left(\frac{k-1}{n}\right)^{r-1}\left(\frac{n-k+1}{n}\right)$$

That is to say, X_k is geometrically distributed on $\mathbb{N}^+$ with parameter $\frac{n-k+1}{n}$.

We will use this in Example 11.2.47 to compute the number of times a person should expect to have to draw coupons from the urn until they win.

Expectation

We motivate the definition of *expectation* (Definition 11.2.34) with the following example.

Example 11.2.33

For each $n \geqslant 1$, let X_n be the average value shown when a fair six-sided die is rolled n times.

When n is small, the value of X_n is somewhat unpredictable. For example, X_1 is uniformly distributed, since it takes each of the values $1, 2, 3, 4, 5, 6$ with equal probability. This is summarised in the following table:

e	1	2	3	4	5	6
$\mathbb{P}\{X_1 = e\}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

The distribution of X_2 is shown in the following table:

e	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
$\mathbb{P}\{X_2 = e\}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

As can be seen, the probabilities increase towards the middle of the table; the extreme values occur with low probability. This effect is exaggerated as n increases. Indeed,

$$\mathbb{P}\{X_n = 1\} = \mathbb{P}\{X_n = 6\} = \frac{1}{6^n}$$

so that $\mathbb{P}\{X_n = 1\} \to 0$ and $\mathbb{P}\{X_n = 6\} \to 0$ as $n \to \infty$; however, it can be shown that for all $\varepsilon > 0$, we have

$$\mathbb{P}\{3.5 - \varepsilon < X_n < 3.5 + \varepsilon\}$$

so that $\mathbb{P}\{|X_n - 3.5| < \varepsilon\} \to 1$ as $n \to \infty$.

Thus when we roll a die repeatedly, we can expect its value to eventually be as close to 3.5 as we like. This is an instance of a theorem called the *law of large numbers*

The value 3.5 in Example 11.2.33 is special because it is the average of the numbers $1,2,3,4,5,6$. More generally, assignments of different probabilities to different values of a random variable X yields a *weighted average* of the possible values. This weighted average, known as the *expectation* of the random variable, behaves in the same way as the number 3.5 did in Example 11.2.33.

Definition 11.2.34

Let $(\Omega, \mathbb{P})$ be a probability space, let $E \subseteq \mathbb{R}$ be countable, and let X be an E-valued random variable on $(\Omega, \mathbb{P})$. The **expectation** (or **expected value**) of X, if it exists, is the real number $\mathbb{E}[X]$ (LaTeX code: `\mathbb{E}`) defined by

$$\mathbb{E}[X] = \sum_{e \in E} e f_X(e)$$

Example 11.2.35

Let X be a random variable representing the value shown when a fair six-sided die is rolled. Then $X \sim \mathrm{U}(\{1,2,3,4,5,6\})$, so that $f_X(k) = \frac{1}{6}$ for all $1 \leqslant k \leqslant 6$, and hence

$$\mathbb{E}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5$$

so the expected value of the die roll is 3.5.

Example 11.2.36

Let $p \in [0,1]$ and let $X \sim \mathrm{B}(1,p)$. Then

$$\mathbb{E}[X] = 0 \cdot (1-p) + 1 \cdot p = p$$

So the expected value of a Bernoulli random variable is equal to the parameter.

Exercise 11.2.37

Let $(\Omega, \mathbb{P})$ be a probability space and let $c \in \mathbb{R}$. Thinking of c as a *constant* real-valued random variable,[b] prove that $\mathbb{E}[c] = c$.

The following lemma provides an alternative method for computing the expectation of a random variable. It will be useful for proving that expectation is *linear* in Theorem 11.2.43.

Lemma 11.2.38

Let $(\Omega, \mathbb{P})$ be a probability space, let E be a countable set and let X be an E-valued random variable on $(\Omega, \mathbb{P})$. Then

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\})$$

Proof

Recall from Lemma 11.2.9 that

$$\Omega = \bigcup_{e \in E} \{X = e\}$$

and the events $\{X = e\}$ are mutually exclusive. Hence

$$\begin{aligned}
\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) &= \sum_{e \in E} \sum_{\omega \in \{X=e\}} X(\omega)\mathbb{P}(\{\omega\}) && \text{by Lemma 11.2.9} \\
&= \sum_{e \in E} e\mathbb{P}\{X = e\} && \text{by (ii) in Proposition 11.1.5} \\
&= \sum_{e \in E} e f_X(e) && \text{by Definition 11.2.5}
\end{aligned}$$

as required. □

Proposition 11.2.39

Let $n \in \mathbb{N}$ and $p \in [0, 1]$, and suppose that X is a random variable such that $X \sim \mathrm{B}(n, p)$. Then $\mathbb{E}[X] = np$.

Proof

Since $X \sim \mathrm{B}(n, p)$, we have $f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ for all $0 \leqslant k \leqslant n$. Hence

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} && \text{by definition of expectation} \\
&= \sum_{k=1}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} && \text{since the } k = 0 \text{ term is zero} \\
&= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} && \text{by Proposition 8.1.39} \\
&= \sum_{\ell=0}^{n-1} n \binom{n-1}{\ell} p^{\ell+1} (1-p)^{(n-1)-\ell} && \text{writing } \ell = k+1 \\
&= np \cdot \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^{\ell} (1-p)^{(n-1)-\ell} && \text{pulling out constant factors} \\
&= np(p + (1-p))^{n-1} && \text{by the binomial theorem} \\
&= np && \text{since } p + (1-p) = 1
\end{aligned}$$

as required. □

Example 11.2.40

A coin which shows heads with probability $\frac{1}{3}$, and tails otherwise, is tossed 12 times. Letting X be the random variable represent the number of heads that show, we see that $X \sim \mathrm{B}(12, \frac{1}{3})$, and hence the expected number of heads that show is equal to

$$\mathbb{E}[X] = 12 \cdot \frac{1}{3} = 4$$

Exercise 11.2.41

Use Exercise 9.3.51 to prove that the expectation of a $\mathbb{N}$-valued random variable which is geometrically distributed with parameter $p \in [0,1]$ is equal to $\frac{1-p}{p}$. Use this to compute the expected number of times a coin must be flipped before the first time heads shows, given that heads shows with probability $\frac{2}{7}$.

Exercise 11.2.42

Prove that the expectation of a $\mathbb{N}^+$-valued random variable which is geometrically distributed with parameter $p \in [0,1]$ is equal to $\frac{1}{p}$.

Theorem 11.2.43 · *Linearity of expectation*

Let $(\Omega, \mathbb{P})$ be a probability space, let $E \subseteq \mathbb{R}$ be countable, let X and Y be E-valued random variables on $(\Omega, \mathbb{P})$, and let $a, b \in \mathbb{R}$. Then

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

Proof

This follows directly from the fact that summation is linear. Indeed,

$$\begin{aligned}
\mathbb{E}[aX+bY] &= \sum_{\omega\in\Omega} (aX+bY)(\omega)\mathbb{P}(\{\omega\}) && \text{by Lemma 11.2.38}\\
&= \sum_{\omega\in\Omega} \Big(aX(\omega)\mathbb{P}(\{\omega\}) + bY(\omega)\mathbb{P}(\{\omega\})\Big) && \text{expanding}\\
&= a\sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\}) + b\sum_{\omega\in\Omega} Y(\omega)\mathbb{P}(\{\omega\}) && \text{by linearity of summation}\\
&= a\mathbb{E}[X] + b\mathbb{E}[Y] && \text{by Lemma 11.2.38}
\end{aligned}$$

as required. □

Example 11.2.44

Let X be a random variable representing the sum of the numbers shown when a fair six-sided die is rolled twice. We can write $X = Y + Z$, where Y is the value of the first die roll and Z is the value of the second die roll. By Example 11.2.35, we have $\mathbb{E}[Y] = \mathbb{E}[Z] = 3.5$. Linearity of expectation then yields

$$\mathbb{E}[X] = \mathbb{E}[Y] + \mathbb{E}[Z] = 3.5 + 3.5 = 7$$

so the expected value of the sum of the two die rolls is 7.

Example 11.2.45

A coin, when flipped, shows heads with probability $p \in [0,1]$. The coin is flipped. If it shows heads, I gain \$10; if it shows tails, I lose \$20. We compute the least value of p that ensures that I do not expect to lose money.

Let X be the random variable which is equal to 0 if tails shows, and 1 if heads shows. then $X \sim \mathrm{B}(1,p)$, so that $\mathbb{E}[X] = p$ by Example 11.2.36. Let Y be the amount of money I gain. Then

$$Y = 10X - 20(1-X) = 30X - 20$$

Hence my expected winnings are

$$\mathbb{E}[Y] = 30\mathbb{E}[X] - 20 = 30p - 20$$

In order for this number to be non-negative, we require $p \geqslant \frac{2}{3}$.

Theorem 11.2.43 generalises by induction to linear combinations of countably many random variables; this is proved in the following exercise

Exercise 11.2.46

Let $(\Omega, \mathbb{P})$ be a probability space, let $E \subseteq \mathbb{R}$ be countable, let $\{X_i \mid i \in I\}$ be a family of E-valued random variables on $(\Omega, \mathbb{P})$, indexed by some countable set I, and let $\{a_n \mid n \in \mathbb{N}\}$ be an I-indexed family of real numbers. Prove that

$$\mathbb{E}\left[\sum_{i \in I} a_i X_i\right] = \sum_{i \in I} a_i \mathbb{E}[X_i]$$

Example 11.2.47

Recall Example 11.2.32: an urn contains $n \geqslant 1$ distinct coupons. Each time you draw a coupon that you have not drawn before, you get a stamp. When you get all n stamps, you win. We find the expected number of times you need to draw a coupon from the urn in order to win.

For each $1 \leqslant k \leqslant n$, let X_k be the random variable representing the number of draws required to draw the k^{th} new coupon, after $k-1$ coupons have been collected. Then the total number of times a coupon must be drawn is $X = X_1 + X_2 + \cdots + X_n$.

We already saw that $X_k \sim \mathrm{Geom}\left(\frac{n-k+1}{n}\right)$ for each $1 \leqslant k \leqslant n$. By Exercise 11.2.42, we have $\mathbb{E}[X_k] = \frac{n}{n-k+1}$ for all $1 \leqslant k \leqslant n$. By linearity of expectation, it follows that

$$\mathbb{E}[X] = \sum_{k=1}^{n} \mathbb{E}[X_k] = \sum_{k=1}^{n} \frac{n}{n-k+1} = n \sum_{i=1}^{n} \frac{1}{i}$$

Section 11.E

Chapter 11 exercises

True–False questions

In Questions 11.E.1 to 11.E.5, determine (with proof) whether the statement is true or false.

11.E.1. In a probability space $(\Omega, \mathbb{P})$, we have $\mathbb{P}(A \mid \Omega) = \mathbb{P}(A)$ for all events A.

11.E.2. The function $\mathbb{P} : \mathscr{P}(\mathbb{N}) \to [0,1]$ defined by $\mathbb{P}(A) = \sum_{n \in A} 2^{-n}$ is a probability measure on $\mathbb{N}$.

11.E.3. The function $\mathbb{P} : \mathscr{P}(\mathbb{Z}) \to [0,1]$ defined by $\mathbb{P}(A) = \sum_{n \in A} 2^{-n}$ is a probability measure on $\mathbb{Z}$.

11.E.4. The function $\mathbb{P} : \mathscr{P}(\mathbb{N}) \to [0,1]$, defined by letting $\mathbb{P}(\varnothing) = \varnothing$ and $\mathbb{P}(A) = 2^{-\min(A)}$ for all $A \neq \varnothing$, is a probability measure on $\mathbb{N}$.

11.E.5. A random variable can have infinite expectation.

Always–Sometimes–Never questions

In Questions 11.E.6 to 11.E.10, determine (with proof) whether the conclusion is always, sometimes or never true under the given hypotheses.

11.E.6. Let $(\Omega, \mathbb{P})$ be a probability space. The only event whose probability is zero is the empty event $\varnothing \subseteq \Omega$.

11.E.7. Let $(\Omega, \mathbb{P})$ be a probability space and let A be an event. Then $\mathbb{P}(A)^2 - \mathbb{P}(A) = 2$.

11.E.8. Let $(\Omega, \mathbb{P})$ be a probability space and let A and B be events with $\mathbb{P}(B) > 0$. Then $\mathbb{P}(A \mid B) \geqslant \mathbb{P}(A)$.

11.E.9. Let $(\Omega, \mathbb{P})$ be a probability space and let A and B be mutually exclusive events with $\mathbb{P}(B) > 0$. Then $\mathbb{P}(A \mid B) = 0$.

11.E.10. Let $(\Omega, \mathbb{P})$ be a probability space and let A, B and C be events with $\mathbb{P}(B) > 0$ and $\mathbb{P}(C) > 0$. Then $\mathbb{P}(A \mid B)\mathbb{P}(B \mid C) = \mathbb{P}(A \mid C)$.

Chapter 12

Additional topics

Section 12.1

Orders and lattices

We saw in Section 5.1 how equivalence relations behave like '$=$', in the sense that they are reflexive, symmetric and transitive.

This section explores a new kind of relation which behaves like '$\leqslant$'. This kind of relation proves to be extremely useful for making sense of mathematical structures, and has powerful applications throughout mathematics, computer science and even linguistics.

Definition 12.1.1

A relation R on a set X is a **partial order** if R is reflexive, antisymmetric and transitive. That is, if:

- (Reflexivity) $x \mathrel{R} x$ for all $x \in X$;
- (Antisymmetry) For all $x, y \in X$, if $x \mathrel{R} y$ and $y \mathrel{R} x$, then $x = y$;
- (Transitivity) For all $x, y, z \in X$, if $x \mathrel{R} y$ and $y \mathrel{R} z$, then $x \mathrel{R} z$.

A set X together with a partial order R on X is called a **partially ordered set**, or **poset** for short, and is denoted (X, R).

When we talk about partial orders, we usually use a suggestive symbol like '$\preccurlyeq$' (LaTeX code: `\preceq`) or '$\sqsubseteq$' (LaTeX code: `\sqsubseteq`).

Example 12.1.2

We have seen many examples of posets so far:

- Any of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$, with the usual order relation $\leqslant$.
- Given a set X, its power set $\mathscr{P}(X)$ is partially ordered by $\subseteq$. Indeed:
  - **Reflexivity.** If $U \in \mathscr{P}(X)$ then $U \subseteq U$.
  - **Antisymmetry.** If $U, V \in \mathscr{P}(X)$ with $U \subseteq V$ and $V \subseteq U$, then $U = V$ by definition of set equality.
  - **Transitivity.** If $U, V, W \in \mathscr{P}(X)$ with $U \subseteq V$ and $V \subseteq W$, then $U \subseteq W$ by Proposition 2.1.12.
- The set $\mathbb{N}$ of natural numbers is partially ordered by the divisibility relation—see Exercises 5.1.23, 5.1.33 and 5.1.40. However, as noted in Exercise 5.1.33, the set $\mathbb{Z}$ of integers is not partially ordered by divisibility, since divisibility is not antisymmetric on $\mathbb{Z}$.
- Any set X is partially ordered by its equality relation. This is called the **discrete order** on X.

Much like the difference between the relations $\leqslant$ and $<$ on $\mathbb{N}$, or between $\subseteq$ and $\subsetneq$ on $\mathscr{P}(X)$, every partial order can be *strictified*, in a precise sense outlined in the following definition and proposition.

Definition 12.1.3

A relation R on a set X is a **strict partial order** if it is irreflexive, asymmetric and transitive. That is, if:

- (Irreflexivity) $\neg(x\,R\,x)$ for all $x \in X$;
- (Asymmetry) For all $x, y \in X$, if $x\,R\,y$, then $\neg(y\,R\,x)$;
- (Transitivity) For all $x, y, z \in X$, if $x\,R\,y$ and $y\,R\,z$, then $x\,R\,z$.

Proposition 12.1.4

Let X be a set. Partial orders $\preccurlyeq$ on X are in natural correspondence with strict partial orders $\prec$ on X, according to the rule:

$$x \preccurlyeq y \;\Leftrightarrow\; (x \prec y \vee x = y) \quad \text{and} \quad x \prec y \;\Leftrightarrow\; (x \preccurlyeq y \wedge x \neq y)$$

Proof

Let P be the set of all partial orders on X and let S be the set of all strict partial orders on X. Define functions

$$f : P \to S \quad \text{and} \quad g : S \to P$$

as in the statement of the proposition, namely:

- Given a partial order $\preccurlyeq$, let $f(\preccurlyeq)$ be the relation $\prec$ defined for $x, y \in X$ by letting $x \prec y$ be true if and only if $x \preccurlyeq y$ and $x \neq y$;
- Given a strict partial order $\prec$, let $g(\prec)$ be the relation $\preccurlyeq$ defined for $x, y \in X$ by letting $x \preccurlyeq y$ be true if and only if $x \prec y$ or $x = y$.

We'll prove that f and g are mutually inverse functions. Indeed:

- f is well-defined. To see this, fix $\preccurlyeq$ and $\prec = f(\preccurlyeq)$ and note that:
  - $\prec$ is irreflexive, since for $x \in X$ if $x \prec x$ then $x \neq x$, which is a contradiction.
  - $\prec$ is asymmetric. To see this, let $x, y \in X$ and suppose $x \prec y$. Then $x \preccurlyeq y$ and $x \neq y$. If also $y \prec x$, then we'd have $y \preccurlyeq x$, so that $x = y$ by antisymmetry of $\preccurlyeq$. But $x \neq y$, so this is a contradiction.
  - $\prec$ is transitive. To see this, let $x, y, z \in X$ and suppose $x \prec y$ and $y \prec z$. Then $x \preccurlyeq y$ and $y \preccurlyeq z$, so that $x \preccurlyeq z$. Moreover, if $x = z$ then we'd also have $z \preccurlyeq x$ by reflexivity of $\preccurlyeq$, so $z \preccurlyeq y$ by transitivity of $\preccurlyeq$, and hence $y = z$ by antisymmetry of $\preccurlyeq$. But this contradicts $y \prec z$.

  So $\prec$ is a strict partial order on X.
- g is well-defined. To see this, fix $\prec$ and $\preccurlyeq = g(\prec)$ and note that:
  - $\preccurlyeq$ is reflexive. This is built into the definition of $\preccurlyeq$.
  - $\preccurlyeq$ is symmetric. To see this, fix $x, y \in X$ and suppose $x \preccurlyeq y$ and $y \preccurlyeq x$. Now if $x \neq y$ then $x \prec y$ and $y \prec x$, but this contradicts asymmetry of $\prec$. Hence $x = y$.
  - $\preccurlyeq$ is transitive. To see this, fix $x, y, z \in X$ and suppose $x \preccurlyeq y$ and $y \preccurlyeq z$. Then one of the following four cases must be true:

* $x = y = z$. In this case, $x = z$, so $x \preccurlyeq z$.
* $x = y \prec z$. In this case, $x \prec z$, so $x \preccurlyeq z$.
* $x \prec y = z$. In this case, $x \prec z$, so $x \preccurlyeq z$.
* $x \prec y \prec z$. In this case, $x \prec z$ by transitivity of $\prec$, so $x \preccurlyeq z$.

In any case, we have that $x \preccurlyeq z$.

So $\preccurlyeq$ is a partial order on X.

- $g \circ f = \mathrm{id}_P$. To see this, let $\prec = f(\preccurlyeq)$ and $\sqsubseteq = g(\prec)$. For $x, y \in X$, we have $x \sqsubseteq y$ if and only if $x \prec y$ or $x = y$, which in turn occurs if and only if $x = y$ or both $x \preccurlyeq y$ and $x \neq y$. This is equivalent to $x \preccurlyeq y$, since if $x = y$ then $x \preccurlyeq y$ by reflexivity. Hence $\sqsubseteq$ and $\preccurlyeq$ are equal relations, so $g \circ f = \mathrm{id}_P$.
- $f \circ g = \mathrm{id}_S$. To see this, let $\preccurlyeq = g(\prec)$ and $\sqsubset = f(\preccurlyeq)$. For $x, y \in X$, we have $x \sqsubset y$ if and only if $x \preccurlyeq y$ and $x \neq y$, which in turn occurs if and only if $x \neq y$ and either $x \prec y$ or $x = y$. Since $x \neq y$ precludes $x = y$, this is equivalent to $x \prec y$. Hence $\prec$ and $\sqsubset$ are equal relations, so $f \circ g = \mathrm{id}_S$.

So f and g are mutually inverse functions, and we have established the required bijection. □

In light of Proposition 12.1.4, we will freely translate between partial orders and strict partial orders wherever necessary. When we do so, we will use $\prec$ (LaTeX code: `\prec`) to denote the 'strict' version, and $\preccurlyeq$ to denote the 'weak' version. (Likewise for $\sqsubset$ (LaTeX code: `\sqsubset`).)

Definition 12.1.5

Let $(X, \preccurlyeq)$ be a poset. A $\preccurlyeq$-**least element** of X (or a **least element of** X **with respect to** $\preccurlyeq$) is an element $\bot \in X$ (LaTeX code: `\bot`) such that $\bot \preccurlyeq x$ for all $x \in X$. A $\preccurlyeq$-**greatest element** of X (or a **greatest element of** X **with respect to** $\preccurlyeq$) is an element $\top \in X$ (LaTeX code: `\top`) such that $x \preccurlyeq \top$ for all $x \in X$.

Example 12.1.6

Some examples of least and greatest elements that we have already seen are:

- In $(\mathbb{N}, \leqslant)$, 0 is a least element; there is no greatest element.
- Let $n \in \mathbb{N}$ with $n > 0$. Then 1 is a least element of $([n], \leqslant)$, and n is a greatest element.
- $(\mathbb{Z}, \leqslant)$ has no greatest or least elements.

Proposition 12.1.7 says that least and greatest elements of posets are unique, if they exist. This allows us to talk about 'the' least or 'the' greatest element of a poset.

Proposition 12.1.7

Let $(X, \preccurlyeq)$ be a poset. If X has a least element, then it is unique; and if X has a greatest element, then it is unique.

Proof

Suppose X has a least element ℓ. We prove that if ℓ' is another least element, then $\ell' = \ell$.

So take another least element ℓ'. Since ℓ is a least element, we have $\ell \preccurlyeq \ell'$. Since ℓ' is a least element, we have $\ell' \preccurlyeq \ell$. By antisymmetry of $\preccurlyeq$, it follows that $\ell = \ell'$.

Hence least elements are unique. The proof for greatest elements is similar, and is left as an exercise. □

Exercise 12.1.8
Let X be a set. The poset $(\mathscr{P}(X), \subseteq)$ has a least element and a greatest element; find both.

Exercise 12.1.9
Prove that the least element of $\mathbb{N}$ with respect to divisibility is 1, and the greatest element is 0.

Definition 12.1.10 · *Supremum*
Let $(X, \preccurlyeq)$ be a poset and let $A \subseteq X$. A $\preccurlyeq$-**supremum** of A is an element $s \in X$ such that

- $a \preccurlyeq s$ for each $a \in A$; and
- If $s' \in X$ with $a \preccurlyeq s'$ for all $a \in A$, then $s \preccurlyeq s'$.

A $\preccurlyeq$-**infimum** of A is an element $i \in X$ such that

- $i \preccurlyeq a$ for each $a \in A$; and
- If $i' \in X$ with $i' \preccurlyeq a$ for all $a \in A$, then $i' \preccurlyeq i$.

Example 12.1.11
The well-ordering principle states that if $U \subseteq \mathbb{N}$ is inhabited then U has a $\leqslant$-infimum, and moreover the infinum of U is an element of U.

Exercise 12.1.12
Let X be a set, and let $U, V \in \mathscr{P}(X)$. Prove that the $\subseteq$-supremum of $\{U, V\}$ is $U \cup V$, and the $\subseteq$-infimum of $\{U, V\}$ is $U \cap V$.

Exercise 12.1.13
Let $a, b \in \mathbb{N}$. Show that $\gcd(a, b)$ is an infimum of $\{a, b\}$ and that $\operatorname{lcm}(a, b)$ is a supremum of $\{a, b\}$ with respect to divisibility.

Example 12.1.14
Define $U = [0, 1) = \{x \in \mathbb{R} \mid 0 \leqslant x < 1\}$. We prove that U has both an infimum and a supremum in the poset $(\mathbb{R}, \leqslant)$.

- **Infimum.** 0 is an infimum for U. Indeed:
  (i) Let $x \in U$. Then $0 \leqslant x$ by definition of U.
  (ii) Let $y \in \mathbb{R}$ and suppose that $y \leqslant x$ for all $x \in U$. Then $y \leqslant 0$, since $0 \in U$.

so 0 is as required.

- **Supremum.** 1 is a supremum for U. Indeed:
  (i) Let $x \in U$. Then $x < 1$ by definition of U, so certainly $x \leqslant 1$.
  (ii) Let $y \in \mathbb{R}$ and suppose that $x \leqslant y$ for all $x \in U$. We prove that $1 \leqslant y$ by contradiction. So suppose it is not the case that $1 \leqslant y$. Then $y < 1$. Since $x \leqslant y$ for all $x \in U$, we have $0 \leqslant y$. But then
  $$0 \leqslant y = \frac{y+y}{2} < \frac{y+1}{2} < \frac{1+1}{2} = 1$$
  But then $\frac{y+1}{2} \in U$ and $y < \frac{y+1}{2}$. This contradicts the assumption that $x \leqslant y$ for all $x \in U$. So it must in fact have been the case that $1 \leqslant y$.

so 1 is as required.

The following proposition proves that suprema and infima are unique, provided they exist.

Proposition 12.1.15
Let $(X, \preccurlyeq)$ is a poset, and let $A \subseteq X$.

(i) If $s, s' \in X$ are suprema of A, then $s = s'$;

(ii) If $i, i' \in X$ are infima of A, then $i = i'$.

Proof
Suppose s, s' are suprema of A. Then:

- $a \preccurlyeq s'$ for all $a \in A$, so $s' \preccurlyeq s$ since s is a supremum of A;
- $a \preccurlyeq s$ for all $a \in A$, so $s \preccurlyeq s'$ since s' is a supremum of A.

Since $\preccurlyeq$ is antisymmetric, it follows that $s = s'$. This proves (i).

The proof of (ii) is almost identical and is left as an exercise to the reader. □

Notation 12.1.16
Let $(X, \preccurlyeq)$ be a poset and let $U \subseteq X$. Denote the $\preccurlyeq$-infimum of U, if it exists, by $\bigwedge U$ (LaTeX code: `\bigwedge`); and denote the $\preccurlyeq$-supremum of U, if it exists, by $\bigvee U$ (LaTeX code: `\bigvee`). Moreover, for $x, y \in X$, write

$$\bigwedge\{x,y\} = x \wedge y \text{ (LaTeX code: \wedge)}, \qquad \bigvee\{x,y\} = x \vee y \text{ (LaTeX code: \vee)}$$

Example 12.1.17
Some examples of Notation 12.1.16 are as follows.

- Let X be a set. In $(\mathscr{P}(X), \subseteq)$ we have $U \wedge V = U \cap V$ and $U \vee V = U \cup V$ for all $U, V \in \mathscr{P}(X)$.
- We have seen that, in $(\mathbb{N}, |)$, we have $a \wedge b = \gcd(a,b)$ and $a \vee b = \operatorname{lcm}(a,b)$ for all $a, b \in \mathbb{N}$.

- In $(\mathbb{R}, \leqslant)$, we have $a \wedge b = \min\{a,b\}$ and $a \vee b = \max\{a,b\}$.

Definition 12.1.18
A **lattice** is a poset $(X, \preccurlyeq)$ such that every pair of elements of X has a $\preccurlyeq$-supremum and a $\preccurlyeq$-infimum.

Example 12.1.19
We have seen that $(\mathscr{P}(X), \subseteq)$, $(\mathbb{R}, \leqslant)$ and $(\mathbb{N}, |)$ are lattices.

Proposition 12.1.20 · *Associativity laws for lattices*
Let $(X, \preccurlyeq)$ be a lattice, and let $x, y, z \in X$. Then

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z \qquad \text{and} \qquad x \vee (y \vee z) = (x \vee y) \vee z$$

Proof
We prove $x \wedge (y \wedge z) = (x \wedge y) \wedge z$; the other equation is dual and is left as an exercise. We prove that the sets $\{x, y \wedge z\}$ and $\{x \wedge y, z\}$ have the same sets of lower bounds, and hence the same infima. So let

$$L_1 = \{i \in X \mid i \preccurlyeq x \text{ and } i \preccurlyeq y \wedge z\} \quad \text{and} \quad L_2 = \{i \in X \mid i \preccurlyeq x \wedge y \text{ and } i \preccurlyeq z\}$$

We prove $L_1 = L = L_2$, where

$$L = \{i \in X \mid i \preccurlyeq x,\ i \preccurlyeq y \text{ and } i \preccurlyeq z\}$$

First we prove $L_1 = L$. Indeed:

- $L_1 \subseteq L$. To see this, suppose $i \in L_1$. Then $i \preccurlyeq x$ by definition of L_1. Since $i \preccurlyeq y \wedge z$, and $y \wedge z \preccurlyeq y$ and $y \wedge z \preccurlyeq z$, we have $i \preccurlyeq y$ and $i \preccurlyeq z$ by transitivity of $\preccurlyeq$.
- $L \subseteq L_1$. To see this, suppose $i \in L$. Then $i \preccurlyeq x$ by definition of L. Moreover, $i \preccurlyeq y$ and $i \preccurlyeq z$ by definition of L, so that $i \preccurlyeq y \wedge z$ by definition of $\wedge$. Hence $i \in L$.

The proof that $L_2 = L$ is similar. Hence $L_1 = L_2$. But $x \wedge (y \wedge z)$ is, by definition of $\wedge$, the $\preccurlyeq$-greatest element of L_1, which exists since $(X, \preccurlyeq)$ is a lattice. Likewise, $(x \wedge y) \wedge z$ is the $\preccurlyeq$-greatest element of L_2.

Since $L_1 = L_2$, it follows that $x \wedge (y \wedge z) = (x \wedge y) \wedge z$, as required. □

In the next exercise, you will prove some properties satisfied by suprema and infima in addition to associativity.

Exercise 12.1.21 · *Properties of suprema and infima*
Let $(X, \preccurlyeq)$ be a lattice. Prove that, for all $x, y \in X$, we have:

(a) (**Idempotence**) $x \wedge x = x$ and $x \vee x = x$;

(b) (**Commutativity**) $x \wedge y = y \wedge x$ and $x \vee y = y \vee x$;

(c) (**Absorption**) $x \vee (x \wedge y) = x$ and $x \wedge (x \vee y) = x$.

Example 12.1.22
It follows from what we've proved that if $a, b, c \in \mathbb{Z}$ then

$$\gcd(a, \gcd(b,c)) = \gcd(\gcd(a,b), c)$$

For example, take $a = 882$, $b = 588$ and $c = 252$. Then

- $\gcd(b,c) = 84$, so $\gcd(a, \gcd(b,c)) = \gcd(882, 84) = 42$;
- $\gcd(a,b) = 294$, so $\gcd(\gcd(a,b), c) = \gcd(294, 252) = 42$.

These are indeed equal.

Distributive lattices and Boolean algebras

One particularly important class of lattice is that of a *distributive lattice*, in which suprema and infima interact in a particularly convenient way. This makes algebraic manipulations of expressions involving suprema and infima particularly simple.

Definition 12.1.23
A lattice $(X, \preccurlyeq)$ is **distributive** if

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \quad \text{and} \quad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$

for all $x, y, z \in X$.

Example 12.1.24
For any set X, the power set lattice $(\mathscr{P}(X), \subseteq)$ is distributive. That is to say that for all $U, V, W \subseteq X$ we have

$$U \cap (V \cup W) = (U \cap V) \cup (U \cap W) \quad \text{and} \quad U \cup (V \cap W) = (U \cup V) \cap (U \cup W)$$

This was the content of Example 2.2.27 and Exercise 2.2.28.

Exercise 12.1.25
Prove that $(\mathbb{N}, |)$ is a distributive lattice.

Definition 12.1.26
Let $(X, \preccurlyeq)$ be a lattice with a greatest element $\top$ and a least element $\bot$, and let $x \in X$. A **complement** for x is an element y such that

$$x \wedge y = \bot \quad \text{and} \quad x \vee y = \top$$

Example 12.1.27
Let X be a set. We show that every element $U \in \mathscr{P}(X)$ has a complement.

Exercise 12.1.28
Let $(X, \preccurlyeq)$ be a distributive lattice with a greatest element and a least element, and let $x \in X$. Prove that, if a complement for x exists, then it is unique; that is, prove that if $y, z \in X$ are complements for X, then $y = z$.

Exercise 12.1.28 justifies the following notation.

Notation 12.1.29
Let $(X, \preccurlyeq)$ be a distributive lattice with greatest and least elements. If $x \in X$ has a complement, denote it by $\neg x$.

Proposition 12.1.30
Let $(X, \preccurlyeq)$ be a distributive lattice and let $x \in X$. If x has a complement $\neg x$, then $\neg x$ has a complement, and $\neg(\neg x) = x$.

Proof
Assume that x has a complement $\neg x$. Then by commutativity of $\wedge$ and $\vee$ and by definition of $\neg$, we have

$$(\neg x) \wedge x \;=\; x \wedge (\neg x) \;=\; \bot \quad \text{and} \quad (\neg x) \vee x \;=\; x \vee (\neg x) \;=\; \top$$

so x satisfies the definition of what it means to be a complement of $\neg x$. By uniqueness of complements (Exercise 12.1.28), we have $x = \neg(\neg x)$. □

Definition 12.1.31
A lattice $(X, \preccurlyeq)$ is **complemented** if every element $x \in X$ has a complement. A **Boolean algebra** is a complemented distributive lattice with a greatest element and a least element.

The many preceding examples and exercises concerning $(\mathscr{P}(X), \subseteq)$ piece together to provide a proof of the following theorem.

Theorem 12.1.32
Let X be a set. Then $(\mathscr{P}(X), \subseteq)$ is a Boolean algebra.

Another extremely important example of a Boolean algebra is known as the *Lindenbaum–Tarski algebra*, which we define in Definition 12.1.35. In order to define it, we need to prove that the definition will make sense. First of all, we fix some notation.

Definition 12.1.33
Let P be a set, thought of as a set of propositional variables. Write $\mathscr{L}(P)$ to denote the set of propositional formulae with propositional variables in P—that is, the elements of $\mathscr{L}(P)$ are strings built from the elements of P, using the operations of conjunction ($\wedge$), disjunction ($\vee$) and negation ($\neg$).

Lemma 12.1.34
Logical equivalence $\equiv$ is an equivalence relation on $\mathscr{L}(P)$.

Proof
By Theorem 1.3.6 we know that, for all $s,t \in \mathscr{L}(P)$, $s \equiv t$ if and only if s and t have the same truth values for all assignments of truth values to their constituent propositional variables. Reflexivity, symmetry and transitivity of $\equiv$ are then immediate. □

In what follows, the set P of propositional variables is fixed; we may moreover take it to be countably infinite, since all strings in $\mathscr{L}(P)$ are finite.

Definition 12.1.35
The **Lindenbaum–Tarski algebra (for propositional logic)** over P is the pair $(A,\vdash)$, where $A = \mathscr{L}(P)/\equiv$ and $\vdash$ is the relation on A defined by $[s]_\equiv \vdash [t]_\equiv$ if and only if $s \Rightarrow t$ is a tautology.

In what follows, we will simply write $[-]$ for $[-]_\equiv$.

Theorem 12.1.36
The Lindenbaum–Tarski algebra is a Boolean algebra.

Proof Sketch proof
There is lots to prove here! Indeed, we must prove:

- $\vdash$ is a well-defined relation on A; that is, if $s \equiv s'$ and $t \equiv t'$ then we must have $[s] \vdash [t]$ if and only if $[s'] \vdash [t']$.
- $\vdash$ is a partial order on A; that is, it is reflexive, antisymmetric and transitive.
- The poset $(A,\vdash)$ is a lattice; that is, it has suprema and infima.
- The lattice $(A,\vdash)$ is distributive, has a greatest element and a least element, and is complemented.

We will omit most of the details, which are left as an exercise; instead, we outline what the components involved are.

The fact that $\vdash$ is a partial order can be proved as follows.

- Reflexivity of $\vdash$ follows from the fact that $s \Rightarrow s$ is a tautology for all propositional formulae s.
- Symmetry of $\vdash$ follows from the fact that, for all propositional formulae s,t, if $s \Leftrightarrow t$ is a tautology then s and t are logically equivalent.
- Transitivity of $\vdash$ follows immediately from transitivity of $\Rightarrow$.

The fact that $(A,\vdash)$ is a lattice can be proved by verifying that:

- Given $[s],[t] \in A$, the infimum $[s] \wedge [t]$ is given by conjunction, namely $[s] \wedge [t] = [s \wedge t]$.

- Given $[s], [t] \in A$, the supremum $[s] \vee [t]$ is given by disjunction, namely $[s] \vee [t] = [s \vee t]$.

Finally, distributivity of suprema and infima in $(A, \vdash)$ follows from the corresponding properties of conjunction and disjunction; $(A, \vdash)$ has greatest element $[p \Rightarrow p]$ and least element $[\neg(p \Rightarrow p)]$, where p is some fixed propositional variable; and the complement of $[s] \in A$ is given by $[\neg s]$. □

De Morgan's laws revisited

We finish this section on orders and lattices with a general version of de Morgan's laws for Boolean algebras, which by Theorems Theorems 12.1.32 and 12.1.36 implies the versions we proved for logical formulae (Theorem 1.3.24) and for sets (Theorem 2.2.44(a)–(b)).

Theorem 12.1.37 · *De Morgan's laws*
Let $(X, \preccurlyeq)$ be a Boolean algebra, and let $x, y \in X$. Then

$$\neg(x \wedge y) = (\neg x) \vee (\neg y) \quad \text{and} \quad \neg(x \vee y) = (\neg x) \wedge (\neg y)$$

Proof of (a)
We prove that $(\neg x) \vee (\neg y)$ satisfies the definition of a complement for $x \wedge y$; then we'll have $(\neg x) \vee (\neg y) = \neg(x \wedge y)$ by Exercise 12.1.28.

So let $z = (\neg x) \vee (\neg y)$. Then

$(x \wedge y) \wedge z$	
$= (x \wedge y) \wedge ((\neg x) \vee (\neg y))$	by definition of z
$= [(x \wedge y) \wedge (\neg x)] \vee [(x \wedge y) \wedge (\neg y)]$	by distributivity
$= [(x \wedge (\neg x)) \wedge y] \vee [x \wedge (y \wedge (\neg y))]$	by associativity and commutativity
$= [\bot \wedge y] \vee [x \wedge \bot]$	by definition of complements
$= \bot \vee \bot$	by definition of $\bot$ and $\wedge$
$= \bot$	by idempotence

Likewise we have

$(x \wedge y) \vee z$	
$= (x \wedge y) \vee ((\neg x) \vee (\neg y))$	by definition of z
$= [x \vee ((\neg x) \vee (\neg y))] \wedge [y \vee ((\neg x) \vee (\neg y))]$	by distributivity
$= [((x \vee (\neg x)) \vee (\neg y)] \wedge [(\neg x) \vee (y \vee (\neg y))]$	by associativity and commutativity
$= [\top \vee (\neg y)] \wedge [(\neg x) \vee \top]$	by definition of complements
$= \top \wedge \top$	by definition of $\top$ and $\wedge$
$= \top$	by idempotence

Since $(x \wedge y) \wedge z = \bot$ and $(x \wedge y) \vee z = \top$, we have

$$(\neg x) \vee (\neg y) \;=\; z \;=\; \neg(x \wedge y)$$

by definition of complements. □

✎ **Exercise 12.1.38**
Prove part (b) of Theorem 12.1.37.

Section 12.2

Inductively defined sets

In Section 4.1, we formalised the idea that the set of natural numbers should be what is obtained by starting with zero and repeating the successor ('plus one') operation—this was done using Peano's axioms (Definition 4.1.1). From these axioms we were able to derive the weak and strong induction principles, which turned out to be extremely powerful for proving results about natural numbers.

We now generalise this idea to other so-called *inductively defined* sets. The definition (Definition 12.2.9) is a little tricky to digest, but the idea is relatively simple: an inductively defined set X is one whose elements are built out of some specified *basic elements* (such as 0) by iterating some specified operations, called *constructors* (such as the successor operation)—every element of X should either be a basic element, or should be built in a unique way out of simpler elements of X by using a constructor.

Each inductively defined set X will have its own induction principle: which says that if a property is true of all of the basic elements of X, and if all of its constructors preserve the truth of the property, then the property is true of all of the elements of X. We will prove this in Theorem 12.2.28.

Before jumping into the definition of an inductively defined set, it is helpful to see some examples. The first example is familiar.

Example 12.2.1

The set $\mathbb{N}$ of natural numbers is *the* canonical example of an inductively defined set. The Peano axioms (Definition 4.1.1) tell us that every element of $\mathbb{N}$ can be obtained by starting from 0 and applying the successor operation ('plus one') some finite number of times. In particular, every natural number is either 0, or is the successor of a unique natural number.

We can think of $\mathbb{N}$ as being the set *generated by* the following rules:

(i) $0 \in \mathbb{N}$; and

(ii) If $n \in \mathbb{N}$, then $n+1 \in \mathbb{N}$.

Certainly these rules are *true*, but that is not enough to fully characterise the set of natural numbers: both rules are also true with $\mathbb{N}$ replaced by $\mathbb{Z}$ or $\mathbb{R}$, for example.

But what it means to say that $\mathbb{N}$ is 'generated by' these rules is that:

- Each time one of the rules is applied, we obtain a *new* natural number. In particular:
  - ◇ Since $0 \in \mathbb{N}$ and rule (ii) must give us a new natural number every time we apply it, we must have $n+1 \neq 0$ for all $n \in \mathbb{N}$. This is exactly what Definition 4.1.1(i) says; note that this fails with $\mathbb{N}$ replaced by $\mathbb{Z}$ or $\mathbb{R}$ since we have $(-1)+1=0$.
  - ◇ If we apply rule (ii) to two different natural numbers, we must obtain two different results. Thus for all $m,n \in \mathbb{N}$, if $m+1=n+1$, then $m=n$. This is exactly what Definition 4.1.1(ii) says.
- Only those things that can be obtained by applying rules (i) and (ii) some finite number of times are natural numbers. This is exactly what Definition 4.1.1(iii) says.

Thus we can think of the natural numbers as being precisely those things that are given to us by applying rules (i) and (ii).

The next example concerns words over an alphabet, which we studied in Section 6.2 in the context of determining when a set is countable—see Definition 6.2.26.

Example 12.2.2

Let Σ be a fixed alphabet. The set Σ^* of words over Σ is built by starting from the empty word ε by appending elements $a \in \Sigma$. Thus Σ^* is generated by the following rules:

(i) $\varepsilon \in \Sigma^*$; and

(ii) For all $a \in \Sigma$, if $w \in \Sigma^*$, then $wa \in \Sigma^*$.

The fact that Σ^* is *generated by* these rules means that:

- Each time we append an element $a \in \Sigma$ to the end of a word $w \in \Sigma^*$, the resulting word wa is a new element of Σ^*; and
- Only those things obtained by appending elements $a \in \Sigma$ to the end of the empty word ε some finite number of times are elements of Σ^*.

We can actually think of rule (ii) as being a family of rules, one rule for each $a \in \Sigma$. For fixed a, the rule says 'for all w, if $w \in \Sigma^*$, then $wa \in \Sigma^*$'. This viewpoint is useful because it allows us to restrict our attention to rules that only have variable elements of the set being inductively defined in the hypothesis.

We alluded to the inductive character of propositional formulae in Section 1.1. Although we were not able to make the idea precise at the time, we are now well on our way to doing so.

Example 12.2.3

Given a set P, consider the set $\mathscr{L}(P)$ of all propositional formulae with propositional variables from P and logical operators from the set $\{\wedge, \vee, \Rightarrow, \neg\}$. Then $\mathscr{L}(P)$ is built out of the elements of P using these logical operators.

Specifically, $\mathscr{L}(P)$ is generated by the following rules:

(i) For each $p \in P$, we have $p \in \mathscr{L}(P)$;

(ii) If $\varphi, \psi \in \mathscr{L}(P)$, then $\varphi \wedge \psi \in \mathscr{L}(P)$;

(iii) If $\varphi, \psi \in \mathscr{L}(P)$, then $\varphi \vee \psi \in \mathscr{L}(P)$;

(iv) If $\varphi, \psi \in \mathscr{L}(P)$, then $\varphi \Rightarrow \psi \in \mathscr{L}(P)$;

(v) If $\varphi \in \mathscr{L}(P)$, then $\neg\varphi \in \mathscr{L}(P)$.

To say that $\mathscr{L}(P)$ is generated by these rules means that every propositional formula is (exclusively) either: a propositional variable, the conjunction of two simpler formulae, the disjunction of two simpler formulae, the implication formed by two simpler formulae, or the negation of a simper formula.

We can interpret (i) as being a family of rules, one for each $p \in P$, where for fixed p, the rule simply says '$p \in \mathscr{L}(P)$'. Just like we said in Example 12.2.2, this is useful because it removes variable elements of sets other than $\mathscr{L}(P)$ from the hypotheses of the rule.

Keeping these examples in mind, we will now work towards defining the notion of an *inductively defined set*. First we formalise the notion of a rule.

Definition 12.2.4

A (**finitary**) **rule** for an inductive definition is an expression of the form

$$(\mathbf{x} \mid \sigma(\mathbf{x})) \quad (\text{LaTeX code: } \texttt{\textbackslash mid})$$

where $\mathbf{x}$ is a (possibly empty) list of variables, and $\sigma(\mathbf{x})$ is some expression which may involve the variables in $\mathbf{x}$.

The number of variables in $\mathbf{x}$ is called the **arity** of the rule, and is denoted by $\mathrm{ar}(\sigma)$ (LaTeX code: `\mathrm{ar}`).

A quick note on terminology: a constructor of arity $r \in \mathbb{N}$ is called an *r-ary* constructor. For $r = 0, 1, 2$, we may say *nullary*, *unary* and *binary*, respectively.

We interpret the rule $(x, y, \ldots \mid \sigma(x, y, \ldots))$ as meaning 'if $x, y, \ldots$ are elements of the set being defined, then $\sigma(x, y, \ldots)$ is an element of the set being defined'.

Note that nullary rules take the form $(\mid a)$, where a is some expression with no variables; we would interpret such a rule as saying 'a is an element of the set being defined', with no hypotheses.

Example 12.2.5

The rules describing the natural numbers are $(\mid 0)$ and $(n \mid n+1)$. The rule $(\mid 0)$ can be interpreted to mean '0 is a natural number', and the rule $(n \mid n+1)$ can be interpreted to mean 'if n is a natural number, then $n+1$ is a natural number'. In the context of Example 12.2.1, this means that $(\mid 0)$ corresponds with rule (i) and $(n \mid n+1)$ corresponds with rule (ii).

Example 12.2.6

Let P be a set of propositional variables. The rules that describe the set $\mathscr{L}(P)$ of logical formulae over P, as described in Example 12.2.3, are given by

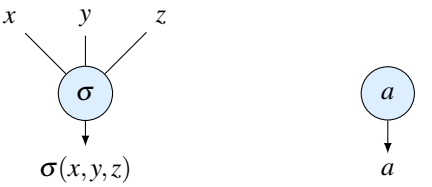
$$\underbrace{(\mid p)}_{\text{one for each } p \in P} \qquad (\varphi, \psi \mid \varphi \wedge \psi) \quad (\varphi, \psi \mid \varphi \vee \psi) \quad (\varphi, \psi \mid \varphi \Rightarrow \psi) \quad (\varphi \mid \neg\varphi)$$

The first of these rules says that every propositional variable $p \in P$ is a propositional formula. The next three say that if φ and ψ are propositional formulae, then so are $\varphi \wedge \psi$, $\varphi \vee \psi$ and $\varphi \Rightarrow \psi$. The last rule says that if φ is a propositional formula, then so is $\neg\varphi$.

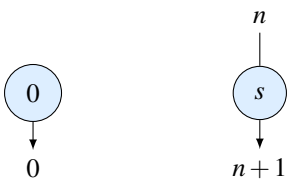
Exercise 12.2.7

Fix an alphabet Σ. Following Example 12.2.2, define rules that describe the set Σ^* of words over Σ. How would your rules need to be changed if the empty word ε were not allowed?

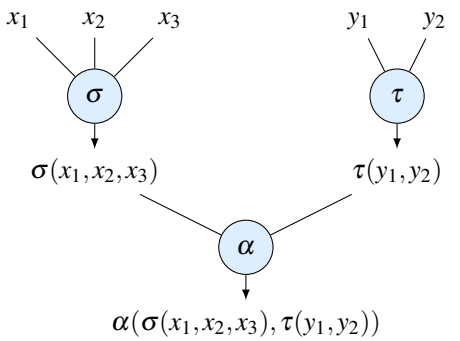
We can represent a rule $\sigma = (\mathbf{x} \mid \sigma(\mathbf{x}))$ diagrammatically by drawing a node with an input edge for each variable in the rule, and a single output output, representing the expression $\sigma(\mathbf{x})$. (Note that a nullary rule has no inputs.) For example:



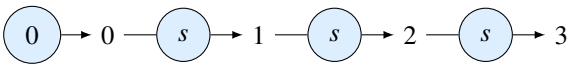
Thus the rules $0 = (\ \mid 0)$ and $s = (n \mid n+1)$ that describe the natural numbers can be expressed as follows:



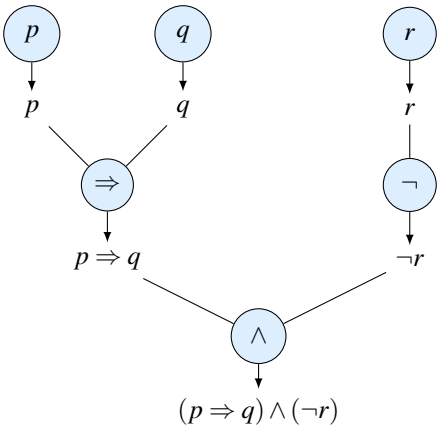
These diagrams can be pieced together to represent the result of applying a rule to the outputs of other rules. For example:



This can be useful for seeing how an element of a set is obtained by applying the rules. For example, the following diagram shows how the natural number 3 is obtained by applying the rules $(\ \mid 0)$ and $(n \mid n+1)$



The following diagram shows how the logical formula $(p \Rightarrow q) \wedge (\neg r)$ is obtained by applying the rules described in Example 12.2.6.



Exercise 12.2.8

Let $\Sigma = \{a,b,c,d\}$. Draw a diagram to represent how the word $cbadbd \in \Sigma^*$ can be obtained by applying the rules you defined in Exercise 12.2.7.

We are now ready to define an inductively defined set.

Definition 12.2.9

An **inductively defined set** is a set A equipped with a set R of rules and, for each rule $\sigma = (\mathbf{x} \mid \sigma(\mathbf{x})) \in R$, a function $f_\sigma : A^{\mathrm{ar}(\sigma)} \to A$, such that:

(a) For each $a \in A$, there is a unique rule $\sigma \in R$ and a unique $\mathrm{ar}(\sigma)$-tuple $\mathbf{a} = (a_0, a_1, \ldots, a_{\mathrm{ar}(\sigma)-1}) \in A^{\mathrm{ar}(\sigma)}$ such that $a = f_\sigma(\mathbf{a})$; and

(b) For all $B \subseteq A$, if $f_\sigma(\mathbf{a}) \in B$ for all $\sigma \in R$ and all $\mathbf{a} \in B^{\mathrm{ar}(\sigma)}$, then $B = A$.

Given a rule σ, the function $f_\sigma : A^{\mathrm{ar}(\sigma)} \to A$ is called the **constructor** associated with σ; we may at times refer to the **arity** of the constructor f_σ to mean the arity of the associated rule σ.

Note that nullary constructors are the same thing as elements of A. Indeed, $A^0 = \{()\}$, where $()$ is the empty list of elements of A, and so specifying a function $f_\sigma : A^0 \to A$ is equivalent to specifying an element $f_\sigma(())\in A$. In this sense, we may regard nullary constructors as *being* the elements of A—we call such elements *basic elements*.

Definition 12.2.10

A **basic element** of an inductively defined set A is an element of A that is the value of a nullary constructor $f_\sigma : A^0 \to A$. If $\sigma = (\mid a)$ is a nullary rule, we will denote this element by a—thus we have $a = f_\sigma(())\in A$ for all nullary rules $\sigma = (\mid a)$.

Considering basic elements separately from constructors of positive arity has its pros and cons, so we will take whichever approach is most convenient for us at any given point in time. Unless otherwise specified, we will not separate nullary constructors from the others.

We have already seen some examples of inductively defined sets—let's prove that they truly *are* inductively defined sets!

Proposition 12.2.11
The set $\mathbb{N}$ of natural numbers is inductively defined by the rules $(\ |\ 0)$ and $(n \mid n+1)$.

Proof
Since $(\ |\ 0)$ is a nullary rule, it will correspond to a basic element of $\mathbb{N}$—it may be no surprise that we take this element to be the natural number 0.

The rule $s = (n \mid n+1)$ induces a function $f_s : \mathbb{N} \to \mathbb{N}$; we take this to be the successor function, defined by $f_s(n) = n+1$ for all $n \in \mathbb{N}$.

Now we must verify conditions (a) and (b) of Definition 12.2.9.

(a) Let $n \in \mathbb{N}$.
- If $n = 0$, then since 0 is a basic element and $0 \neq n+1 = f_s(n)$ for any $n \in \mathbb{N}$, we have that '0' is the unique expression of 0 as a (nullary) constructor applied to (no) elements of $\mathbb{N}$.
- If $n > 0$, then $n-1 \in \mathbb{N}$ and $n = (n-1)+1 = f_s(n-1)$. Moreover if $m \in \mathbb{N}$ and $n = f_s(m)$, then $n = m+1$, so that $m = n-1$. Moreover $n \neq 0$, meaning that there is a unique rule (namely s) and a unique natural number m (namely $m = n-1$) such that $n = f_s(m)$.

(b) Let X be a set, and assume that $0 \in X$ and $f_s(n) \in X$ for all $n \in \mathbb{N}$. Then condition (iii) of Definition 4.1.1 ensures that $\mathbb{N} \subseteq X$.

Thus $\mathbb{N}$ is inductively defined by $(\ |\ 0)$ and $(n \mid n+1)$, as required. □

Example 12.2.12
Let P be a set of propositional variables. In order to exhibit $\mathscr{L}(P)$ as an inductively defined set, we should be slightly more precise about the role of *brackets* in propositional formulae than we have been so far.

For example, if we take the rules discussed in Example 12.2.3 literally, then $p \wedge q \vee r$ would be a valid logical formula. This is problematic for two reasons: first, does it mean $(p \wedge q) \vee r$, or $p \wedge (q \vee r)$? These are not logically equivalent, so the distinction matters. Second, this causes the 'uniqueness' part of Definition 12.2.9 to fail, since $p \wedge q \vee r = f_\wedge(p, f_\vee(q,r)) = f_\vee(f_\wedge(p,q), r)$.

To remedy this, we will require parentheses to be added whenever we introduce a logical operator. Thus the rules defining a logical formula are:

$$\underbrace{(\ |\ p)}_{\text{one for each } p \in P} \qquad (\varphi, \psi \mid (\varphi \wedge \psi)) \quad (\varphi, \psi \mid (\varphi \vee \psi)) \quad (\varphi, \psi \mid (\varphi \Rightarrow \psi)) \quad (\varphi \mid (\neg \varphi))$$

Thus $p \wedge q \vee r$ is not a valid element of $\mathscr{L}(P)$, but $((p \wedge q) \vee r)$ and $(p \wedge (q \vee r))$ are.

Exercise 12.2.13
Let Σ be an alphabet. Prove that Σ^* is inductively defined by the rules $(\ |\ \varepsilon)$ and $\sigma_a = (w \mid wa)$ for $a \in \Sigma$.

Exercise 12.2.14
Prove that $\mathbb{N}$ is inductively defined by the rules $(\ |\ 0)$, $(\ |\ 1)$ and $(n \mid n+2)$.

Defining new inductively defined sets

Now that we have seen what it means for an *already defined* set to be inductively defined, it would be useful to be able to *make definitions* of inductively defined sets. Based on what we have seen so far, in order to define an inductively defined set, all we should need to do is specify a set of rules that tell us how to generate its elements.

We now prove that it really is that simple!

Definition 12.2.15
Let R be a set of rules. The set **generated by** R is defined by $A = \bigcup_{n\in\mathbb{N}} A_n$, where the sets A_n for $n \in \mathbb{N}$ are defined recursively by:

- $A_0 = \{a \mid (\ |\ a) \in R\}$; and
- $A_{n+1} = \left\{ \sigma(\mathbf{a}) \;\middle|\; \mathbf{a} \in A_n^{\mathrm{ar}(\sigma)},\ (\mathbf{x} \mid \sigma(\mathbf{x})) \in R \right\}$.

That is, A_0 is the set of symbols to the right of the bar '|' in the nullary rules in R, and A_{n+1} is the set of all expressions obtained by substituting the elements of A_n symbol-by-symbol for the variables in the expressions to the right of the bar the other rules.

Example 12.2.16
Let N be the set generated by the rules $(\ |\ z)$ and $(n \mid s(n))$. Then

- $N_0 = \{z\}$, since $(\ |\ z)$ is the only nullary rule.
- N_1 is the result of applying the rules to the elements of N_0, of which there is only one, namely z. Applying the rule $(\ |\ z)$ (to no elements, since it is nullary) gives z; applying the rule $(n \mid s(n))$ to z gives $s(z)$. So $N_1 = \{z, s(z)\}$.
- Continuing, we get $N_2 = \{z, s(z), s(s(z))\}$, and so on.

We thus have

$$N = \bigcup_{n\in\mathbb{N}} N_n = \{z, s(z), s(s(z)), s(s(s(z))), \dots\}$$

This looks an awful lot like the set of all natural numbers; and indeed it is, provided we interpret $z = 0$ and $s(n) = n+1$ (like we did in Section 4.1).

Example 12.2.17
Let A be the set generated by the rules $(\ |\ \star)$ and $(x, y \mid [x, y])$. Then

- $A_0 = \{\star\}$;

- $A_1 = \{\star,\ [\star,\star]\}$;
- $A_2 = \{\star,\ [\star,\star],\ [\star,[\star,\star]],\ [[\star,\star],\star],\ [[\star,\star],[\star,\star]]\}$;
- … and so on.

Thus for each $n \in \mathbb{N}$, the set A_n consists of all parenthesised lists of '$\star$'s, where the list has length between 1 and 2^n (inclusive). This can be proved by induction on n.

Hence $A = \bigcup_{n\in\mathbb{N}} A_n$ is the set of all (finite) such lists.

Exercise 12.2.18
Let R be a set of rules for an inductive definition and let A be the set generated by R. Prove that if R has no nullary rules, then A is empty.

Exercise 12.2.19
Let R be a set of rules for an inductive definition, and let A be the set generated by R. Prove that if R is countable, then A is countable.

Exercise 12.2.20
Let R be a set of rules. Prove that the set A generated by R is inductively defined; for each rule $\sigma = (\mathbf{x} \mid \sigma(\mathbf{x}))$, the constructor $f_\sigma : A^{\mathrm{ar}(\sigma)} \to A$ is defined by

$$f_\sigma(\mathbf{a}) \;=\; \sigma(\mathbf{a})$$

where $\sigma(\mathbf{a})$ denotes the result of substituting the specific elements of A listed in $\mathbf{a}$ for the respective variables listed in $\mathbf{x}$. [For example, if $\sigma = (x,y \mid [x \odot y])$ and $\mathbf{a} = (2,5)$, then $\sigma(2,5) = [2 \odot 5]$.]

Structural recursion

The recursion theorem for the natural numbers (Theorem 4.1.2) says that we can define a function $h : \mathbb{N} \to X$ by specifying the value of $h(0)$, and for each $n \in \mathbb{N}$, specifying the value of $h(n+1)$ in terms of the value of $h(n)$. This makes intuitive sense: since every natural number is obtained from 0 by adding one some finite number of times, if we know the value of $h(0)$, and we know how the value of h changes when we add one to its argument, then we should know the value of $h(n)$ for all natural numbers n.

It turns out that there is nothing special about $\mathbb{N}$ here: exactly the same argument demonstrates that for *any* inductively defined set A, if we know what a function $h : A \to X$ does to its basic elements, and we know how its values change when we apply a constructor to its arguments, then we should know the value of $h(a)$ for all $a \in A$.

The proof of the structural recursion theorem is one of the times where treating basic elements ($\equiv$ nullary constructors) separately from constructors of positive arity makes the proof more difficult; so we will phrase it entirely in terms of constructors.

Theorem 12.2.21 · *Structural recursion theorem*

Let A be an inductively defined set, let X be a set, and for each rule σ, let

$$h_\sigma : A^{\mathrm{ar}(\sigma)} \times X^{\mathrm{ar}(\sigma)} \to X$$

Then there is a unique function $h : A \to X$ such that

$$h(f_\sigma(\mathbf{a})) \;=\; h_\sigma(\mathbf{a}, h(\mathbf{a}))$$

for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$, where $h(\mathbf{a}) = (h(a_1), \ldots, h(a_{\mathrm{ar}(\sigma)}))$.

Proof

Let $D = \{a \in X \mid h(a) \text{ is defined}\} \subseteq A$. Let σ be a rule let $\mathbf{a} \in D^{\mathrm{ar}(\sigma)}$. Define

$$h(f_\sigma(\mathbf{a})) = h_\sigma(\mathbf{a}, h(\mathbf{a}))$$

Note that this is defined, since the assumption that $\mathbf{a} \in D^{\mathrm{ar}(\sigma)}$ implies that $h(a_i)$ is defined for all $i \in [\mathrm{ar}(\sigma)]$. Therefore $f_\sigma(\mathbf{a}) \in D$. By Definition 12.2.9(b) we have $D = A$, so that $h(a)$ is defined for all $a \in A$, so that we have a function $h : A \to X$ as required.

To see that h is unique, let $h' : A \to X$ and suppose that h' satisfies the same equations as h, meaning that

$$h'(f_\sigma(\mathbf{a})) = h_\sigma(\mathbf{a}, h'(\mathbf{a}))$$

for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$.

Let $E = \{a \in A \mid h'(a) = h(a)\}$, let σ be a rule, and let $\mathbf{a} \in E^{\mathrm{ar}(\sigma)}$. Then $h'(a_i) = h(a_i)$ for all $i \in [\mathrm{ar}(\sigma)]$, so that

$$h'(f_\sigma(\mathbf{a})) = h_\sigma(\mathbf{a}, h'(\mathbf{a})) = h_\sigma(\mathbf{a}, h(\mathbf{a})) = h(f_\sigma(\mathbf{a}))$$

and hence $f_\sigma(\mathbf{a}) \in E$. By Definition 12.2.9(b) again, we have $E = A$, so that $h'(a) = h(a)$ for all $a \in A$. Therefore $h' = h$ by function extensionality. □

Strategy 12.2.22 · *Defining functions by structural recursion*

Let A be an inductively defined set and let X be a set. In order to specify a function $h : A \to X$, it suffices to define $h(a)$ for all basic elements $a \in A$, and for each rule σ, to define $h(f_\sigma(\mathbf{a}))$ in terms of the values of $h(a_i)$ for each $i \in [\mathrm{ar}(\sigma)]$.

Example 12.2.23

Let $\mathscr{L}(P)$ be the inductively defined set of propositional formulae over a set P of propositional variables. Define $h : \mathscr{L}(P) \to \mathbb{N}$ recursively as follows:

(i) Let $h(p) = 0$ for all $p \in P$;

(ii) For all $\varphi, \psi \in \mathscr{L}(P)$, let $h(\varphi \wedge \psi) = h(\varphi) + h(\psi) + 1$;

(iii) For all $\varphi, \psi \in \mathscr{L}(P)$, let $h(\varphi \vee \psi) = h(\varphi) + h(\psi) + 1$;

(iv) For all $\varphi, \psi \in \mathscr{L}(P)$, let $h(\varphi \Rightarrow \psi) = h(\varphi) + h(\psi) + 1$;

(v) For all $\varphi \in \mathscr{L}(P)$, let $h(\neg\varphi) = h(\varphi) + 1$.

By the structural recursion theorem, this completely determines the function h.

For example

$$\begin{aligned} & h((p \Rightarrow q) \wedge (\neg r)) \\ &= h(p \Rightarrow q) + h(\neg r) + 1 && \text{by (ii)} \\ &= [h(p) + h(q) + 1] + [h(r) + 1] + 1 && \text{by (iv) and (v)} \\ &= [0 + 0 + 1] + [0 + 1] + 1 && \text{by (i)} \\ &= 3 \end{aligned}$$

More generally, for each $\varphi \in \mathscr{L}(P)$, the value $h(\varphi)$ is the number of logical operators that appear in φ. We can prove this by *structural induction*—more on this soon.

Definition 12.2.24

Let A be an inductively defined set. The **rank** $\mathrm{rk}(a)$ (LaTeX code: `\mathrm{rk}`) of an element $a \in A$ is defined by recursively by letting $\mathrm{rk}(a) = 0$ for all basic elements a, and

$$\mathrm{rk}(f_\sigma(\mathbf{a})) = \max\{\mathrm{rk}(a_i) \mid i \in [\mathrm{ar}(\sigma)]\} + 1$$

for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$.

Intuitively, the rank of an element $a \in A$ tells us how many 'layers' of constructors (of positive arity) appear in the unique expression of a as constructors applied to basic elements.

Example 12.2.25

Let P be a set of propositional variables, let $p, q, r, s \in P$, and define

$$\varphi = ((p \wedge q) \vee (r \Rightarrow (\neg s))) \in \mathscr{L}(P)$$

Intuitively, s is the innermost propositional variable in φ, because in order to access s we must first go through $\vee$, and then $\Rightarrow$, and then $\neg$; therefore we have three layers of constructors, so we expect to have $\mathrm{rk}(\varphi) = 3$. We will verify that we are correct.

Since $p, q, r, s \in P$ are basic elements of $\mathscr{L}(P)$, we have

$$\mathrm{rk}(p) = \mathrm{rk}(q) = \mathrm{rk}(r) = \mathrm{rk}(s) = 0$$

Now we can iteratively find the ranks of the subformulae of φ:

- $\mathrm{rk}((p \wedge q)) = \max\{0, 0\} + 1 = 0 + 1 = 1$;
- $\mathrm{rk}((\neg s)) = \max\{\mathrm{rk}(s)\} + 1 = \max\{0\} + 1 = 0 + 1 = 1$;
- $\mathrm{rk}(r \Rightarrow (\neg s)) = \max\{\mathrm{rk}(r), \mathrm{rk}((\neg s))\} + 1 = \max\{0, 1\} + 1 = 1 + 1 = 2$;

and so finally we have

$$\mathrm{rk}(\varphi) = \max\{\mathrm{rk}((p \wedge q), (r \Rightarrow (\neg s))\} + 1 = \max\{1, 2\} + 1 = 2 + 1 = 3$$

as expected.

Example 12.2.26

The rank of a natural number n is n. Indeed we have $\mathrm{rk}(0) = 0$ since 0 is a basic element, and for all $n \in \mathbb{N}$ we have $\mathrm{rk}(n+1) = \max\{\mathrm{rk}(n)\} + 1 = \mathrm{rk}(n) + 1$. By the recursion theorem (take your pick from Theorem 4.1.2 or Theorem 12.2.21), we have $\mathrm{rk}(n) = n$ for all $n \in \mathbb{N}$.

Exercise 12.2.27

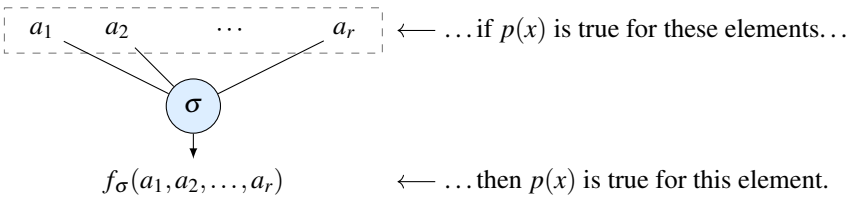
Let Σ be an alphabet and let Σ^* be the inductively defined set of words over Σ. Prove that for all $w \in \Sigma$, the rank $\mathrm{rk}(w)$ is equal to the length of W. [We regard the empty string ε to have length 0, despite the fact that the placeholder character ε is used to denote it.]

Structural induction

We now derive the structural induction principle, which is used for proving that a property $p(x)$ is true for all elements x of an inductively defined set A.

The idea behind structural induction is the same as the idea behind weak induction: since every element $a \in A$ is of the form $f_\sigma(\mathbf{a})$ for some (possibly nullary) rule σ, if we can prove that the truth of p is preserved by applying each constructor f_σ, then $p(x)$ must be true for all $x \in A$. [In the case of nullary constructors, this amounts to checking that $p(a)$ is true for each basic element $a \in A$.]

Thus in order to prove $\forall x \in A,\ p(x)$, we need to prove that for every rule $\sigma\ldots$



Let's prove that this intuition is valid.

Theorem 12.2.28 · *Structural induction principle*

Let A be an inductively defined set and let $p(x)$ be a logical formula with free variable $x \in A$. If for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ we have

$$[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a}))$$

then $p(x)$ is true for all $x \in A$.

Proof

Assume that for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ we have

$$[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a}))$$

Let $X = \{a \in A \mid p(a)\} \subseteq A$. Then for all $a \in A$ we have $a \in X$ if and only if $p(a)$ is true. Thus we have

$$[\forall i \in [\mathrm{ar}(\sigma)],\ a_i \in X] \Rightarrow f_\sigma(\mathbf{a}) \in X$$

But this is exactly the hypothesis of condition (b) of Definition 12.2.9, and so $A \subseteq X$.

Hence $A = X$, so that $p(a)$ is true for all $a \in A$, as required. □

Let's digest the statement of Theorem 12.2.28.

For a nullary rule $\sigma = (\ \mid a)$, the arity $\mathrm{ar}(\sigma)$ is 0, so $\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)$ is equivalent to $\forall i \in \varnothing,\ p(a_i)$, which is vacuously true (see Exercise 2.1.24), so that

$$[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a})) \quad \equiv \quad p(a)$$

Therefore, saying that $[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a}))$ is true for all rules σ and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ is equivalent to saying:

- $p(a)$ is true for all basic elements a; and
- $[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a}))$ is true for all rules σ of *positive* arity and all $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$.

This suggests the following proof strategy.

❖ **Strategy 12.2.29** · *Proof by structural induction*

In order to prove a proposition of the form $\forall a \in A,\ p(a)$, where A is an inductively defined set, it suffices to prove for all rules σ that

$$[\forall i \in [\mathrm{ar}(\sigma)],\ p(a_i)] \Rightarrow p(f_\sigma(\mathbf{a}))$$

Equivalently, it suffices to:

- For each basic element $a \in A$, prove $p(a)$—this is the **base case** for a;
- For each constructor σ of arity $r > 0$, let $a_1, a_2, \ldots, a_r \in A$ and assume that $p(a_i)$ is true for all $i \in [r]$, and prove that $p(f_\sigma(a_1, a_2, \ldots, a_r))$ is true—this is the **induction step** for σ.

The assumption $\forall i \in [r],\ p(a_i)$ in the induction step for σ is called the **induction hypothesis** for σ.

✎ **Example 12.2.30**

The structural induction principle for the inductively defined set $\mathbb{N}$ is exactly the same as the weak induction principle (Theorem 4.2.1). It says that to prove $\forall n \in \mathbb{N},\ p(n)$ is true, it suffices to:

- Prove $p(0)$ is true (since 0 is the only basic element); and
- Let $n \in \mathbb{N}$, assume that $p(n)$ is true, and prove that $p(n+1)$ is true (since the successor operation is the only constructor of positive arity).

Thus we recover weak induction as a special case of structural induction.

Example 12.2.31

Fix a set P of propositional variables and let $\mathscr{L}(P)$ be the inductively defined set of (properly bracketed) propositional formulae over P, as in Example 12.2.12.

We prove that every propositional formula $\varphi \in \mathscr{L}(P)$ has the same number of open brackets '(' as closed brackets ')'.

- (**Base cases**) The basic elements of $\mathscr{L}(P)$ are the propositional variables $p \in P$. So let $p \in P$; this is a logical formula with no open brackets and no closed brackets, so the number of open brackets is equal to the number of closed brackets.
- (**Induction step** for $\wedge$) Let $\varphi, \psi \in \mathscr{L}(P)$ and assume that φ and ψ each have the same number of open brackets as closed brackets. Say φ has a open brackets and a closed brackets, and ψ has b open brackets and b closed brackets, where $a, b \in \mathbb{N}$. Now

  $$f_\wedge(\varphi, \psi) \;=\; (\varphi \wedge \psi)$$

  Since φ contributes a open brackets and ψ contributes b open brackets, this has $a+b+1$ open brackets; likewise it has $a+b+1$ closed brackets.

  This completes the induction step for $\wedge$.
- (**Induction steps** for $\vee$ and $\Rightarrow$) These are identical to the induction step for $\wedge$—just replace $\wedge$ by the necessary logical operator throughout.
- (**Induction step** for $\neg$) Let $\varphi \in \mathscr{L}(P)$ and assume that φ has the same number of open brackets as closed brackets—say it has a of each, where $a \in \mathbb{N}$. Then

  $$f_\neg(\varphi) \;=\; (\neg\varphi)$$

  Since φ contributes a open brackets, this has $a+1$ open brackets; likewise it has $a+1$ closed brackets.

  This completes the induction step for $\neg$.

So by structural induction, it follows that every propositional formula $\varphi \in \mathscr{L}(P)$ has the same number of open brackets as closed brackets.

Exercise 12.2.32

Let P be a set of propositional variables. Prove by structural induction that the rank of a propositional formula $\varphi \in \mathscr{L}(P)$ is equal to the number of logical operators in φ.

A nice application of structural induction is to provide a formula for the *totient* of an integer n (Definition 7.3.26). We proved such a formula in Section 7.3, but that proof relied on the heavy machinery of the Chinese remainder theorem (Theorem 7.3.46)—here we give a proof without it.

Theorem 7.3.59

Let n be a nonzero integer. Then

$$\varphi(n) = |n| \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

where the product is indexed over distinct positive primes p dividing n.

Proof

Recall that $\varphi(-n) = \varphi(n)$ for all $n \in \mathbb{Z}$, so we may assume without loss of generality that $n \geqslant 0$—otherwise just replace n by $-n$ throughout. Moreover

$$\varphi(0) \;=\; 0 \;=\; 0 \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

so for the rest of the proof, assume that $n > 0$.

Let P be the set of all positive prime numbers, and let

$$A \;=\; \bigcup_{n \in \mathbb{N}} P^n \;=\; \{(p_1, p_2, \ldots, p_k) \mid k \in \mathbb{N},\ p_1, p_2, \ldots, p_k \in P\}$$

The elements of A are precisely lists of positive primes of finite length. Note that A is inductively defined by the rules

$$(\ |\ ()\ ) \quad \text{and} \quad \sigma_q = (x \mid (x, q)) \text{ for each } q \in P$$

That is, the empty list $()$ is an element of A, and for each $(p_1, p_2, \ldots, p_k) \in A$ and $q \in P$, we have $(p_1, p_2, \ldots, p_k, q) \in A$.

By the fundamental theorem of arithmetic (Theorem 7.2.12) there is a surjection $\Pi : A \to \{n \in \mathbb{Z} \mid n > 0\}$ defined by

$$\Pi(p_1, p_2, \ldots, p_k) \;=\; \prod_{i=1}^{k} p_i \;=\; p_1 \times p_2 \times \cdots \times p_k$$

for all $k \in \mathbb{N}$ and $p_1, p_2, \ldots, p_k \in P$. Note in particular that $\Pi(\,()\,) = 1$, where $()$ is the empty list.

We prove by structural induction on $(p_1, p_2, \ldots, p_k) \in A$ that the integer $n = \Pi(p_1, p_2, \ldots, p_k) > 0$ satisfies the formula in the statement of the theorem.

- (**Base case**) The unique basic element of A is the empty list $()$. Since $\Pi(\,()\,) = 1$, we must prove that the equation in the statement of the theorem is satisfied when $n = 1$. Well there are no primes p such that $p \mid 1$, and so the product $\displaystyle\prod_{p|1} \left(1 - \frac{1}{p}\right)$ is the empty product, which is equal to 1. Thus
$$\varphi(1) \;=\; 1 \;=\; 1 \cdot 1 \;=\; 1 \cdot \prod_{p|1} \left(1 - \frac{1}{p}\right)$$
as required.
- (**Induction step** for $q \in P$) Fix $(p_1, p_2, \ldots, p_k) \in A$, let $n = \Pi(p_1, p_2, \ldots, p_k)$ and assume that
$$\varphi(n) \;=\; n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$$
Note that $\Pi(p_1, p_2, \ldots, p_k, q) = qn$, and so we need to prove that
$$\varphi(qn) \;=\; qn \cdot \prod_{p|qn} \left(1 - \frac{1}{p}\right)$$
Now either $q \mid n$ or $q \nmid n$.

◇ Assume $q \mid n$. Then n and qn have the same prime divisors, and so for all $p \in P$ we have $p \mid n$ if and only if $p \mid qn$. Therefore:

$$\begin{aligned}
\varphi(qn) &= q\varphi(n) && \text{by Exercise 7.3.28} \\
&= qn \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right) && \text{by induction hypothesis} \\
&= qn \cdot \prod_{p|qn} \left(1 - \frac{1}{p}\right) && \text{as observed above}
\end{aligned}$$

as required.

◇ Assume $q \nmid n$. Then for all $p \in P$ we have $p \mid qn$ if and only if $p \mid n$ or $p = q$. Therefore:

$$\begin{aligned}
\varphi(qn) &= (q-1)\varphi(n) && \text{by Exercise 7.3.29} \\
&= (q-1)n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right) && \text{by induction hypothesis} \\
&= qn\left(1 - \frac{1}{q}\right) \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right) && \text{rearranging} \\
&= qn \cdot \prod_{p|qn} \left(1 - \frac{1}{p}\right) && \text{absorbing } 1 - \frac{1}{q} \text{ into the product}
\end{aligned}$$

as required.

In both cases, the formula in the statement of the theorem is satisfied by the integer qn.

By structural induction, the result follows. □

Uniqueness of inductive definitions

It would be nice if an inductive definition completely characterises the set that it defines. But this is slightly too much to ask; for example, the sets

$$\{0, 1, 2, 3, \dots\} \quad \text{and} \quad \{\varepsilon, \bullet, \bullet\bullet, \bullet\bullet\bullet, \dots\}$$

are both inductively defined by the rules $(\ \mid z)$ and $(n \mid s(n))$. In the first we take the basic element z to be the natural number 0, and for each natural number n we take $s(n)$ to be the natural number $n+1$; in the second, we take the basic element z to be the empty word ε, and for each string n we take $s(n)$ to be the string '$n\bullet$' (so for example $s(\bullet\bullet) = \bullet\bullet\bullet$).

But there is evidently a way of translating between these sets: we can define a function from the first to the second by sending each natural number n to the string $\bullet\bullet\cdots\bullet$, where there are n '$\bullet$'s in the string.

Thus the sets $\{0, 1, 2, 3, \dots\}$ and $\{\varepsilon, \bullet, \bullet\bullet, \bullet\bullet\bullet, \dots\}$ are 'essentially the same'—they are inductively defined by the same rules, and so the only real difference is how we *label* their elements.

The next theorem demonstrates that the same is true of inductively defined sets in general. That is,

if two sets A and B are inductively defined by the same rules, then the only way that they can differ is in the labels we use to denote their elements. Formally, this 'relabelling' establishes a bijection $h : A \to B$ between the two sets.

In fact, we prove something stronger: not only is there a bijection between the two sets, but the bijection respects the constructors—that is, if we apply a constructor in A to some elements and relabel the result, that is equivalent to relabelling the elements first and then applying the corresponding constructor in B. Even better, this bijection is the unique bijection that does so.

Thus to sum up what the following theorem tells us: inductively defined sets are *uniquely* determined by their defining rules, up to a *unique* bijection that is compatible with the constructors—this is as close to absolute uniqueness as we could possibly hope to get!

Theorem 12.2.33 · *Uniqueness of inductively defined sets*

Let A and B be two sets that are inductively defined by the same set of rules. For each rule σ, let $f_\sigma : A^{\mathrm{ar}(\sigma)} \to A$ and $g_\sigma : B^{\mathrm{ar}(\sigma)} \to B$ be the respective constructors.

There is a unique bijection $h : A \to B$ such that

$$h(f_\sigma(\mathbf{a})) \;=\; g_\sigma(h(\mathbf{a}))$$

where $h(\mathbf{a}) = (h(a_0), \ldots, h(a_{\mathrm{ar}(\sigma)-1})) \in B^{\mathrm{ar}(\sigma)}$.

Proof

Define $h : A \to B$ by structural recursion as follows: given a rule σ given $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ such that $h(a_i)$ has been defined for all $i \in [\mathrm{ar}(\sigma)]$, define

$$h(f_\sigma(\mathbf{a})) \;=\; g_\sigma(h(\mathbf{a}))$$

We just need to prove that h is a bijection, since evidently the other condition on h is satisfied by construction.

So define $k : B \to A$ by structural recursion on the same way; note that now we have

$$k(g_\sigma(\mathbf{b})) \;=\; f_\sigma(k(\mathbf{b}))$$

where $k(\mathbf{b}) = (k(b_0), \ldots, k(b_{\mathrm{ar}(\sigma)-1})) \in A^{\mathrm{ar}(\sigma)}$.

We prove that $k(h(a)) = a$ for all $a \in A$ by structural induction. To this end, let σ be a rule, let $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ and suppose that $k(h(a_i)) = a_i$ for all $i \in [\mathrm{ar}(\sigma)]$. Note that with the notation conventions we are using, this induction hypothesis can be more concisely stated as saying that $k(h(\mathbf{a})) = \mathbf{a}$.

Let $a = f_\sigma(\mathbf{a})$. Then

$$\begin{aligned}
k(h(a)) &= k(h(f_\sigma(\mathbf{a}))) && \text{by definition of } a \\
&= k(g_\sigma(h(\mathbf{a}))) && \text{by construction} \\
&= f_\sigma(k(h(\mathbf{a})) && \text{by construction} \\
&= f_\sigma(\mathbf{a}) && \text{by induction hypothesis} \\
&= a && \text{by definition of } a
\end{aligned}$$

This completes the induction step. So we have $k(h(a)) = a$ for all $a \in A$.

A similar proof by structural induction reveals that $h(k(b)) = b$ for all $b \in B$.

Thus k is an inverse for h, so that h is a bijection.

The fact that h is the *unique* such bijection is immediate from the fact that it is defined by structural recursion, with the recursive formula being equivalent to the assertion that h respects constructors. □

Example 12.2.34

Let Σ be an alphabet. Then leaf-labelled rooted planar binary trees over Σ are essentially the same as parenthesisations of lists of elements of Σ of positive length.

Indeed, let R be the set of rules defined by

$$R = \{(\mid a) \mid a \in \Sigma\} \cup \{(t_1, t_2 \mid [t_1, t_2])\}$$

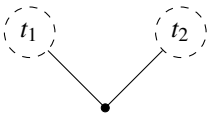
Like in Example 12.2.17, the inductively defined set A generated by R is given by the set of all parenthesisations of lists of elements of Σ. For example

$$[[[a_1, a_2], a_3], [a_4, a_5]] \in A$$

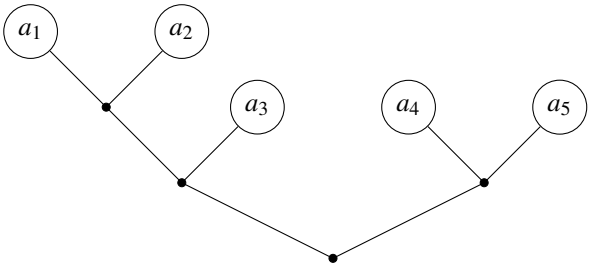
where $a_1, a_2, a_3, a_4, a_5 \in \Sigma$.

But the set T of all leaf-labelled rooted planar binary trees over Σ is also inductively defined by the rules in R. The basic element a corresponding to the rule $(\mid a)$ is precisely the tree consisting of just its root, which is labelled by the element $a \in \Sigma$:

and given trees t_1, t_2, the tree corresponding to $[t_1, t_2]$ is given by forming a tree with a root and two branches, pasting t_1 onto the left branch, and pasting t_2 onto the right branch:



By Theorem 12.2.33, there is a unique bijection $A \to T$ that is compatible with the constructors in each set. For example, the element $[[[a_1, a_2], a_3], [a_4, a_5]] \in A$ corresponds with the following tree:



⋆ Quotient-inductive sets

Some sets appear to be inductively defined, in the sense that its elements can be constructed from some basic elements by applying some constructors, except there may be more than one way of expressing each of its elements as a constructor applied to some elements of the set.

For example, consider the set $\mathbb{Z}$ of integers. Every integer n can be obtained from 0 by either adding 1 or subtracting 1 some number of times. Thus we'd like to say that $\mathbb{Z}$ is inductively defined by the rules

$$(\mid 0) \quad s = (x \mid x+1) \quad p = (x \mid x-1)$$

Here s stands for 'successor' and p stands for 'predecessor'. However, strictly speaking, $\mathbb{Z}$ is not inductively defined by these rules; for example

$$2 = 1+1 = f_s(1) \quad \text{and} \quad 2 = 3-1 = f_p(3)$$

and so 2 does not have a *unique* expression as a constructor applied to an integer.

In fact, the inductively defined set A generated by the three rules above consists of all strings of the form

$$0 \pm 1 \pm 1 \pm \cdots \pm 1$$

with each $\pm$ sign being either $+$ or $-$. We recover the corresponding integer by evaluating the string; this evaluation function defines a *surjection* $q : A \to \mathbb{Z}$; for example

$$q(0+1+1) = 2 \quad \text{and} \quad q(0+1+1+1-1+1-1-1+1) = 2$$

As made precise in Theorem 5.2.35, the fact that there is a surjection $A \to \mathbb{Z}$ means that we can think of $\mathbb{Z}$ as being a quotient of A by a suitable equivalence relation.

To see what this equivalence relation is, note that $(n+1)-1 = n$ and $(n-1)+1 = n$ for all $n \in \mathbb{Z}$. Thus in a string $0 \pm 1 \pm \cdots \pm 1$, we can add or remove instances of '$+1-1$' or '$-1+1$' in the string as we please, and we will obtain the same integer. So define an equivalence relation $\sim$ on A by letting $a \sim b$ if and only if the string b can be obtained from a by adding or removing '$+1-1$' or '$-1+1$' some number of times; for example

$$0+1+1 \quad \sim \quad 0+1-1+1+1 \quad \sim \quad 0+1+1+1-1+1-1-1+1$$

Then $\sim$ is an equivalence relation on A, and the evaluation function $q : A \to \mathbb{Z}$ descends to a bijection $\bar{q} : A/\!\sim \to \mathbb{Z}$.

We will call sets obtained in this way *quotient-inductive sets*.

Definition 12.2.35

Let A be an inductively defined set and let σ be a rule. A function $h : A \to X$ with domain A **respects** σ if, for all $\mathbf{a}, \mathbf{b} \in A^{\mathrm{ar}(\sigma)}$, we have

$$[\forall i \in [\mathrm{ar}(\sigma)],\ h(a_i) = h(b_i)] \quad \Rightarrow \quad h(f_\sigma(\mathbf{a})) = h(f_\sigma(\mathbf{b}))$$

Definition 12.2.36

A **quotient-inductive set** is a set Q equipped with a surjection $q : A \to Q$ for some inductively defined set A, such that q respects all of the rules of A.

The function q is called the **quotient map**; the **basic elements**, **constructors** and **rules** of Q are the images under q of those of A.

We will abuse notation slightly in the following way. Given a rule σ and $\mathbf{u} \in Q^{\mathrm{ar}(\sigma)}$, we will write $f_\sigma(\mathbf{u})$ instead of $q(f_\sigma(\mathbf{a}))$, where $\mathbf{a} \in A^{\mathrm{ar}(\sigma)}$ is such that $q(a_i) = u_i$ for each $i \in [\mathrm{ar}(\sigma)]$. Note that this is well-defined: the elements a_i for $i \in [\mathrm{ar}(\sigma)]$ exist since q is surjective, and the value of $q(f_\sigma(\mathbf{a}))$ is independent on the choices of a_i picked since q respects the rules of A.

Example 12.2.37

As we discussed, the rules $(\ |\ 0)$, $s = (x\ |\ x+1)$ and $p = (x\ |\ x-1)$ give rise to an inductively defined set A whose elements are strings of the form $0 \pm 1 \pm 1 \pm \cdots \pm 1$, with each '$\pm$' sign replaced with either $+$ or $-$.

The function $q : A \to \mathbb{Z}$ given by evaluating the string is evidently a surjection: indeed, given $n \in \mathbb{Z}$, if $n \geqslant 0$ then $n = q(0 + \underbrace{1 + \cdots + 1}_{n \text{ '1's}})$, and if $n < 0$ then $n = q(0 - \underbrace{1 - \cdots - 1}_{|n| \text{ '1's}})$.

Thus $\mathbb{Z}$ is a quotient-inductive set, with basic element 0 and constructors '$+1$' and '-1'.

Exercise 12.2.38

Prove that the set $\mathbb{Z}^+$ of all positive integers is a quotient-inductive set given by the rules $(\ |\ 1)$ and $(x\ |\ x \cdot p)$ for each positive prime $p \in \mathbb{Z}$. Describe the corresponding inductively defined set A and the quotient map $q : A \to \mathbb{Z}^+$ explicitly.

The reason we generalise the notion of an inductively defined set to that of a quotient-inductive set is that we can generalise the structural induction principle to such sets.

Note, however, that the structural recursion theorem depends in an essential way on the uniqueness of the representation of elements in terms of constructors, so we cannot expect the structural recursion theorem to generalise to quotient-inductive sets.

Theorem 12.2.39 · *Structural induction principle for quotient-inductive sets*

Let Q be a quotient-inductive set and let $p(x)$ be a logical formula with free variable $x \in Q$. If for all rules σ and all $\mathbf{x} \in Q^{\mathrm{ar}(\sigma)}$ we have

$$[\forall i \in [\mathrm{ar}(\sigma)],\ p(x_i)] \Rightarrow p(f_\sigma(\mathbf{x}))$$

then $p(x)$ is true for all $x \in Q$.

Proof
Let $q : A \to Q$ be the quotient map from the inductively defined set A, and let $\overline{p}(x)$ be the logical formula with free variable $x \in A$ defined by letting $\overline{p}(x)$ mean $p(q(x))$.

Since q is surjective, if $\overline{p}(x)$ is true for all $x \in A$, then $p(x)$ is true for all $x \in Q$. But the assumptions in the statement of the theorem imply that for all rules σ we have

$$\forall \mathbf{x} \in A^{\mathrm{ar}(\sigma)},\ [\forall i \in [\mathrm{ar}(\sigma)],\ \overline{p}(x_i)] \Rightarrow \overline{p}(f_\sigma(\mathbf{x})$$

and so $\overline{p}(x)$ is true for all $x \in A$ by (the usual version of) structural induction on A. □

Section 12.E

Chapter 12 exercises

Appendices

Appendix A

Proof-writing

Section A.1

Elements of proof-writing

Prior to taking a first course in pure mathematics, most people's exposure to the subject has a *computational* focus: the problems are questions that have a single correct answer, and the solutions consist of supplying the correct answer together with some explanation of how it was obtained ('showing your work'). Typically this amounts to a step-by-step sequence of calculations, algebraic manipulations and applications of formulae that result in the desired answer.

Pure mathematics has a different flavour: the task we face is to use the knowledge we have already established in order to discover new properties of the objects we are studying, and then communicate our discoveries in the form of a proof.

A given (true) statement may have many possible proofs, and the proofs may be qualitatively different from one another in many ways. Factors that may differ from one proof to another include length, verbosity, detail, motivation, proof strategies used, balance of words and notation, ...—the list goes on.

To complicate matters further, it might be that a proof is suitable in one setting but not another. For example, a detailed proof that $2+2=4$ using the definitions of '2', '4' and '$+$' is appropriate when studying the set of natural numbers from an axiomatic perspective (see Proposition 4.1.5), but this level of detail would not be appropriate when using the fact that $2+2=4$ in a counting argument (as in Section 8.1).

With all of this going on, it is difficult to know what is expected of us when we are asked to prove something.

The goal of this section is to shed some light on the following question:

What makes for an effective proof?

Learning how to write a proof is like learning any new style of writing: it is an iterative process with a cycle of practice, feedback and reflection. Writing a good proof requires patience and sincere effort, but the ends very much justify the means.

This section should not be read in isolation: it should be read either *after* working through a few chapters in the main part of the book, or as a reference for proof-writing in parallel with the rest of the book.

Proofs as prose

Someone with little mathematical background might look at a calculus textbook and ask 'where are all the numbers?', surprised by the sheer quantity of letters and symbols that appear. In the same vein, someone with little exposure to pure mathematics might look at this book and ask 'where are all the equations?'—it can be surprising to look at a mathematical text and see so many... *words*.

This is because we—pure mathematicians—use proofs as a tool for communicating our ideas, and we simply *need* to use words in order to make ourselves comprehensible to one another. Further-

more, in order to convince the wider mathematical community that our results are correct, another mathematician should be able to read and understand the proofs that we write, and then be able to communicate our arguments to other mathematicians.

This brings us to the first, and perhaps most important, writing principle for proofs.

❖ **Writing Principle A.1.1**
Mathematical proofs should be written (and read) as prose—in particular, it should be possible to read a proof aloud using full sentences.

To illustrate, consider the following proof that the composite of two injective functions is injective.

❝ **Extract A.1.2**
X, Y, Z sets, $f : X \to Y$, $g : Y \to Z$, f injective, g injective, $a, b \in X$:

$$\begin{aligned} g(f(a)) &= g(f(b)) && \text{def } \circ \\ f(a) &= f(b) && \because g \text{ injective} \\ a &= b && \because f \text{ injective} \end{aligned}$$

$\therefore g \circ f$ injective

This proof has many assets: it is correct, all of the variables used are quantified, and the steps are justified. But try reading it aloud. You will soon see that this proof does not read as prose—at least not easily.

For a short proof like this, that may be inconsequential, but for a more extended proof, or in a fully fledged mathematical paper (or textbook), it will not do. It is the duty of the proof-writer—that's you!—to make the reader feel as if they are being spoken to.

With this in mind, compare Extract A.1.2 with Extract A.1.3 below.

❝ **Extract A.1.3**
Let X, Y and Z be sets, let $f : X \to Y$ and $g : Y \to Z$, and assume that f and g are injective. Let $a, b \in X$ and assume that $(g \circ f)(a) = (g \circ f)(b)$. Then

$$\begin{aligned} & g(f(a)) = g(f(b)) && \text{by definition of } \circ \\ \Rightarrow\ & f(a) = f(b) && \text{by injectivity of } g \\ \Rightarrow\ & a = b && \text{by injectivity of } f \end{aligned}$$

Hence $g \circ f$ is injective, as required.

Despite the fact that Extract A.1.3 is not written exclusively using words, it is much easier to read it as prose, substituting phrases for the notation where needed—we will see more of this in Writing Principle A.1.8.

Here is a transcription of how Extract A.1.3 might be read aloud, with all the variables' letter names spelt out in italics.

Let X, Y and Z be sets, let F be a function from X to Y and G be a function from Y to Z, and assume that F and G are injective. Let A and B be elements of X and assume that G F of A is equal to G F of B. Then G of F of A equals G of F of B by definition of function composition, so F of A equals F of B by injectivity of G, and so A equals B by injectivity of F. Hence G F is injective, as required.

✎ **Exercise A.1.4**

Consider the following proof that the product of two odd integers is odd.

$a, b \in \mathbb{Z}$ both odd $\Rightarrow a = 2k+1, b = 2\ell+1, k, \ell \in \mathbb{Z}$

$$ab = (2k+1)(2\ell+1) = 4k\ell + 2k + 2\ell + 1 = 2(2k\ell + k + \ell) + 1$$

$2k\ell + k + \ell \in \mathbb{Z} \therefore ab$ odd

Rewrite the proof so that it is easier to read it as prose. Read it aloud, and transcribe what you read.

Us, ourselves and we

Different fields of study have different conventions about how their results should be communicated, and these conventions change over time. One such convention in mathematics, at least in recent decades, is the use of the first person plural ('we show that…').

There are several theories for why this practice has developed; for example, it is more modest than the first person singular ('I show that…'), and it implies that the reader is included in the proof process. But the *reasons* for using the first person plural are largely irrelevant for us—what matters is that this is the convention that prevails.

Some texts, particularly older ones, may deviate from this practice, such as by using the third person impersonal ('one shows that…') or the passive voice ('it is shown that…').

But although not universal, the first person plural is certainly the most commonplace, and so we shall adopt it.

❖ **Writing Principle A.1.5**

It is customary for mathematical proofs to be written in the first person plural.

Open almost any page in this textbook and you will see the first person plural everywhere. The next extract was taken almost at random to illustrate.

❝ **Extract A.1.6** · *taken from Theorem 7.1.26*

We proved in Theorem 7.1.14 that a greatest common divisor of a and b is a least element of the set

$$X = \{au + bv \mid u, v \in \mathbb{Z},\ au + bv > 0\}$$

So let $u, v \in \mathbb{Z}$ be such that $au + bv = d$. Then

$$a(ku) + b(kv) = k(au + bv) = kd = c$$

and so letting $x = ku$ and $y = kv$, we see that the equation $ax + by = c$ has a solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.

Some publishers require that mathematical variables not appear immediately after a punctuation mark, such as a comma or full stop. The first person plural comes to the rescue in the form of the phrase 'we have', which can usually be inserted after the offending punctuation mark in order to separate it from a statement that begins with a variable.

Extract A.1.7 · *taken from Example 5.1.21*

Let R be the relation on $\mathbb{R}$ defined for $a, b \in \mathbb{R}$ by $a \, R \, b$ if and only if $b - a \in \mathbb{Q}$. Then R is reflexive, since for all $a \in \mathbb{R}$, we have $a - a = 0 \in \mathbb{Q}$, so that $a \, R \, a$.

Without the phrase 'we have', this would have read as follows, violating the convention that variables should not follow punctuation marks.

> … for all $a \in \mathbb{R}$, $a - a = 0 \in \mathbb{Q}$, so …

Mathematical notation

What sets mathematics apart from most other areas of study is its heavy use of notation. Even in this introductory book, so much new notation is introduced that it can quickly become overwhelming. When writing a proof, it is sometimes important to take a step back and remember the reason why it is used.

Writing Principle A.1.8

Mathematical notation should be used in moderation with the goal of improving readability and precision.

The material covered in this textbook introduces a huge quantity of new notation, and it can be tempting to overuse it in proofs. Of course, how much notation is too much or too little is largely a matter of taste, but there is a balance to be struck.

To illustrate, what follows are three proofs that for all $a \in \mathbb{Z}$, if a^2 is divisible by 3, then a is divisible by 3.

The first proof can be read as prose, as long as you're willing to try hard enough to translate notation on the fly, but it is very heavy on the notation. Quantifiers and logical operators are *everywhere*.

Extract A.1.9

Let $a \in \mathbb{Z}$ and assume $3 \mid a^2$. Then $\exists k \in \mathbb{Z},\ a^2 = 3k$ by Definition 7.1.5. Also $\exists q, r \in \mathbb{Z},\ 0 \leqslant r < 3 \wedge a = 3q + r$ by Theorem 7.1.1. Now

- $r = 0 \Rightarrow a^2 = (3q)^2 = 3(3q^2) \Rightarrow \exists k \in \mathbb{Z},\ a^2 = 3k$ ✓
- $r = 1 \Rightarrow a^2 = (3q+1)^2 = 3(3q^2+2q)+1 \Rightarrow \neg\exists k \in \mathbb{Z},\ a^2 = 3k \rightsquigarrow$ contradiction by Theorem 7.1.1
- $r = 2 \Rightarrow a^2 = (3q+2)^2 = 3(3q^2+4q+1)+1 \Rightarrow \neg\exists k \in \mathbb{Z},\ a^2 = 3k \rightsquigarrow$ contradiction by Theorem 7.1.1

Now $(r = 0 \vee r = 1 \vee r = 2) \wedge (r \neq 1 \wedge r \neq 2) \Rightarrow r = 0 \Rightarrow 3 \mid a$, as required.

On the other extreme, the next proof is very easy to read as prose, but its sheer verbosity makes the actual mathematics harder to follow.

Extract A.1.10

Suppose the square of an integer a is divisible by three. Then a^2 can be expressed as three multiplied by another integer k by definition of division. Furthermore, by the division theorem, a can itself be expressed as three multiplied by the quotient q of a upon division by three, plus the remainder r upon division by three; and the remainder r is greater than equal to zero and is less than three.

Now if the remainder r is zero, then $a^2 = (3q)^2 = 3(3q^2)$, which is consistent with our assumption that a^2 is three multiplied by an integer, since $3q^2$ is an integer. If the remainder r is one, then $a^2 = (3q+1)^2 = 3(3q^2+2q)+1$, which implies by the division theorem that a^2 cannot be expressed as three multiplied by an integer, contradicting our assumption. If the remainder r is equal to two, then $a^2 = (3q+2)^2 = 3(3q^2+4q+1)+1$, which again implies by the division theorem that a^2 cannot be expressed as three multiplied by an integer, contradicting our assumption again.

Since the only case that did not lead to a contradiction was that in which the remainder r was equal to zero, it follows that a is divisible by three.

The next extract strikes a better balance. It uses enough words to make reading it as prose easier than in Extract A.1.9, but it uses enough notation to keep it much more concise and navigable than Extract A.1.10.

Extract A.1.11

Let $a \in \mathbb{Z}$ and assume that $3 \mid a^2$. Then $a^2 = 3k$ for some $k \in \mathbb{Z}$ by Definition 7.1.5. Moreover, by the division theorem (Theorem 7.1.1), there exist $q, r \in \mathbb{Z}$ with $0 \leqslant r < 3$ such that $a = 3q + r$. Now

- If $r = 0$, then $a^2 = (3q)^2 = 3(3q^2)$. Letting $k = 3q^2$, we see that this is consistent with our assumption that $3 \mid a$.
- If $r = 1$, then $a^2 = (3q+1)^2 = 3(3q^2+2q)+1$. By the division theorem, it follows that $3 \nmid a^2$, contradicting our assumption.
- If $r = 2$, then $a^2 = (3q+2)^2 = 3(3q^2+4q)+1$. By the division theorem again, it follows that $3 \nmid a^2$, contradicting our assumption.

Since $r = 0$ in the only case that is consistent with our assumption, it follows that $a = 3q$, and so $3 \mid a$, as required.

Observe that one of the main differences between the notation-heavy Extract A.1.9 and the more reasonable Extract A.1.11 is that the latter does not use logical operators or quantifiers—they are replaced by their corresponding English translations.

While logical operators and quantifiers—particularly $\Rightarrow$, $\Leftrightarrow$ and $\forall$—do have their place in proofs, it is often best to tread lightly. The role of symbolic logic is to help you to figure out how to prove a proposition, and how to word its proof (more in this later); but the proof in the end should not typically contain much in the way of logical notation.

Think of logical notation as being like an engine in a car, a circuit board in a computer, or an internal organ in a guinea pig—they make it work, but you don't want to see them!

In line with Writing Principle A.1.1, when we use mathematical notation, we should be careful to make sure that what we write can still be read as prose. To illustrate, let's recall Extract A.1.3.

Extract A.1.3

Let X, Y and Z be sets, let $f : X \to Y$ and $g : Y \to Z$, and assume that f and g are injective. Let $a, b \in X$ and assume that $(g \circ f)(a) = (g \circ f)(b)$. Then

$$\begin{aligned} & g(f(a)) = g(f(b)) && \text{by definition of } \circ \\ \Rightarrow\ & f(a) = f(b) && \text{by injectivity of } g \\ \Rightarrow\ & a = b && \text{by injectivity of } f \end{aligned}$$

Hence $g \circ f$ is injective, as required.

The uses of notation are highlighted in the following transcription.

> Let X, Y and Z be sets, let F be a function from X to Y and G be a function from Y to Z, and assume that F and G are injective. Let A and B be elements of X and assume that G F of A is equal to G F of B. Then G of F of A equals G of F of B by definition of function composition, so F of A equals F of B by injectivity of G, and so A equals B by injectivity of F. Hence G F is injective, as required.

Observe that the kind of phrase that is read aloud depends on how the notation is being used within a larger sentence. For example, consider the following extract.

Extract A.1.12

Let $n \in \mathbb{Z}$ and suppose that n is even. Then $n = 2k$ for some $k \in \mathbb{Z}$.

Read aloud, we would say something like:

> Let N be an integer and suppose that N is even. Then N equals two K for some integer K.

Despite the fact that '$n \in \mathbb{N}$' and '$k \in \mathbb{N}$' here differ only in the letter that is used, they are read aloud differently because they are playing different roles in the sentence—the former is used as a verb phrase, and the latter as a noun phrase.

✎ **Exercise A.1.13**

Consider the following notation-heavy proof that $X \cap (Y \cup Z) \subseteq (X \cap Y) \cup (X \cap Z)$ for all sets X, Y and Z.

Let X, Y and Z be sets and let $a \in X \cap (Y \cup Z)$. Then

$$a \in X \wedge a \in Y \cup Z \Rightarrow a \in X \wedge (a \in Y \vee a \in Z)$$

- **Case 1:** $a \in Y \Rightarrow (a \in X \wedge a \in Y) \Rightarrow a \in X \cap Y \Rightarrow a \in (X \cap Y) \cup (X \cap Z)$.
- **Case 2:** $a \in Z \Rightarrow (a \in X \wedge a \in Z) \Rightarrow a \in X \cap Y \Rightarrow a \in (X \cap Y) \cup (X \cap Z)$.

$\therefore \forall a,\ a \in X \cap (Y \cup Z) \Rightarrow a \in (X \cap Y) \cup (X \cap Z)$

$\therefore X \cap (Y \cup Z) \subseteq (X \cap Y) \cup (X \cap Z)$.

Read the proof aloud and transcribe what you said. Then rewrite the proof with a more appropriate balance of notation and text.

The Greek alphabet

Greek letters are often used as variables in mathematical texts—sometimes the 26 letters in the Latin alphabet just aren't enough! The Greek alphabet has 24 letters and, like the Latin alphabet, it has two cases.

Name	Upper	Lower
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ε or ϵ
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ
Iota	I	ι
Lambda	Λ	λ
Kappa	K	κ
Mu	M	μ

Name	Upper	Lower
Nu	N	ν
Xi	Ξ	ξ
Omicron	O	o
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Upsilon	Υ	υ
Phi	Φ	φ or ϕ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

Note that several of the upper-case Greek letters are identical to upper-case Latin letters—in mathematics, these are not typically distinguished so that, for example, the letter 'H' will always be interpreted as an upper-case Latin letter H, rather than an upper-case Greek letter *eta*. For the same reason, the lower-case Greek letter *omicron* is not distinguished from the lower-case Latin letter O.

❖ **LaTeX tip**

In order to use Greek letters as mathematical variables using LaTeX:

- For the upper-case letters that are identical to a letter in the Latin alphabet, use the `\mathrm` command together with the Latin letter. For example, upper-case *rho* can be input as `\mathrm{P}`, even though *rho* corresponds phonemically with the Latin letter R.
- For the upper-case letters that are not identical to a letter in the Latin alphabet, the LaTeX command is given by their Greek name with an upper-case first letter. For example, the command `\Gamma` produces the ouput Γ.
- For the lower-case letters (except *epsilon* and *phi*—see below), the LaTeX command is given by their Greek name in lower case. For example, the command `\eta` produces the output η.

The variant forms of *epsilon* and *phi*, respectively ε (`\varepsilon`) and φ (`\varphi`), are preferred over ϵ (`\epsilon`) and ϕ (`\phi`). This is to better distinguish them from the symbols $\in$ (element symbol) and $\varnothing$ (empty set), respectively.

Section A.2

Vocabulary for proofs

The focus of Chapter 1 was on examining the logical structure of a proposition and using this to piece together a proof.

For example, in order to prove that every prime number greater than two is odd, we observe that 'every prime number greater than two is odd' takes the form

$$\forall n \in \mathbb{Z},\ [(n \text{ is prime} \wedge n > 2) \Rightarrow n \text{ is odd}]$$

By piecing together the proof strategies in Sections 1.1 and 1.2, we can see what a proof of this must look like:

- By Strategy 1.2.10, we must assume $n \in \mathbb{Z}$ and, without assuming anything about n other than that it is an integer, derive '$(n$ is prime $\wedge\ n > 2) \Rightarrow n$ is odd';
- By Strategy 1.1.22, we must assume 'n is prime $\wedge\ n > 2$' and derive that n is odd;
- By Strategy 1.1.9, we may separately assume that n is prime and $n > 2$.

Thus a proof that every prime number greater than two is odd would assume $n \in \mathbb{Z}$, assume that n is prime and $n > 2$, and then derive that n is odd.

All of this tells us how to *structure* a proof, but it does not tell us *what to write* in such a proof—that is the goal of this section.

This section provides some basic vocabulary and templates that can be used in proofs. We will use some notation conventions for introducing these template:

- [Square | brackets | and | bars] will be used where one of several choices can be made. For example, if you see

  [then | therefore | so | hence]

  it means that any of the words 'then', 'therefore', 'so' or 'hence' can be used.
- (Round brackets) will be used where a word or phrase is optional. For example if you see

  Let $x \in X$ (be arbitrary)

  it means that either 'Let $x \in X$' or 'Let $x \in X$ be arbitrary' can be used.
- ⟨Angle brackets⟩ will be used to provide instructions. For example if you see

  ⟨insert proof of p here⟩

  then you should write out a proof of the proposition p being referred to.

Breaking down a proof

As we discussed in Section 1.1, at every stage in a proof, there is some set of *assumptions* and some set of *goals*. The assumptions are the propositions that we may take to be true, either because we already proved them or because they are being temporarily assumed; and the goals are the propositions that remain to be deduced in order for the proof to be complete.

The words we use indicate to the reader how the assumptions and goals are changing. Thus the words we use allow the reader to follow our logical reasoning and verify our correctness.

For the next few pages, we will examine the proof strategies governing logical operators and quantifiers, as discussed in Chapter 1, and identify some words and phrases that can be used in a proof to indicate which strategy is being used.

Deductive reasoning

At its core, a proof is a sequence of deductions: starting with a proposition that is already known or assumed to be true, we deduce something new, and continue deducing new things until the proposition we deduce is the result we are trying to prove.

Each time we make a deduction, it should be clear why that deduction is valid, so it is good practice to justify each deduction we make by either *stating* or *citing* the reason.

✦ **Vocabulary A.2.1**

The following construction can be used to indicate that an assumption p is being used to deduce a goal q.

> [**then** | **therefore** | **so** (**that**) | **hence**] ⟨state q here⟩ [**by** ⟨cite p here⟩ | **since** ⟨state p here⟩]
> — *or* —
> **we know that** ⟨state p here⟩, **and so** ⟨state q here⟩
> — *or* —
> **it follows from** ⟨cite p here⟩ **that** ⟨state q here⟩

If p was the last thing to be proved in the proof, it may not be necessary to cite it or state it again explicitly—that can be inferred.

Here are a couple of examples of Vocabulary A.2.1 in action.

❝ **Extract A.2.2** · *taken from Proposition 2.2.8*

… Then $a \in Y$ since $X \subseteq Y$, so that $a \in X \cap Y$ by definition of intersection …

Notice the choice of words used to *state* versus to *cite* assumptions; in both the previous and next example, we used ‘since’ to state, and ‘by’ to cite.

Extract A.2.3 · *taken from Theorem 9.1.19*

If $a > 0$, then by Example 9.1.14 and Exercise 9.1.15, we have

$$\vec{x} \cdot \left(\frac{b}{a}\vec{x}\right) = \frac{b}{a}\|\vec{x}\|^2$$

which is non-negative if and only if $b \geqslant 0$, since we are assuming that $a \geqslant 0$.

Exercise A.2.4

In the following proof that every multiple of four is even, identify all instances where an assumption is stated or cited in order to justify a step in the proof.

> Let $n \in \mathbb{Z}$ and suppose that n is divisible by 4. Then by definition of divisibility we have $n = 4k$ for some $k \in \mathbb{Z}$. But then n is even, since $n = 4k = 2(2k)$ and $2k \in \mathbb{Z}$.

Assuming implications: reducing a problem to another problem

One of the ways that an assumption of the form $p \Rightarrow q$ is useful is that it 'reduces' the problem of proving q to that of proving p.

Vocabulary A.2.5

The following construction can be used to indicate to a reader that you are invoking an assumption of the form $p \Rightarrow q$ to prove a goal q by instead proving p.

> [**Since** ⟨state $p \Rightarrow q$ here⟩ | **By** ⟨cite $p \Rightarrow q$ here⟩], (**in order to prove** ⟨state q here⟩,) **it suffices to prove** ⟨state p here⟩.

We used this construction in the proof of the strong induction principle:

Extract A.2.6 · *taken from Theorem 4.3.2*

Notice that $q(n)$ implies $p(n)$ for all $n \geqslant n_0$, since given $n \geqslant n_0$, if $p(k)$ is true for all $n_0 \leqslant k \leqslant n$, then in particular $p(k)$ is true when $k = n$.

So it suffices to prove $q(n)$ is true for all $n \geqslant n_0$.

In the proof of Proposition 8.1.38, we used Vocabulary A.2.5 in the very first sentence to guide the rest of the proof.

Extract A.2.7 · *taken from Proposition 8.1.38*

First note that $\binom{n}{k} = \left|\binom{[n]}{k}\right|$ and $\binom{n}{n-k} = \left|\binom{[n]}{n-k}\right|$, so in order to prove $\binom{n}{k} = \binom{n}{n-k}$ it suffices by Strategy 6.1.15 to find a bijection $f : \binom{[n]}{k} \to \binom{[n]}{n-k}$.

Proving implications: introducing assumptions

Several kinds of logical formulae are proved by introducing new assumptions into a proof. For example:

- Strategy 1.1.22 says that an implication $p \Rightarrow q$ can be proved by assuming p and deriving q.
- Strategy 1.2.10 says that a universally quantified proposition $\forall x \in X,\ p(x)$ can be proved by introducing a new variable x, assuming $x \in X$, and deriving $p(x)$.
- Strategy 1.1.36 says that a negation $\neg p$ can be proved by assuming p and deriving a contradiction.

Vocabulary A.2.8

The words **assume** and **suppose** can be used to introduce a new assumption p into a proof.

> | [**assume** | **suppose**] ⟨state p here⟩.

The proposition p may then be used in the proof.

In the following extract, observe how the introduced assumption is used later in the proof.

Extract A.2.9 · *taken from Theorem 3.1.25*

Assume $A = B$ and let $x \in X$. Then

$$\begin{aligned} \chi_A(x) = 1 &\Leftrightarrow x \in A && \text{by definition of } \chi_A \\ &\Leftrightarrow x \in B && \text{since } A = B \\ &\Leftrightarrow \chi_B(x) = 1 && \text{by definition of } \chi_B \end{aligned}$$

Proving conjunctions: breaking into steps

Often a goal in a proof has the form $p \wedge q$—for example, in order to prove a function $f : X \to Y$ is a bijection, we can prove that f is injective **and** f is surjective; and in order to prove that a relation $\sim$ is an equivalence relation, we can prove that $\sim$ is reflexive **and** symmetric **and** transitive.

In these cases, we can split into steps.

Vocabulary A.2.10

To indicate to a reader that you are proving a conjunction $p \wedge q$ by proving p and q individually, you can say that you are breaking into **steps**. For example:

> - **Step 1:** (⟨state p here⟩) ⟨insert proof of p here⟩.
> - **Step 2:** (⟨state q here⟩) ⟨insert proof of q here⟩.

This can be generalised to conjunctions of more than two propositions. Explicitly enumerated steps are not usually necessary, as long as it is clear what you are aiming to achieve in each step.

A common example of where steps are used is in proving propositions of the form $p \Leftrightarrow q$, which is shorthand for $(p \Rightarrow q) \wedge (q \Rightarrow p)$. In these cases, the two steps are the proofs of $p \Rightarrow q$ and $q \Rightarrow p$; the steps can then be labelled as $(\Rightarrow)$ and $(\Leftarrow)$, respectively.

Extract A.2.11 · *taken from Example 1.1.31*

Our goal is now to prove that 8 divides n if and only if 8 divides n'.

- $(\Rightarrow)$ Suppose 8 divides n. Since 8 divides n'', it follows from Exercise 0.40 that 8 divides $an + bn''$ for all $a, b \in \mathbb{Z}$. But

$$n' = n - (n - n') = n - n'' = 1 \cdot n + (-1) \cdot n''$$

  so indeed 8 divides n', as required.

- $(\Leftarrow)$ Suppose 8 divides n'. Since 8 divides n'', it follows from Exercise 0.40 that 8 divides $an' + bn''$ for all $a, b \in \mathbb{Z}$. But

$$n = n' + (n - n') = n' + n'' = 1 \cdot n' + 1 \cdot n''$$

  so indeed 8 divides n, as required.

Proofs of set equality by double containment also follow this format; in Section 2.1 we denoted the steps by $(\subseteq)$ and $(\supseteq)$, respectively.

Assuming disjunctions: breaking into cases

By Strategy 1.1.16, in order to use an assumption of the form $p \vee q$ ('p or q') to deduce a goal r, it suffices to show that r may be deduced from each of p and q—the idea here is that we may not know which of p or q is true, but that is fine since we derive r in both cases.

Vocabulary A.2.12

To indicate to a reader that you are using an assumption $p \vee q$ to prove a goal r, you can say that you are breaking into **cases**. For example:

> We know that either ⟨state p here⟩ or ⟨state q here⟩.
> - **Case 1:** Assume that ⟨state p here⟩**.** ⟨insert proof of r here⟩.
> - **Case 2:** Assume that ⟨state q here⟩**.** ⟨insert proof of r here⟩.
>
> In both cases we see that ⟨state r here⟩, as required.

Like with proofs involving steps (Vocabulary A.2.10), the explicit enumeration of cases is not usually necessary.

In the following extract, the assumption made is $(k \leqslant n) \vee (k > n)$, which is valid by the law of excluded middle (Strategy 1.1.43), and the goal is $2^k y \in D$.

Extract A.2.13 · *taken from Example 3.2.18*

Since $y \in D$, we must have $y = \frac{a}{2^n}$ for some $n \in \mathbb{N}$.

- If $k \leqslant n$ then $n-k \in \mathbb{N}$ and so $2^k y = \frac{a}{2^{n-k}} \in D$.
- If $k > n$ then $k-n > 0$ and $2^k y = 2^{k-n} a \in \mathbb{Z}$; but $\mathbb{Z} \subseteq D$ since if $a \in \mathbb{Z}$ then $a = \frac{a}{2^0}$. So again we have $2^k y \in D$.

In both cases we have $2^k y \in D$; and $f(2^k y) = y$, so that f is surjective.

Sometimes proofs by cases are extremely compact—so much so that you might not notice it—like in the next extract.

Extract A.2.14 · *taken from Proposition 7.2.7*

Since $a \mid p$, we have $a \in \{1, -1, p, -p\}$. If $a = \pm 1$, then a is a unit; if $a = \pm p$, then $b = \pm 1$, so that b is a unit. In any case, either a or b is a unit, and hence p is irreducible.

Note that cases were not explicitly labelled, but it was clear from context that cases were what were being used.

Proving negations: proof by contradiction

Vocabulary A.2.15

The following construction can be used in order to indicate that an assumption p is being introduced with the view of deriving a contradiction, thereby proving that p is false (that is, $\neg p$ is true).

> **Towards a contradiction, assume** ⟨state p here⟩.
> — *or* —
> **Assume** ⟨state p here⟩. **We will derive a contradiction.**

The following construction can be used in order to indicate that a contradiction to an assumption q has been reached from a contradictory assumption p.

> **This contradicts** ⟨cite q here⟩. **Therefore** ⟨state $\neg p$ here⟩.
> — *or* —
> . . . , **contrary to** ⟨cite q here⟩, **so that** ⟨state $\neg p$ here⟩.

Explicit reference to the proposition being contradicted is not always necessary if it is clear from context.

> **This is** [**a contradiction** | **nonsense** | **absurd** | **impossible**]. Therefore ⟨state $\neg p$ here⟩.

Extract A.2.16 · *taken from Theorem 6.1.24*

We proceed by contradiction. Suppose $\mathbb{N}$ is finite. Then $|\mathbb{N}| = n$ for some $n \in \mathbb{N}$, and hence $\mathbb{N}$ is either empty (nonsense, since $0 \in \mathbb{N}$) or, by Lemma 6.1.23, it has a greatest element g. But $g+1 \in \mathbb{N}$ since every natural number has a successor, and $g+1 > g$, so this contradicts maximality of g. Hence $\mathbb{N}$ is infinite.

In the next extract, the contradictory assumption was made a long time before the contradiction is reached, so special care is taken to reiterate that a contradiction was reached, and what the conclusion of the contradiction is.

Extract A.2.17 · *taken from Theorem 9.3.56*

Towards a contradiction, suppose that $e \in \mathbb{Q}$.

[*... many lines of proof omitted...*]

But this implies that $0 < c < 1$, which is nonsense since $c \in \mathbb{Z}$.

We have arrived at a contradiction, so it follows that e is irrational.

Proofs by contradiction need not be long and drawn out—sometimes you just need to say that something is false and give a brief justification. This is illustrated in the next extract.

Extract A.2.18 · *taken from Theorem 7.1.1*

... if $r \geqslant b$, then $r-b \in \mathbb{N}$ and $a = qb+r = (q+1)b+(r-b)$, which implies that $r-b \in X$; but $r-b < r$ since $b > 0$, so this contradicts minimality of r. Therefore $r < b$ ...

Proving universally quantified statements: introducing variables

As discussed in Strategy 1.2.10, proving that a property $p(x)$ hold for *all* elements x of a set X is done by introducing a new variable $x \in X$ and, assuming nothing about that variable other than that it is an element of X, deriving the truth of $p(x)$.

This sounds scary, but introducing the variable x is very easy to do.

Vocabulary A.2.19

The following constructions can be used to introduce a new variable x, referring to an arbitrary element of a set X.

> **Let** $x \in X$ (**be arbitrary**).
> — *or* —
> [**Take** | **Fix**] (**an** (**arbitrary**) **element**) $x \in X$.
> — *or* —
> **Given** $x \in X$, ...

Explicit use of the word 'arbitrary' can be useful to drive home the point that nothing is assumed about x other than that it is an element of X. In practice, it is optional.

If you turn to almost any page in the book, you will see an instance of Vocabulary A.2.19.

For example, in the next extract, the goal was to prove the universally quantified proposition meaning that $g \circ f$ is injective, that is $\forall a, b \in X, [(g \circ f)(a) = (g \circ f)(b) \Rightarrow a = b]$.

> **Extract A.2.20** · *taken from Proposition 3.2.4*
> ... let $a, b \in X$. We need to prove that
>
> $$(g \circ f)(a) = (g \circ f)(b) \quad \Rightarrow \quad a = b$$
>
> ...

Variables can also be introduced mid-sentence using the word 'given', like in the next extract.

> **Extract A.2.21** · *taken from Example 6.1.11*
> We need to prove $f(k) \in X$ for all $k \in [2n+1]$. Well given $k \in [2n+1]$, we have $0 \leqslant k \leqslant 2n$, and so
>
> $$-n \leqslant \underbrace{k-n}_{=f(k)} \leqslant 2n-n = n$$
>
> so that $f(k) \in X$ as claimed.

Proving existentially quantified statements: making definitions

When proving a statement of the form $\exists x \in X,\ p(x)$, we usually have to define an element $a \in X$ that makes $p(a)$ true, and then prove that $p(a)$ is true.

> **Vocabulary A.2.22**
> The following construction can be used to indicate that you are proving that there exists an element x of a set X such that $p(x)$ is true.
>
> > **[Define | Let]** ⟨define a here⟩. ⟨insert proof of $p(a)$ here⟩.
> > (**[It follows that | So | Therefore]** ⟨state $\exists x \in X,\ p(x)$ here⟩.)

Proofs of divisibility (see Definition 7.1.5) often use this construction, as in the following example.

> **Extract A.2.23** · *taken from Theorem 7.2.11*
> Define $q = ku + bv$; then
>
> $$b = abu + pbv = pku + pbv = p(ku + bv) = qp$$
>
> so $p \mid b$, as required.

Here is another example, which uses the word 'let'.

Extract A.2.24 · *taken from Theorem 9.1.16*

If $\vec{y} \neq \vec{0}$ then let $a = \|\vec{y}\|^2$ and $b = \vec{x} \cdot \vec{y}$; otherwise, let $a = 0$ and $b = 1$. In both cases, we have $a\vec{x} = b\vec{y}$ and a, b are not both zero.

Assuming existentially quantified statements: choosing elements

Another reason why variables might be introduces is because an *assumption* is existentially quantified. For example, if we know that 8 is even, then we know that $8 = 2k$ *for some* $k \in \mathbb{Z}$.

Vocabulary A.2.25

The following construction can be used to indicate that you are invoking an assumption of the form $\exists x \in X,\ p(x)$.

> | **Let** $a \in X$ **be such that** ⟨state $p(a)$ here⟩.

Again, proofs involving division yield several examples of when this construction is used.

Extract A.2.26 · *taken from Theorem 7.1.38*

... so by Proposition 7.1.35, we have $a \mid y_0 - y$ and $b \mid x - x_0$. Let $k, \ell \in \mathbb{Z}$ be such that $x - x_0 = kb$ and $y_0 - y = \ell a$.

Here is an example from Section 9.2 on the convergence of sequences; here, the existentially quantified assumption allowing us to introduce the variables a and b was cited.

Extract A.2.27 · *taken from Theorem 9.2.58*

Let (x_n) be a sequence of real numbers and let $a, b \in \mathbb{R}$ be such that $a < x_n < b$ for each $n \geqslant 0$—the numbers a and b exist since the sequence (x_n) is bounded.

'Without loss of generality'

Sometimes we need to prove a result in one of several cases, but the proofs in each case turn out to be very similar. This might be because the proofs are the same but with two variables swapped; or it might be because one case easily reduces to the other.

In such cases, instead of writing out the whole proof again with the minor changes made, we can inform the reader that this is happening and tell them how to recover the cases we do not prove.

Vocabulary A.2.28

When invoking an assumption of the form $p \vee q$, the phrase **without loss of generality** can be helpful to avoid splitting into cases when a proof in each case is essentially identical:

> | **(We may) assume** ⟨state p here⟩ **(without loss of generality)—otherwise** ⟨say how to modify the proof if q were true instead of p⟩.

The phrase ‘without loss of generality’ is so widespread that it is sometimes abbreviated to **wlog** (or WLOG), but this is best reserved for informal, hand-written proofs.

We only used the phrase ‘without loss of generality’ explicitly once in this book so far; this is recalled in the next extract.

Extract A.2.29 · *taken from proof of Theorem 7.3.59 in Section 12.2*

Recall that $\varphi(-n) = \varphi(n)$ for all $n \in \mathbb{Z}$, so we may assume without loss of generality that $n \geqslant 0$—otherwise just replace n by $-n$ throughout.

Nonetheless, we used the construction in Vocabulary A.2.28 at other times, as illustrated in the next extract:

Extract A.2.30 · *taken from Theorem 7.1.1*

We may assume that $r_0 \leqslant r_1$—otherwise, swap the roles of (q_0, r_0) and (q_1, r_1) in what follows.

Keeping track of everything

It can be difficult when writing a proof to keep track of what you have proved and what you have left to prove, particularly if you partially prove a result and then come back to it after a break.

On the other hand, it can be very difficult when *reading* a proof to keep track of what has been proved and what is left to prove!

As such, when writing a proof, it is of benefit both to you and to your future readers to insert language into your proofs that clarifies what has been done and what is left to do.

Vocabulary A.2.31

The following phrases can be used to indicate that the next goal in a proof is to prove a proposition p.

> **[We want | Our goal (now) is | It remains] to [show | prove]** ⟨state p here⟩
> — *or* —
> **To see that** ⟨state p here⟩, **note that** …

The first extract we look at is taken from the middle of a very long proof in Section 9.3. The entire highlighted sentence could be deleted and the proof would still be *correct*, but it would be much harder to follow—the sentence beginning ‘So let $K \geqslant N \ldots$’ might look as if it came out of nowhere.

Ext

… It remains to prove that $\left| \sum_{n=0}^{K} a_{\sigma(n)} - A \right| < \varepsilon$ for all $K \geqslant N$. So let $K \geqslant N$, …

Here is an example where we reiterate what we are about to prove before we prove it. Again, the highlighted text could be deleted without making the proof any less correct, but including it makes the proof easier to navigate.

Extract A.2.33 · *taken from Lemma 10.2.16*

To see that H is a bijection, note that the function $K : Y^X \to [\lambda]^{[\kappa]}$ defined by $K(\varphi) = k_\varphi = g^{-1} \circ \varphi \circ f$ is a bijection ...

When a key goal in a proof is reached, it is useful to say so.

Vocabulary A.2.34

The following phrases can be used to reiterate that a goal p has been proved.

| **[Hence | So (that) | Therefore | It follows that]** ⟨state p here⟩(, **as required**).

The next proof extract is a very typical example.

Extract A.2.35 · *taken from Theorem 4.1.2*

By condition (iii) of Definition 4.1.1, we have $\mathbb{N} \subseteq D$, so that $f(n)$ is defined for all $n \in \mathbb{N}$, as required.

Sometimes it might not be obvious that we have achieved the goal that we set out to achieve, in which case saying *why* the conclusion is 'as required' is also important.

This is illustrated in the next extract, where the goal was to prove that $|X \cup Y| = |X| + |Y| - |X \cap Y|$.

Extract A.2.36 · *taken from Proposition 6.1.18*

Hence $|X \cup Y| = m + n = |X| + |Y|$, which is as required since $|X \cap Y| = 0$.

Case study: proofs by induction

Many of the strategies for structuring and writing proofs are illustrated in proofs by induction. The strategy of proof by weak induction is described by the weak induction principle, recalled next.

Theorem 4.2.1 · *Weak induction principle*

Let $p(n)$ be logical formula with free variable $n \in \mathbb{N}$, and let $n_0 \in \mathbb{N}$. If

(i) $p(n_0)$ is true; and

(ii) For all $n \geqslant n_0$, if $p(n)$ is true, then $p(n+1)$ is true;

then $p(n)$ is true for all $n \geqslant n_0$.

Expressed as a single logical formula, the weak induction principle says that

$$[p(n_0) \wedge (\forall n \geqslant n_0, (p(n) \Rightarrow p(n+1))] \;\Rightarrow\; [\forall n \geqslant n_0,\ p(n)]$$

Let's pretend we know nothing about proofs by induction, and use the vocabulary in this section to construct a template for ourselves.

The goal is to prove $\forall n \geqslant n_0,\ p(n)$. We are going to invoke the weak induction principle (Theorem 4.2.1), and so using Vocabulary A.2.5 our proof should look something like this:

> ⟨insert proof of $p(n_0) \wedge (\forall n \geqslant n_0, (p(n) \Rightarrow p(n+1))$ here⟩
> It follows by the weak induction principle that $p(n)$ is true for all $n \geqslant n_0$.

Since our goal is the conjunction of two formulae, using Vocabulary A.2.10 tells us that we should break up the proof into steps:

> - **Step 1.** ⟨insert proof of $p(n_0)$ here⟩
> - **Step 2.** ⟨insert proof of $\forall n \geqslant n_0,\ p(n) \Rightarrow p(n+1)$ here⟩
>
> It follows by the weak induction principle that $p(n)$ is true for all $n \geqslant n_0$.

The goal in Step 2 is universally quantified, so we now need to introduce a variable n satisfying the condition that (n is a natural number and) $n \geqslant n_0$. We can do so using Vocabulary A.2.19.

> - **Step 1.** ⟨insert proof of $p(n_0)$ here⟩
> - **Step 2.** Fix $n \geqslant n_0$. ⟨insert proof of $p(n) \Rightarrow p(n+1)$ here⟩
>
> It follows by the weak induction principle that $p(n)$ is true for all $n \geqslant n_0$.

The goal in Step 2 is an implication, so using Vocabulary A.2.8, we obtain the following.

> - **Step 1.** ⟨insert proof of $p(n_0)$ here⟩
> - **Step 2.** Fix $n \geqslant n_0$ and assume ⟨state $p(n)$ here⟩.
>
>   ⟨insert proof of $p(n+1)$ here⟩
>
> It follows by the weak induction principle that $p(n)$ is true for all $n \geqslant n_0$.

To help the reader navigate the proof, we can use Vocabulary A.2.31 to reiterate the goal that we want to prove in Step 2.

> - **Step 1.** ⟨insert proof of $p(n_0)$ here⟩
> - **Step 2.** Fix $n \geqslant n_0$ and assume ⟨state $p(n)$ here⟩.
>
>   We need to prove ⟨state $p(n+1)$ here⟩.
>
>   ⟨insert proof of $p(n+1)$ here⟩
>
> It follows by the weak induction principle that $p(n)$ is true for all $n \geqslant n_0$.

This is now a perfectly good template for a proof of $\forall n \geqslant n_0,\ p(n)$. But this is exactly what a proof by induction looks like: 'Step 1' is the base case, 'Step 2' is the induction step, the assumption $p(n)$ in Step 2 is the induction hypothesis, and the goal $p(n+1)$ in Step 2 is the induction goal!

Appendix B

Mathematical miscellany

There have been a number of times in the book where we have avoided delving too deeply into the more technical or obscure aspects of a definition or proof. Usually this was because exploring these aspects was not central to the topic at hand, or because the details involved were sufficiently messy that providing all the details would obfuscate the main ideas being discussed.

This appendix provides a home for the comments we didn't make, the theorems we didn't prove, the details we didn't provide and the obscurities we didn't explore.

We begin with a quick glance at the *foundations of mathematics* in Section B.1. We will provide the axioms for Zermelo–Fraenkel set theory (ZF), which encodes all mathematical objects as sets and allows us to derive all mathematical objects from a collection of axioms. We will also demonstrate how to encode natural numbers, integers, rational numbers and complex numbers within this framework.

Section B.1

Set theoretic foundations

Naïve set theory and Russell's paradox

Naïve set theory formalises what we mean when we say 'a set is a collection of objects'. It consists of just two axioms:

- **Extensionality:** If two sets contain exactly the same elements, then they are equal.
- **Comprehension:** For any logical formula $p(x)$ with free variable x, there is a set whose elements are precisely those objects x such that $p(x)$ is true.

Note that the comprehension axiom implies that anything that we can define using set-builder notation $\{x \mid p(x)\}$ is a set.

Unfortunately, naïve set theory is not consistent—the most famous proof of this is due to Bertrand Russell, stated next.

Theorem B.1.1 · *Russell's paradox*
Under naïve set theory, the set $R = \{x \mid x \notin x\}$ satisfies neither $R \in R$ nor $R \notin R$. Thus naïve set theory is inconsistent.

Proof
The definition of R implies that $R \in R$ if and only if $R \notin R$, so that both the assertion that $R \in R$ and the assertion that $R \notin R$ lead to a contradiction. □

Several resolutions to Russell's paradox were suggested in the first half of the 20th century. Some did this by restricting how set-builder notation can be used, so that not *all* expressions of the form $\{x \mid p(x)\}$ define sets; others did this by assigning a *type* to each object and restricting how objects of one type may interact with objects of another type.

The foundation that was eventually adopted by the mathematical community at large is *Zermelo–Fraenkel set theory* (ZF), these days usually with the addition of the axiom of choice (ZFC). As such, this is the foundational system presented here.

It is worth noting, however, that other systems continue to be studied by mathematicians, logicians, philosophers and computer scientists. In particular, *Martin–Löf type theory* is a popular mathematical foundation used by computer scientists, as it much more closely mimics how programming languages work.

Zermelo–Fraenkel set theory

Zermelo–Fraenkel set theory overcomes Russell's paradox by deleting the axiom of comprehension and replacing it with axioms of two kinds: axioms that posit the existence of some sets (such as

the empty set), and axioms that describe how to construct new sets from sets already known to exist (such as their power set). In this theory, everything is a set, and so all mathematical objects—numbers, functions, relations, and so on—are encoded as sets of some kind. This is in contrast to *type theories*, which describe mathematical objects in terms of their intrinsic properties.

The first two axioms of ZF set theory that we introduce are the *axiom of extensionality* and the *axiom of foundation*. They concern the behaviour of the set elementhood relation $\in$, with the axiom of extensionality describing how it relates to equality of sets (as in Axiom 2.1.14), and axiom of foundation

Axiom B.1.2 · *Axiom of extensionality*

If two sets have the same elements, then they are equal.

$$\forall X, \forall Y, [(\forall a, a \in X \Leftrightarrow a \in Y) \Rightarrow X = Y]$$

A consequence of the axiom of extensionality is that two sets can be proved to be equal by proving that they contain the same elements.

Axiom B.1.3 · *Axiom of foundation*

Every inhabited set has an $\in$-minimal element.

$$\forall X, [(\exists y, y \in X) \Rightarrow \exists x, (x \in X \wedge \forall u \in X, u \notin x)]$$

The axiom of foundation states that $\in$ is a well-founded relation.

The axiom of foundation is mysterious at first sight, but it captures the idea that every set should built up from $\varnothing$ using set theoretic operations.

Lemma B.1.4

Under the axiom of foundation, we have $X \notin X$ for all sets X.

Proof

Let X be a set. The set $\{X\}$ is inhabited since $X \in \{X\}$, so by the axiom of foundation there is some $x \in \{X\}$ such that $u \notin x$ for all $u \in \{X\}$. But the only element of $\{X\}$ is X itself, so this says exactly that $X \notin X$. □

Exercise B.1.5

Use the axiom of foundation to prove that there is no sequence of sets $X_0, X_1, X_2, \ldots$ such that $X_{n+1} \in X_n$ for all $n \in \mathbb{N}$.

The next few axioms of ZF set theory posit the existence of certain sets or constructions of sets.

Axiom B.1.6 · *Empty set axiom*

There is a set with no elements.

$$\exists X, \forall x, x \notin X$$

The empty set axiom asserts the existence of $\varnothing$.

Axiom B.1.7 · *Pairing axiom*

For any two sets x and y, there is a set containing only x and y.

$$\forall x, \forall y, \exists X, \forall u, [u \in X \Leftrightarrow (u = x \vee u = y)]$$

The axiom of pairing asserts the existence of sets of the form $\{x, y\}$.

Axiom B.1.8 · *Union axiom*

The union of any family of sets exists and is a set.

$$\forall F, \exists U, \forall x, [x \in U \Leftrightarrow \exists X, (x \in X \wedge X \in F)]$$

The axiom of union asserts that if $F = \{X_i \mid i \in I\}$ is a family of sets then the set $U = \bigcup_{i \in I} X_i$ exists.

Axiom B.1.9 · *Power set axiom*

The set of all subsets of a set is a set.

$$\forall X, \exists P, \forall U, [U \in P \Leftrightarrow \forall u, (u \in U \Rightarrow u \in X)]$$

The axiom of power set asserts the existence of $\mathscr{P}(X)$ for all sets X.

Assuming only the previously stated axioms, it is entirely plausible that every set be finite. This isn't good news for us, since we want to be able to reason about infinite sets, such as the set $\mathbb{N}$ of natural numbers. The *axiom of infinity* asserts the existence of an infinite set using a clever set theoretic construction called the *successor set* operation.

Definition B.1.10

Given a set X, the **successor set** of X is the set X^+ defined by

$$X^+ = X \cup \{X\}$$

Lemma B.1.11

Let X and Y be sets. If $X^+ = Y^+$, then $X = Y$.

Proof

Assume $X^+ = Y^+$. Then

- We have $X \in X^+$, so $X \in Y^+ = Y \cup \{Y\}$, and so either $X = Y$ or $X \in Y$;
- We have $Y \in Y^+$, so $Y \in X^+ = X \cup \{X\}$, and so either $Y = X$ or $Y \in X$.

If $X = Y$ then we're done. Otherwise, we must have $X \in Y$ and $Y \in X$. But then we can define a sequence of sets by letting

$$X_n = \begin{cases} X & \text{if } n \text{ is even} \\ Y & \text{if } n \text{ is odd} \end{cases}$$

for all $n \in \mathbb{N}$. This sequence satisfies $X_{n+1} \in X_n$ for all $n \in \mathbb{N}$, since if n is even then

$$X_{n+1} = Y \in X = X_n$$

and if n is odd then

$$X_{n+1} = X \in Y = X_n$$

This contradicts Exercise B.1.5, so we must have $X = Y$, as claimed. □

Axiom B.1.12 · *Axiom of infinity*

There is an inhabited set containing successor sets of all of its elements.

$$\exists X, [(\exists u, u \in X) \wedge \forall x, (x \in X \Rightarrow x^+ \in X)]$$

Intuitively, the axiom of infinity tells us that there is a set X which contains (at least) a family of elements of the form u, u^+, u^{++}, u^{+++}, and so on—each of these elements must be distinct by Lemma B.1.11, so that X must be infinite.

Axiom B.1.13 · *Axiom of replacement*

The image of any set under any function is a set. That is, for each logical formula $p(x,y)$ with two free variables x, y, we have

$$\forall X, [(\forall x \in X, \exists! y, p(x,y)) \Rightarrow \exists Y, \forall y, y \in Y \Leftrightarrow \exists x \in X, p(x,y)]$$

Axiom B.1.14 · *Axiom of separation*

For any logical formula $p(x)$ with one free variable, and any set X, there is a set consisting of the elements of X satisfying $p(x)$.

$$\forall X, \exists U, \forall x, [x \in U \Leftrightarrow (x \in X \wedge p(x))]$$

The axiom of separation asserts the existence of sets of the form $\{x \in X \mid p(x)\}$.

Exercise B.1.15

Prove that the axioms of infinity and separation imply the existence of an empty set.

In light of Exercise B.1.15, the empty set axiom (Axiom B.1.6) is in fact redundant, in the presence of the other axioms. We keep it around for the sake of convenience.

Grothendieck universes

In Section 2.1 one of the first things we defined was a *universal set*, which we promptly forgot about and mentioned as little as possible. In this short subsection we briefly introduce the notion of a *Grothendieck universe*, named after the interesting (and influential) mathematician Alexander Grothendieck.

Definition B.1.16

A **Grothendieck universe** is a set $\mathscr{U}$ satisfying the following properties:

(i) The elements of $\mathscr{U}$ are sets;

(ii) For all $X \in \mathscr{U}$, if $x \in X$, then $x \in \mathscr{U}$;

(iii) $\mathbb{N}_{\mathsf{vN}} \in \mathscr{U}$ (see Construction B.2.5);

(iv) For all $X \in \mathscr{U}$, we have $\mathscr{P}(X) \in \mathscr{U}$;

(v) For all $I \in \mathscr{U}$ and all $\{X_i \mid i \in I\} \subseteq \mathscr{U}$, we have $\bigcup_{i \in I} X_i \in \mathscr{U}$.

The existence of a Grothendieck universe is not implied by the axioms of Zermelo–Frankel set theory (with or without the axiom of choice)—if it were, it would violate *Gödel's incompleteness theorem*, a result that is even further beyond the scope of this book than Grothendieck universes are!

Theorem B.1.17

Let $\mathscr{U}$ be a Grothendieck universe. The axioms of Zermelo–Fraenkel set theory are satisfied relative to $\mathscr{U}$. If the axiom of choice is assumed, then that is also satisfied relative to $\mathscr{U}$.

The upshot of Theorem B.1.17 is that although there is no universal set, if we assume the existence of a Grothendieck universe $\mathscr{U}$, then for the purposes of this book, we may relativise everything we do to $\mathscr{U}$ and *pretend* that $\mathscr{U}$ is indeed a universal set. And this is exactly what we did in Section 2.1.

Section B.2

Constructions of the number sets

In the most commonly used foundation of mathematics, Zermelo–Fraenkel set theory with the axiom of choice (discussed at length in Section B.1), every mathematical object is a set. This raises the question of: what about the other mathematical objects? For example, if numbers are sets, what sets are they? What about functions and relations?

For functions and relations, there is an easy cop-out: we simply identify them with their graphs. For example, we can pretend that the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ for all $x \in \mathbb{R}$ 'is' the set $\{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = x^2\}$.

For numbers, however, the answer is not quite so simple. For example, what set should the number number 3 be? And does it matter if we consider 3 to be a natural number or a real number?

This section presents one of many possible ways of encoding numbers—that is, natural numbers, integers, rational numbers, real numbers and complex numbers—as sets.

The natural numbers

We can use the framework provided by Zermelo–Fraenkel set theory (Section B.1) to provide set theoretic constructions of the number sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Indeed, if we want to reason about mathematics within the confines of ZF, we must encode everything (including numbers) as sets!

We will begin with a set theoretic construction of the natural numbers—that is, we will construct a notion of natural numbers in the sense of Definition 4.1.1. We will encode the natural numbers as sets, called *von Neumann natural numbers*. We will identify the natural number 0 with the empty set $\varnothing$, and we will identify the successor operation s with an operation involving sets.

Definition B.2.1

A **von Neumann natural number** is any set obtainable from $\varnothing$ by repeatedly taking successor sets (see Definition B.1.10). Write $0_{\mathsf{vN}} = \varnothing$ and $(n+1)_{\mathsf{vN}} = (n_{\mathsf{vN}})^+$; that is

$$0_{\mathsf{vN}} = \varnothing, \quad 1_{\mathsf{vN}} = \varnothing^+, \quad 2_{\mathsf{vN}} = \varnothing^{++}, \quad 3_{\mathsf{vN}} = \varnothing^{+++}, \quad 4_{\mathsf{vN}} = \varnothing^{++++}, \quad \dots$$

Example B.2.2

The first three von Neumann natural numbers are:

- $0_{\mathsf{vN}} = \varnothing$;
- $1_{\mathsf{vN}} = \varnothing^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\}$;
- $2_{\mathsf{vN}} = \varnothing^{++} = \{\varnothing\}^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}$.

Exercise B.2.3
Write out the elements of 3_{vN} $(= \varnothing^{+++})$ and of 4_{vN}.

Exercise B.2.4
Recall the definition of von Neumann natural numbers from Definition B.2.1. Prove that $|n_{\mathsf{vN}}| = n$ for all $n \in \mathbb{N}$.

Construction B.2.5
We construct the set $\mathbb{N}_{\mathsf{vN}}$ of all von Neumann natural numbers as follows. Let X be an arbitrary set satisfying the axiom of infinity (Axiom B.1.12), and then define $\mathbb{N}_{\mathsf{vN}}$ to be the intersection of all subsets of X that also satisfy the axiom of infinity—that is:

$$\mathbb{N}_{\mathsf{vN}} = \{x \in X \mid \forall U \in \mathscr{P}(X), [U \text{ satisfies the axiom of infinity} \Rightarrow x \in U]\}$$

The existence of $\mathbb{N}_{\mathsf{vN}}$ follows from the axioms of power set (Axiom B.1.9) and separation (Axiom B.1.14).

Theorem B.2.6
The set $\mathbb{N}_{\mathsf{vN}}$, zero element 0_{vN} and successor function $s : \mathbb{N}_{\mathsf{vN}} \to \mathbb{N}_{\mathsf{vN}}$ defined by $\mathsf{s}(n_{\mathsf{vN}}) = n^+_{\mathsf{vN}}$ for all $n_{\mathsf{vN}} \in \mathbb{N}_{\mathsf{vN}}$, define a notion of natural numbers.

Proof
We must verify Peano's axioms, which are conditions (i)–(iii) of Definition 4.1.1.

To prove (i), observe that for all sets X we have $X^+ = X \cup \{X\}$, so that $X \in X^+$. In particular, we have $n_{\mathsf{vN}} \in n^+_{\mathsf{vN}}$ for all $n_{\mathsf{vN}} \in \mathbb{N}_{\mathsf{vN}}$, and hence $n^+_{\mathsf{vN}} \neq \varnothing = 0_{\mathsf{vN}}$.

For (ii), let $m_{\mathsf{vN}}, n_{\mathsf{vN}} \in \mathbb{N}_{\mathsf{vN}}$ and assume that $m^+_{\mathsf{vN}} = n^+_{\mathsf{vN}}$. Then $m_{\mathsf{vN}} = n_{\mathsf{vN}}$ by Lemma B.1.11.

For (iii), let X be a set and suppose that $0_{\mathsf{vN}} \in X$ and, for all $n_{\mathsf{vN}} \in \mathbb{N}_{\mathsf{vN}}$, if $n_{\mathsf{vN}} \in X$, then $n^+_{\mathsf{vN}} \in X$. Then X satisfies the axiom of infinity (Axiom B.1.12), and so by Construction B.2.5 we have $\mathbb{N}_{\mathsf{vN}} \subseteq X$. □

In light of Theorem B.2.6, we may declare 'the natural numbers' to be the von Neumann natural numbers, and have done with it. As such, you can—if you want—think of all natural numbers in these notes as *being* their corresponding von Neumann natural number. With this in mind, we now omit the subscript 'vN', leaving implicit the fact that we are referring to von Neumann natural numbers.

However, there are many other possible notions of natural numbers. In Theorem B.2.8, we prove that any two notions of natural numbers are essentially the same, and so the specifics of how we actually define $\mathbb{N}$, the zero element and successor operation, are irrelevant for most purposes.

First we will prove the following handy lemma, which provides a convenient means of proving when a function is the identity function (Definition 3.1.12).

Lemma B.2.7

Let $(\mathbb{N}, z, s)$ be a notion of natural numbers, and let $j : \mathbb{N} \to \mathbb{N}$ be a function such that $j(z) = 0$ and $j(s(n)) = s(j(n))$ for all $n \in \mathbb{N}$. Then $j = \mathrm{id}_{\mathbb{N}}$.

Proof

By Theorem 4.1.2, there is a unique function $i : \mathbb{N} \to \mathbb{N}$ such that $i(z) = 0$ and $i(s(n)) = s(i(n))$ for all $n \in \mathbb{N}$. But then:

- $j = i$ by uniqueness of i, since j satisfies the same conditions as i; and
- $\mathrm{id}_{\mathbb{N}} = i$ by uniqueness of i, since $\mathrm{id}_{\mathbb{N}}(z) = z$ and $\mathrm{id}_{\mathbb{N}}(s(n)) = s(n) = s(\mathrm{id}_{\mathbb{N}}(n))$ for all $n \in \mathbb{N}$.

Hence $j = \mathrm{id}_{\mathbb{N}}$, as required. □

Theorem B.2.8

Any two notions of natural numbers are essentially the same, in a very strong sense. More precisely, if $(\mathbb{N}_1, z_1, s_1)$ and $(\mathbb{N}_2, z_2, s_2)$ are notions of natural numbers, then there is a unique bijection $f : \mathbb{N}_1 \to \mathbb{N}_2$ such that $f(z_1) = z_2$ and $f(s_1(n)) = s_2(f(n))$ for all $n \in \mathbb{N}_1$.

Proof

By applying the recursion theorem (Theorem 4.1.2) for $(\mathbb{N}_1, z_1, s_1)$, with $X = \mathbb{N}_2$, $a = z_2$ and $h : \mathbb{N}_1 \times \mathbb{N}_2 \to \mathbb{N}_2$ defined by $h(m,n) = s_2(n)$ for all $m \in \mathbb{N}_1$ and $n \in \mathbb{N}_2$, we obtain a function $f : \mathbb{N}_1 \to \mathbb{N}_2$ such that $f(z_1) = z_2$ and $f(s_1(n)) = s_2(f(n))$ for all $n \in \mathbb{N}_1$. This also gives us uniqueness of f, so it remains only to prove that f is a bijection.

Likewise, by applying Theorem 4.1.2 to $(\mathbb{N}_2, z_2, s_2)$, with $X = \mathbb{N}_1$, $a = z_1$ and $h : \mathbb{N}_2 \times \mathbb{N}_1 \to \mathbb{N}_1$ defined by $h(m,n) = s_1(n)$ for aall $m \in \mathbb{N}_2$ and $n \in \mathbb{N}_1$, we obtain a (unique!) function $g : \mathbb{N}_2 \to \mathbb{N}_1$ such that $g(z_2) = z_1$ and $g(s_2(n)) = s_1(g(n))$ for all $n \in \mathbb{N}_2$.

But then $g(f(z_1)) = g(z_2) = z_1$ and, for all $n \in \mathbb{N}_1$, we have

$$g(f(s_1(n))) = g(s_2(f(n))) = s_2(g(f(n)))$$

and so $g \circ f = \mathrm{id}_{\mathbb{N}_1}$ by Lemma B.2.7. Likewise $f \circ g = \mathrm{id}_{\mathbb{N}_2}$. Hence g is an inverse for f, so that f is a bijection, as required. □

With this settled, we can now say with confidence that it does not matter exactly which notion of natural numbers we are referring to when we write '$\mathbb{N}$' or '0' or '24': if you and I are using two different notions of natural numbers, then we can translate between our notions using the unique bijection from Theorem B.2.8.

Just for fun, though, let's explore a couple of other possible constructions and conventions that could be used.

Exercise B.2.9 demonstrates that the von Neumann construction (Construction B.2.5) is not the only way of constructing a notion of natural numbers from a set theoretic foundation.

✎ **Exercise B.2.9** · *Zermelo natural numbers*
Set $0_{\mathsf{Z}} = \varnothing$, and for all objects x define $s_{\mathsf{Z}}(x) = \{x\}$. Let

$$\mathbb{N}_{\mathsf{Z}} = \{0_{\mathsf{Z}},\ s_{\mathsf{Z}}(0_{\mathsf{Z}}),\ s_{\mathsf{Z}}(s_{\mathsf{Z}}(0_{\mathsf{Z}})),\ s_{\mathsf{Z}}(s_{\mathsf{Z}}(s_{\mathsf{Z}}(0_{\mathsf{Z}}))),\ \dots\}$$

Prove that $(\mathbb{N}_{\mathsf{Z}}, 0_{\mathsf{Z}}, s_{\mathsf{Z}})$ is a notion of natural numbers; this construction of the natural numbers is called the **Zermelo construction**.

In this book we have used the convention that the natural numbers start at 0; however, in Chapter 0 we briefly mentioned that some authors choose to start them at 1 instead. Exercise B.2.10 shows that in fact we can, in theory, adopt a convention where *any* integer is the least natural number, and everything will work out just fine—or, at least, it will from the perspective of Peano's axioms, so that we can still use recursion and induction.

✎ **Exercise B.2.10** · *Natural numbers with an arbitrary base*
Let $b \in \mathbb{Z}$, and define $\mathbb{Z}_{\geqslant b} = \{n \in \mathbb{Z} \mid n \geqslant b\}$; for example, $\mathbb{Z}_{\geqslant 1} = \{1, 2, 3, 4, 5, \dots\}$. Let $s_b : \mathbb{Z}_{\geqslant b} \to \mathbb{Z}_{\geqslant b}$ be the function defined by $s_b(n) = n + 1$ for all $n \in \mathbb{Z}_{\geqslant b}$.

Prove that $(\mathbb{Z}_{\geqslant b}, b, s_b)$ is a notion of natural numbers, and find an explicit description of the unique bijection $f : \mathbb{N}_{\mathsf{vN}} \to \mathbb{Z}$ described in the statement of Theorem B.2.8.

Merely constructing the *set* of natural numbers is not enough: we also need to provide constructions of every concept that relates to natural numbers, such as addition, multiplication, exponentiation, combinatorial coefficients, inequalities, and so on.

Fortunately for us, we already did much of the hard work in Section 4.1, where we provided recursive definitions of addition, multiplication, exponentiation, binomial coefficients and factorials.

It does not take much additional work to define other such concepts. For example, consider the '$\leqslant$' relation on $\mathbb{N}$. We could define '$m \leqslant n$' to mean $\exists k \in \mathbb{N}, m + k = n$, but this is slightly problematic because it is sensitive to the notion of natural numbers that we are considering—it gives the correct notion when we take $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ (see Exercise B.2.11) but not when we take $\mathbb{N} = \{1, 2, 3, 4, \dots\}$. We can solve this problem by defining '$m \leqslant n$' for $m, n \in \mathbb{N}$ using an iterated recursion on m and n—specifically, we declare the following:

- $0 \leqslant 0$ is true;
- For all $m \in \mathbb{N}$, $m + 1 \leqslant 0$ is false;
- For all $n \in \mathbb{N}$, $0 \leqslant n + 1$ is true; and
- For all $m, n \in \mathbb{N}$, $m + 1 \leqslant n + 1$ is true if and only if $m \leqslant n$ is true.

✎ **Exercise B.2.11**
Prove that the recursive definition of $\leqslant$ is equivalent to the definition of '$m \leqslant n$' as '$\exists k \in \mathbb{N}, m + k = n$', provided we use the convention that $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.

Integers

Now that we have constructed the natural numbers, it is time to construct the integers. The intuition behind the definition we are about to give is that an integer tells us the difference between two natural numbers. Thus '3 as an integer' refers to increase of 3 from one natural number to another (say 0 to 3, or 7 to 10); and '-4 as an integer' refers to a decrease of 4 from one natural number to another (say 4 to 0, or 10 to 6).

A first attempt might be to define $\mathbb{Z}$ to be $\mathbb{N}\times\mathbb{N}$, and interpret $(a,b)\in\mathbb{N}\times\mathbb{N}$ to represent the integer $b-a$. Unfortunately this doesn't quite work, since for instance the integer 3 would be represented as $(0,3)$ and as $(7,10)$—and, more generally, as $(n,n+3)$ for all $n\in\mathbb{N}$.

So instead, we will declare two pairs (a,b), $(c,d)\in\mathbb{N}\times\mathbb{N}$ to represent the same integer whenever $b-a=d-c$. To make this formal, a couple of things need to happen:

(1) We need to make precise what we mean by 'declare two pairs to represent the same integer'—this will be done by defining an equivalence relation on $\mathbb{N}\times\mathbb{N}$ and then passing to the quotient (see Section 5.2).

(2) We cannot use subtraction in our definition of the equivalence relation, or in our proof that it *is* an equivalence relation, since we have not yet defined a notion of subtraction for $\mathbb{N}$.

Fortunately for us, (2) is easy to resolve: we can state the equation $b-a=d-c$ for $m,n\in\mathbb{N}$ equivalently as $a+d=b+c$; this is the equation that we will use to define the equivalence relation on $\mathbb{N}\times\mathbb{N}$ in (1).

Lemma B.2.12

The relation $\sim$ on $\mathbb{N}\times\mathbb{N}$, defined for $(a,b),(c,d)\in\mathbb{N}\times\mathbb{N}$ by $(a,b)\sim(c,d)$ if and only if $a+d=b+c$, is an equivalence relation.

Proof

- (**Reflexivity**) Let $(a,b)\in\mathbb{N}\times\mathbb{N}$. Then $a+b=b+a$, so that $(a,b)\sim(a,b)$.
- (**Symmetry**) Let $(a,b),(c,d)\in\mathbb{N}\times\mathbb{N}$ and assume that $(a,b)\sim(c,d)$. Then $a+d=b+c$, so $c+b=d+a$, and so $(c,d)\sim(a,b)$.
- (**Transitivity**) Let $(a,b),(c,d),(e,f)\in\mathbb{N}\times\mathbb{N}$ and assume that $(a,b)\sim(c,d)$ and $(c,d)\sim(e,f)$. Then $a+d=b+c$ and $c+f=d+e$. But then

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

so cancelling $c+d$ from both sides gives $a+f=b+e$, and so $(a,b)\sim(e,f)$, as required.

We should be careful to note that we did not assume the existence of a subtraction operation when we cancelled $c+d$ from both sides of the equation: the fact that $u+w=v+w$ implies $u=v$ for all $u,v,w\in\mathbb{N}$ can be proved by induction on w using only Peano's axioms.

Now that we have an equivalence relation on $\mathbb{N}\times\mathbb{N}$, we can take $\mathbb{Z}$ to be the resulting set of equivalence classes (that is, the quotient).

Construction B.2.13

The **set of integers** is the set $\mathbb{Z}$ defined by

$$\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim$$

where $\sim$ is the equivalence relation on $\mathbb{N} \times \mathbb{N}$ defined by

$$(a,b) \sim (c,d) \text{ if and only if } a+d = b+c$$

for all $(a,b),(c,d) \in \mathbb{N} \times \mathbb{N}$.

We interpret the element $[(a,b)]_\sim \in \mathbb{Z}$ to be the integer $b-a$. So for example we have

$$3 = [(0,3)]_\sim = [(7,10)]_\sim \quad \text{and} \quad -4 = [(4,0)]_\sim = [(10,6)]_\sim$$

A couple of remarks are now in order.

The first remark is that we can define addition and multiplication operations on $\mathbb{Z}$ using those on $\mathbb{N}$. Formally, we define:

- $[(a,b)]_\sim + [(c,d)]_\sim = [(a+c,b+d)]_\sim$—the intuition here is that for all $a,b,c,d \in \mathbb{N}$ we have $(b-a)+(d-c) = (b+d)-(a+c)$.
- $[(a,b)]_\sim \cdot [(c,d)]_\sim = [(ad+bc,ac+bd)]_\sim$—the intuition here is that for all $a,b,c,d \in \mathbb{N}$ we have $(b-a)(d-c) = (ac+bd)-(ad+bc)$.

Since the operations $+$ and $\cdot$ are defined in terms of representatives of equivalence classes, we should (but won't) check that these operations are well-defined—for example, we need to verify that if $[(a,b)]_\sim = [(a',b')]_\sim$ and $[(c,d)]_\sim = [(c',d')]_\sim$, then $[(a+c,b+d)]_\sim = [(a'+c',b'+d')]_\sim$.

What makes the integers special is that we can also negate them, which in turn allows us to subtract them. With this in mind, we may define negation and subtraction operations on $\mathbb{Z}$ as follows:

- $-[(a,b)]_\sim = [(b,a)]_\sim$—the intuition here is that for all $a,b \in \mathbb{N}$ we have $-(b-a) = a-b$.
- $[(a,b)]_\sim - [(c,d)]_\sim = [(a+d,b+c)]_\sim$—the intuition here is that for all $a,b,c,d \in \mathbb{N}$ we have $(b-a)-(d-c) = (b+c)-(a+d)$.

The negation and subtraction operations may be defined in terms of one another: they are related by the identities

$$-[(a,b)]_\sim = [(0,0)]_\sim - [(a,b)]_\sim \quad \text{and} \quad [(a,b)]_\sim - [(c,d)]_\sim = [(a,b)]_\sim + (-[(c,d)]_\sim)$$

Again, we should but will not prove that the negation and subtraction operations are well-defined.

The second and perhaps most alarming remark is that this construction of $\mathbb{Z}$ means that we do not have $\mathbb{N} \subseteq \mathbb{Z}$: the elements of $\mathbb{Z}$ are equivalence classes of pairs of natural numbers, meaning that the natural number '3' is not equal to the integer '3' (which is really equal to $[(0,3)]_\sim$). However, we will regard $\mathbb{N}$ as being a subset of $\mathbb{Z}$ by identifying each natural number n with the element $[(0,n)]_\sim \in \mathbb{Z}$. What justifies this identification is the following exercise.

Exercise B.2.14

Prove that every element of $\mathbb{Z}$, as defined in Construction B.2.13, is equal to exactly one of the following: $[(0,0)]_\sim$, or $[(0,n)]_\sim$ for some $n > 0$, or $[(n,0)]_\sim$ for some $n > 0$.

Exercise B.2.14 implies that the 'inclusion' function $i : \mathbb{N} \to \mathbb{Z}$ defined by $i(n) = [(0,n)]_\sim$ for all $n \in \mathbb{N}$ is injective. When $n \in \mathbb{N}$, we will usually just write 'n' to denote the integer $i(n) = [(0,n)]_\sim$.

Note that $i(m+n) = i(m) + i(n)$ and $i(mn) = i(m)i(n)$, so when we write something like '$m+n$' for $m,n \in \mathbb{N}$, it doesn't matter whether we're first adding m and n as natural numbers and then interpreting the result as an integer, or first interpreting m and n as integers and then adding the result using the addition operation for integers.

Note also that the identification of $n \in \mathbb{N}$ with $i(n) = [(0,n)]_\sim \in \mathbb{Z}$ gives $-n = [(n,0)]_\sim$ for all $n \in \mathbb{N}$, and so Exercise B.2.14 implies that every integer is equal to exactly one of 0, n or $-n$ for some positive natural number n.

The next result proves that $\mathbb{Z}$ is a *ring*—what this means in essence is that the operations of addition, multiplication and negation satisfy the basic properties that we take for granted when doing arithmetic with integers.

Theorem B.2.15 · *$\mathbb{Z}$ is a ring*

(a) (**Associativity of addition**) $(a+b)+c = a+(b+c)$ for all $a,b,c \in \mathbb{Z}$.

(b) (**Commutativity of addition**) $a+b = b+a$ for all $a,b \in \mathbb{Z}$.

(c) (**Additive identity**) $a+0 = a$ for all $a \in \mathbb{Z}$.

(d) (**Negation**) $a+(-a) = 0$ for all $a \in \mathbb{Z}$.

(e) (**Associativity of multiplication**) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a,b,c \in \mathbb{Z}$.

(f) (**Commutativity of multiplication**) $a \cdot b = b \cdot a$ for all $a,b \in \mathbb{Z}$.

(g) (**Multiplicative identity**) $a \cdot 1 = a$ for all $a \in \mathbb{Z}$.

(h) (**Distributivity**) $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ for all $a,b,c \in \mathbb{Z}$.

Proof

Many pages would be wasted by writing out the proof of this theorem in full detail. Instead we will just prove part (h); any particularly masochistic readers are invited to prove parts (a)–(g) on their own.

To see that the distributivity law holds, let $a,b,c \in \mathbb{Z}$, and let $m,n,p,q,r,s \in \mathbb{N}$ be such that $a = [(m,n)]_\sim$, $b = [(p,q)]_\sim$ and $c = [(r,s)]_\sim$. (We will omit the subscript $\sim$ in what follows.)

Then:

$$\begin{aligned}
&a\cdot(b+c)\\
&=[(m,n)]\cdot([(p,q)]+[(r,s)])\\
&=[(m,n)]\cdot[(p+r,q+s)]\\
&=[(m(q+s)+n(p+r),m(p+r)+n(q+s))]\\
&=[(mq+ms+np+nr,mp+mr+nq+ns)]\\
&=[(mq+np,mp+nq)]+[(ms+nr,mr+ns)]\\
&=[(m,n)]\cdot[(p,q)]+[(m,n)]\cdot[(r,s)]\\
&=(a\cdot b)+(a\cdot c)
\end{aligned}$$

as required. □

Rational numbers

Just as we constructed the integers from the natural numbers by formalising what we mean by the 'difference' of two natural numbers, we will construct the rational numbers from the integers by formalising what we mean by the 'ratio' of two integers.

Since rational numbers should take the form $\frac{a}{b}$ with $a,b\in\mathbb{Z}$ and $b\neq 0$, a first attempt at constructing $\mathbb{Q}$ might be to take $\mathbb{Q}=\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$, and letting the pair (a,b) represent the fraction $\frac{a}{b}$. But just as an integer cannot be expressed uniquely as the difference of two natural numbers, a rational number cannot be expressed uniquely as the ratio of two integers—for example, we have $\frac{1}{2}=\frac{2}{4}$.

This means that we must identify $(a,b)\in\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$ with $(c,d)\in\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$ whenever $\frac{a}{b}=\frac{c}{d}$. But wait! We have not yet defined a division operation, so we cannot use it in our construction of $\mathbb{Q}$—so instead, we will identify (a,b) with (c,d) whenever $ad=bc$, noting that this will be equivalent to $\frac{a}{b}=\frac{c}{d}$ once we have defined a division operation.

Lemma B.2.16
The relation $\sim$ on $\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$, defined for $(a,b),(c,d)\in\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$ by letting $(a,b)\sim(c,d)$ if and only if $ad=bc$, is an equivalence relation.

Sketch of proof
The proof is essentially the same as that of Lemma B.2.12, but with addition replaced by multiplication; the proof of transitivity uses fact that for $u,v,w\in\mathbb{Z}$ with $u\neq 0$, we have $uv=uw\Rightarrow v=w$, which can be proved by induction on $|u|\geqslant 1$ without using a division operation. □

Construction B.2.17
The **set of rational numbers** is the set $\mathbb{Q}$ defined by

$$\mathbb{Q}=(\mathbb{Z}\times(\mathbb{Z}\setminus\{0\}))/\sim$$

where $\sim$ is the equivalence relation on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ defined by

$$(a,b) \sim (c,d) \text{ if and only if } ad = bc$$

for all $(a,b),(c,d) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$.

We will write '$\frac{a}{b}$' to denote the element $[(a,b)]_\sim \in \mathbb{Q}$, noting that $\frac{a}{b} = \frac{c}{d}$ if and only if $ad = bc$. We can now define addition, multiplication, negation and subtraction operations on $\mathbb{Q}$ in terms of those on $\mathbb{Z}$—namely, given $a,b,c,d \in \mathbb{Z}$ with $b,d \neq 0$, we define

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \quad -\frac{a}{b} = \frac{-a}{b}, \quad \frac{a}{b} - \frac{c}{d} = \frac{ad-bc}{bd}$$

As always, we should (but won't) verify that these operations are well-defined.

Again, we do not have $\mathbb{Z} \subseteq \mathbb{Q}$, but we can embed $\mathbb{Z}$ into $\mathbb{Q}$ in a way that respects the arithmetic operations of addition, multiplication, negation and subtraction.

Exercise B.2.18
Prove that the function $i : \mathbb{Z} \to \mathbb{Q}$ defined by $i(a) = \frac{a}{1}$ is an injection, and for all $a,b \in \mathbb{Z}$ we have $i(a+b) = i(a)+i(b)$, $i(ab) = i(a)i(b)$, $i(-a) = -i(a)$ and $i(a-b) = i(a)-i(b)$.

In light of this, given $a \in \mathbb{Z}$, we will typically just write 'a' for the rational number $i(a) = \frac{a}{1} \in \mathbb{Q}$, so we may pretend that $\mathbb{Z} \subseteq \mathbb{Q}$ even though this is not strictly the case.

The construction of $\mathbb{Q}$ also allows for reciprocal and division operations to be defined. Note that $\frac{c}{d} = 0$ if and only if $c = c \cdot 1 = 0 \cdot d = 0$; so if $a,b,c,d \in \mathbb{Q}$ with $b,c,d \neq 0$, then we define

$$\left(\frac{c}{d}\right)^{-1} = \frac{d}{c} \quad \textbf{and} \quad \frac{a}{b} \div \frac{c}{d} = \frac{ad}{bc}$$

The usual comment about well-definedness applies. Note that for all $a,b \in \mathbb{Z}$ with $b \neq 0$ we have

$$a \div b = \frac{a}{1} \div \frac{b}{1} = \frac{a \cdot 1}{1 \cdot b} = \frac{a}{b}$$

So we will dispense with the '$\div$' symbol and simply use the fraction notation.

Note also that the reciprocal and division operations may be defined in terms of one another: they are related by the identities

$$y^{-1} = \frac{1}{y} \quad \textbf{and} \quad \frac{x}{y} = x \cdot y^{-1}$$

for all $x,y \in \mathbb{Q}$ with $y \neq 0$.

The next result proves that $\mathbb{Q}$ is a *field*, which is like a ring (see Theorem B.2.15), meaning that rational numbers can be added, subtracted, multiplied and divided.

Theorem B.2.19 · *$\mathbb{Q}$ is a field*

(a) (**Associativity of addition**) $(a+b)+c = a+(b+c)$ for all $a,b,c \in \mathbb{Q}$.

(b) (**Commutativity of addition**) $a+b = b+a$ for all $a,b \in \mathbb{Q}$.

(c) (**Additive identity**) $a+0 = a$ for all $a \in \mathbb{Q}$.

(d) (**Negation**) $a+(-a) = 0$ for all $a \in \mathbb{Q}$.

(e) (**Associativity of multiplication**) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a,b,c \in \mathbb{Q}$.

(f) (**Commutativity of multiplication**) $a \cdot b = b \cdot a$ for all $a,b \in \mathbb{Q}$.

(g) (**Multiplicative identity**) $a \cdot 1 = a$ for all $a \in \mathbb{Q}$.

(h) (**Reciprocal**) $a \cdot a^{-1} = 1$ for all $a \in \mathbb{Q}$ with $a \neq 0$.

(i) (**Distributivity**) $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ for all $a,b,c \in \mathbb{Q}$.

(j) (**Nontriviality**) $0 \neq 1$.

Proof

As with Theorem B.2.15, we will only prove the distributivity property (i); the reader is invited to prove the other parts.

So let $a,b,c \in \mathbb{Q}$, and let $m,n,p,q,r,s \in \mathbb{Z}$ with $n,q,s \neq 0$ be such that $a = \frac{m}{n}$. $b = \frac{p}{q}$ and $c = \frac{r}{s}$; remember that these fractions are actually equivalence classes, as for example $\frac{m}{n} = [(m,n)]_\sim$, as defined in Construction B.2.17). Then using the definitions of addition and multiplication defined above, and making liberal use of (a)–(h), we have:

$$\begin{aligned}
a \cdot (b+c) &= \frac{m}{n} \cdot \left(\frac{p}{q} + \frac{r}{s} \right) \\
&= \frac{m}{n} \cdot \frac{ps+qr}{qs} \\
&= \frac{mpq+mqr}{nqs} \\
&= \frac{mpq+mqr}{nqs} \cdot 1 \\
&= \frac{mpq+mqr}{nqs} \cdot \frac{n}{n} \\
&= \frac{(mp)(ns)+(nq)(mr)}{(nq)(ns)} \\
&= \frac{mp}{nq} + \frac{mr}{ns} \\
&= \frac{m}{n} \cdot \frac{p}{q} + \frac{m}{n} \cdot \frac{r}{s} \\
&= ab+ac
\end{aligned}$$

as required. □

Axiomatising the real numbers

Above we constructed the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$, defined arithemtic operations and proved that they satisfy the expected properties. In the case of the real numbers, we will reverse the direction: we will build up a set of properties that real numbers and their associated arithmetic operations should satisfy, and we will discuss how to construct the set of real numbers from $\mathbb{Q}$ afterwards.

Axioms B.2.20 · *Field axioms*

Let X be a set equipped with elements 0 ('zero') and 1 ('unit'), and binary operations $+$ ('addition') and $\cdot$ ('multiplication'). The structure $(X, 0, 1, +, \cdot)$ is a **field** if it satisfies the following axioms:

- **Zero and unit**
  - (F1) $0 \neq 1$.
- **Axioms for addition**
  - (F2) (Associativity) $x + (y + z) = (x + y) + z$ for all $x, y, z \in X$.
  - (F3) (Identity) $x + 0 = x$ for all $x \in X$.
  - (F4) (Inverse) For all $x \in X$, there exists $y \in X$ such that $x + y = 0$.
  - (F5) (Commutativity) $x + y = y + x$ for all $x, y \in X$.
- **Axioms for multiplication**
  - (F6) (Associativity) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ for all $x, y, z \in X$.
  - (F7) (Identity) $x \cdot 1 = x$ for all $x \in X$.
  - (F8) (Inverse) For all $x \in X$ with $x \neq 0$, there exists $y \in X$ such that $x \cdot y = 1$.
  - (F9) (Commutativity) $x \cdot y = y \cdot x$ for all $x, y \in X$.
- **Distributivity**
  - (F10) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ for all $x, y, z \in X$.

Example B.2.21

The rationals $\mathbb{Q}$ and the reals $\mathbb{R}$ both form fields with their usual notions of zero, unit, addition and multiplication. However, the integers $\mathbb{Z}$ do not, since for example 2 has no multiplicative inverse.

Example B.2.22

Let $p > 0$ be prime. The set $\mathbb{Z}/p\mathbb{Z}$ (see Definition 5.2.18) is a field, with zero element $[0]_p$ and unit element $[1]_p$, and with addition and multiplication defined by

$$[a]_p + [b]_p = [a + b]_p \quad \text{and} \quad [a]_p \cdot [b]_p = [ab]_p$$

for all $a, b \in \mathbb{Z}$. Well-definedness of these operations is immediate from Theorem 5.2.11 and the modular arithmetic theorem (Theorem 7.3.3).

The only axiom which is not easy to verify is the multiplicative inverse axiom (F8). Indeed, if $[a]_p \in \mathbb{Z}/p\mathbb{Z}$ then $[a]_p \neq [0]_p$ if and only if $p \nmid a$. But if $p \nmid a$ then $a \perp p$, so a has a multiplicative inverse u modulo p. This implies that $[a]_p \cdot [u]_p = [au]_p = [1]_p$. So (F8) holds.

Exercise B.2.23
Let $n > 0$ be composite. Prove that $\mathbb{Z}/n\mathbb{Z}$ is not a field, where zero, unit, addition and multiplication are defined as in Example B.2.22.

Axiom B.2.20 tell us that every element of a field has an additive inverse, and every *nonzero* element of a field has a multiplicative inverse. It would be convenient if inverses were *unique* whenever they exist. Proposition B.2.24 proves that this is the case.

Proposition B.2.24 · *Uniqueness of inverses*
Let $(X, 0, 1, +, \cdot)$ be a field and let $x \in X$. Then

(a) Suppose $y, z \in X$ are such that $x + y = 0$ and $x + z = 0$. Then $y = z$.

(b) Suppose $x \neq 0$ and $y, z \in X$ are such that $x \cdot y = 1$ and $x \cdot z = 1$. Then $y = z$.

Proof of (a)
By calculation, we have

$$\begin{aligned} y &= y + 0 && \text{by (F3)} \\ &= y + (x + z) && \text{by definition of } z \\ &= (y + x) + z && \text{by associativity (F2)} \\ &= (x + y) + z && \text{by commutativity (F5)} \\ &= 0 + z && \text{by definition of } y \\ &= z + 0 && \text{by commutativity (F5)} \\ &= z && \text{by (F3)} \end{aligned}$$

so indeed $y = z$.

The proof of (b) is essentially the same and is left as an exercise. □

Since inverses are unique, it makes sense to have notation to refer to them.

Notation B.2.25
Let $(X, 0, 1, +, \cdot)$ be a field and let $x \in X$. Write $-x$ for the (unique) additive inverse of x and, if $x \neq 0$ write x^{-1} for the (unique) multiplicative inverse of x.

Example B.2.26
In the fields $\mathbb{Q}$ and $\mathbb{R}$, the additive inverse $-x$ of an element x is simply its negative, and the multiplicative inverse x^{-1} of some $x \neq 0$ is simply its reciprocal $\frac{1}{x}$.

Example B.2.27
Let $p > 0$ be prime and let $[a]_p \in \mathbb{Z}/p\mathbb{Z}$. Then $-[a]_p = [-a]_p$ and, if $p \nmid a$, then $[a]_p^{-1} = [u]_p$, where u is any integer satisfying $au \equiv 1 \bmod p$.

Exercise B.2.28

Let $(X,0,1,+,\cdot)$ be a field. Prove that $-(-x)=x$ for all $x\in X$, and that $(x^{-1})^{-1}=x$ for all nonzero $x\in X$.

Example B.2.29

Let $(X,0,1,+,\cdot)$ be a field. We prove that if $x\in X$ then $x\cdot 0=0$. Well, $0=0+0$ by (F3). Hence $x\cdot 0=x\cdot(0+0)$. By distributivity (F10), we have $x\cdot(0+0)=(x\cdot 0)+(x\cdot 0)$. Hence

$$x\cdot 0=(x\cdot 0)+(x\cdot 0)$$

Let $y=-(x\cdot 0)$. Then

$$\begin{aligned}
0&=x\cdot 0+y && \text{by (F4)}\\
&=((x\cdot 0)+(x\cdot 0))+y && \text{as above}\\
&=(x\cdot 0)+((x\cdot 0)+y) && \text{by associativity (F2)}\\
&=(x\cdot 0)+0 && \text{by (F4)}\\
&=x\cdot 0 && \text{by (F3)}
\end{aligned}$$

so indeed we have $x\cdot 0=0$.

Exercise B.2.30

Let $(X,0,1,+,\cdot)$ be a field. Prove that $(-1)\cdot x=-x$ for all $x\in X$, and that $(-x)^{-1}=-(x^{-1})$ for all nonzero $x\in X$.

What makes the real numbers useful is not simply our ability to add, subtract, multiply and divide them; we can also compare their size—indeed, this is what gives rise to the informal notion of a *number line*. Axiom B.2.31 make precise exactly what it means for the elements of a field to be assembled into a 'number line'.

Axioms B.2.31 · *Ordered field axioms*

Let X be a set, $0,1\in X$ be elements, $+,\cdot$ be binary operations, and $\leqslant$ be a relation on X. The structure $(X,0,1,+,\cdot,\leqslant)$ is an **ordered field** if it satisfies the field axioms (F1)–(F10) (see Axiom B.2.20) and, additionally, it satisfies the following axioms:

- **Linear order axioms**

  (PO1) (Reflexivity) $x\leqslant x$ for all $x\in X$.

  (PO2) (Antisymmetry) For all $x,y\in X$, if $x\leqslant y$ and $y\leqslant x$, then $x=y$.

  (PO3) (Transitivity) For all $x,y,z\in X$, if $x\leqslant y$ and $y\leqslant z$, then $x\leqslant z$.

  (PO4) (Linearity) For all $x,y\in X$, either $x\leqslant y$ or $y\leqslant x$.

- **Interaction of order with arithmetic**

  (OF1) For all $x,y,z\in X$, if $x\leqslant y$, then $x+z\leqslant y+z$.

  (OF2) For all $x,y\in X$, if $0\leqslant x$ and $0\leqslant y$, then $0\leqslant xy$.

Example B.2.32

The field $\mathbb{Q}$ of rational numbers and and the field $\mathbb{R}$ of real numbers, with their usual notions of ordering, can easily be seen to form ordered fields.

Example B.2.33

We prove that, in any ordered field, we have $0 \leqslant 1$. Note first that either $0 \leqslant 1$ or $1 \leqslant 0$ by linearity (PO4). If $0 \leqslant 1$ then we're done, so suppose $1 \leqslant 0$. Then $0 \leqslant -1$; indeed:

$$\begin{aligned} 0 &= 1+(-1) && \text{by (F4)} \\ &\leqslant 0+(-1) && \text{by (OF1), since } 1 \leqslant 0 \\ &= (-1)+0 && \text{by commutativity (F5)} \\ &= -1 && \text{by (F3)} \end{aligned}$$

By (OF2), it follows that $0 \leqslant (-1)(-1)$. But $(-1)(-1) = 1$ by Exercise B.2.30, and hence $0 \leqslant 1$. Since $1 \leqslant 0$ and $0 \leqslant 1$, we have $0 = 1$ by antisymmetry (PO2). But this contradicts axiom (F1). Hence $0 \leqslant 1$. In fact, $0 < 1$ since $0 \neq 1$.

We have seen that $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields (Examples B.2.26 and B.2.32), and that $\mathbb{Z}/p\mathbb{Z}$ is a field for $p > 0$ prime (Example B.2.22). The following proposition is an interesting result proving that there is no notion of 'ordering' under which the field $\mathbb{Z}/p\mathbb{Z}$ can be made into an ordered field!

Proposition B.2.34

Let $p > 0$ be prime. There is no relation $\leqslant$ on $\mathbb{Z}/p\mathbb{Z}$ which satisfies the ordered field axioms.

Proof

We just showed that $[0] \leqslant [1]$. It follows that, for all $a \in \mathbb{Z}$, we have $[a] \leqslant [a]+[1]$; indeed:

$$\begin{aligned} [a] &= [a]+[0] && \text{by (F3)} \\ &\leqslant [a]+[1] && \text{by (OF1), since } [0] \leqslant [1] \\ &= [a+1] && \text{by definition of } + \text{ on } \mathbb{Z}/p\mathbb{Z} \end{aligned}$$

It is a straightforward induction to prove that $[a] \leqslant [a+n]$ for all $n \in \mathbb{N}$. But then we have

$$[1] \leqslant [1+(p-1)] = [p] = [0]$$

so $[0] \leqslant [1]$ and $[1] \leqslant [0]$. This implies $[0] = [1]$ by antisymmetry (PO2), contradicting axiom (F1). □

Exercise B.2.35

Let $(X, 0, 1, +, \cdot)$ be a field. Prove that if X is finite, then there is no relation $\leqslant$ on X such that $(X, 0, 1, +, \cdot, \leqslant)$ is an ordered field.

Theorem B.2.36 below summarises some properties of ordered fields which are used in our proofs. Note, however, that this is certainly *not* an exhaustive list of elementary properties of ordered fields that we use—to explicitly state and prove all of these would not make for a scintillating read.

Theorem B.2.36

Let $(X, 0, 1, +, \cdot, \leqslant)$ be an ordered field. Then

(a) For all $x, y \in X$, $x \leqslant y$ if and only if $0 \leqslant y - x$;

(b) For all $x \in X$, $-x \leqslant 0 \leqslant x$ or $x \leqslant 0 \leqslant -x$;

(c) For all $x, x', y, y' \in X$, if $x \leqslant x'$ and $y \leqslant y'$, then $x + y \leqslant x' + y'$;

(d) For all $x, y, z \in X$, if $0 \leqslant x$ and $y \leqslant z$, then $xy \leqslant xz$;

(e) For all nonzero $x \in X$, if $0 \leqslant x$, then $0 \leqslant x^{-1}$.

(f) For all nonzero $x, y \in X$, if $x \leqslant y$, then $y^{-1} \leqslant x^{-1}$.

Proof of (a), (b) and (e)

(a) ($\Rightarrow$) Suppose $x \leqslant y$. Then by additivity (OF1), $x + (-x) \leqslant y + (-x)$, that is $0 \leqslant y - x$. ($\Leftarrow$) Suppose $0 \leqslant y - x$. By additivity (OF1), $0 + x \leqslant (y - x) + x$; that is, $x \leqslant y$.

(b) We know by linearity (PO4) that either $0 \leqslant x$ or $x \leqslant 0$. If $0 \leqslant x$, then by (OF1) we have $0 + (-x) \leqslant x + (-x)$, that is $-x \leqslant 0$. Likewise, if $x \leqslant 0$ then $0 \leqslant -x$.

(e) Suppose $0 \leqslant x$. By linearity (PO4), either $0 \leqslant x^{-1}$ or $x^{-1} \leqslant 0$. If $x^{-1} \leqslant 0$, then by (d) we have $x^{-1} \cdot x \leqslant 0 \cdot x$, that is $1 \leqslant 0$. This contradicts Example B.2.33, so we must have $0 \leqslant x^{-1}$.

The proofs of the remaining properties are left as an exercise. □

We wanted to characterise the reals completely, but so far we have failed to do so—indeed, Example B.2.32 showed that both $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields, so the ordered field axioms do not suffice to distinguish $\mathbb{Q}$ from $\mathbb{R}$. The final piece in the puzzle is *completeness*. This single additional axiom distinguishes $\mathbb{Q}$ from $\mathbb{R}$, and in fact completely characterises $\mathbb{R}$ (see Theorem B.2.38).

Axioms B.2.37 · *Complete ordered field axioms*

Let X be a set, $0, 1 \in X$ be elements, $+, \cdot$ be binary operations, and $\leqslant$ be a relation on X. The structure $(X, 0, 1, +, \cdot, \leqslant)$ is a **complete ordered field** if it is an ordered field—that is, it satisfies axioms (F1)–(F10), (PO1)–(PO4) and (OF1)–(OF2) (see Axioms B.2.20 and B.2.31)—and, in addition, it satisfies the following **completeness axiom**:

(C1) Let $A \subseteq X$. If A has an upper bound, then it has a least upper bound. Specifically, if there exists $u \in X$ such that $a \leqslant u$ for all $a \in A$, then there exists $s \in X$ such that

- $a \leqslant s$ for all $a \in A$; and
- If $s' \in X$ is such that $a \leqslant s'$ for all $a \in A$, then $s \leqslant s'$.

We call such a value $s \in X$ a **supremum** for A.

Theorem B.2.38

The real numbers $(\mathbb{R}, 0, 1, +, \cdot, \leqslant)$ form a complete ordered field. Moreover, any two complete ordered fields are essentially the same.

The notion of 'sameness' alluded to in Theorem B.2.38 is more properly called *isomorphism*. A proof of this theorem is intricate and far beyond the scope of this book, so is omitted. What it tells us is that it doesn't matter exactly how we define the reals, since any complete ordered field will do. We can therefore proceed with confidence that, no matter what notion of 'real numbers' we settle on, everything we prove will be true of that notion. This is for the best, since we haven't actually defined the set $\mathbb{R}$ of real numbers at all!

The two most common approaches to constructing a set of real numbers are:

- **Dedekind reals.** In this approach, real numbers are identified with particular subsets of $\mathbb{Q}$—informally speaking, $r \in \mathbb{R}$ is identified with the set of rational numbers less than r.
- **Cauchy reals.** In this approach, real numbers are identified with equivalence classes of sequences of rational numbers—informally speaking, $r \in \mathbb{R}$ is identified with the set of sequences of rational numbers which converge to r (in the sense of Definition 9.2.15).

Section B.3

Limits of functions

At the end of Section 9.1 we mentioned the use of *limits* of functions without properly defining what we meant. This admittedly brusque section is dedicated to making what we meant mathematical precise.

Limits

Definition B.3.1

Let $D \subseteq \mathbb{R}$. A **limit point** of D is a real number a such that, for all $\delta > 0$, there exists some $x \in D$ such that $0 < |x - a| < \delta$.

Lemma B.3.2

Let $D \subseteq \mathbb{R}$. A real number a is a limit point of D if and only if there is a sequence (x_n) of elements of D, which is not eventually constant, such that $(x_n) \to a$.

Proof

($\Rightarrow$) Let $a \in \mathbb{R}$ and assume that a is a limit point of D. For each $n \geqslant 1$, let x_n be some element of D such that $0 < |x_n - a| < \dfrac{1}{n}$.

Evidently $(x_n) \to a$: indeed, given $\varepsilon > 0$, letting $N \geqslant \max\{1, \frac{1}{\varepsilon}\}$ gives $|x_n - a| < \varepsilon$ for all $n \geqslant N$.

Moreover, the sequence (x_n) is not eventually constant: if it were, there would exist $N \geqslant 1$ and $b \in \mathbb{R}$ such that $x_n = b$ for all $n \geqslant N$. But then by the squeeze theorem (Theorem 9.2.38), we'd have

$$0 \leqslant \lim_{n\to\infty} |x_n - a| = |b - a| \leqslant \lim_{n\to\infty} \frac{1}{n} = 0$$

and so $b = a$. But this contradicts the fact that $|x_n - a| > 0$ for all $n \geqslant 1$.

($\Leftarrow$) Let $a \in \mathbb{R}$ and assume that there is a sequence (x_n) of elements of D, which is not eventually constant, such that $(x_n) \to a$. Then for all $\delta > 0$ there exists some $N \in \mathbb{N}$ such that $|x_n - a| < \varepsilon$ for all $n \geqslant N$. Since (x_n) is not eventually constant, there is some $n \geqslant N$ such that $|x_n - a| > 0$—otherwise (x_n) would be eventually constant with value a! But then $x_n \in D$ and $0 < |x_n - a| < \delta$, so that a is a limit point of D. □

Definition B.3.3

Let $D \subseteq \mathbb{R}$. The **closure** of D is the set $\overline{D}$ (LaTeX code: `\overline{D}`) defined by

$$\overline{D} = D \cup \{a \in \mathbb{R} \mid a \text{ is a limit point of } D\}$$

That is, $\overline{D}$ is given by D together with its limit points.

Example B.3.4

We have $\overline{(0,1)} = [0,1]$. Indeed, $(0,1) \subseteq \overline{(0,1)}$ since $D \subseteq \overline{D}$ for all $D \subseteq \mathbb{R}$. Moreover the sequences $(\frac{1}{n})$ and $(1-\frac{1}{n})$ are non-constant, take values in $(0,1)$, and converge to 0 and 1 respectively, so that $0 \in \overline{(0,1)}$ and $1 \in \overline{(0,1)}$. Hence $[0,1] \subseteq \overline{(0,1)}$.

Given $a \in \mathbb{R}$, if $a > 1$, then letting $\delta = 1 - a > 0$ reveals that $|x-a| \geqslant \delta$ for all $x \in D$; and likewise, if $a < 0$, then letting $\delta = -a > 0$ reveals that $|x-a| \geqslant \delta$ for all $x \in D$. Hence no element of $\mathbb{R} \setminus [0,1]$ is an element of D, so that $\overline{(0,1)} = [0,1]$.

Exercise B.3.5

Let $a, b \in \mathbb{R}$ with $a < b$. Prove that $\overline{(a,b)} = \overline{(a,b]} = \overline{[a,b)} = [a,b]$.

Convention B.3.6

For the rest of this section, whenever we declare $f : D \to \mathbb{R}$ to be a function, it will be assumed that the domain D is a subset of $\mathbb{R}$, and that every point of D is a limit point of D. In other words, D has no *isolated points*, which are points separated from all other elements of D by a positive distance. For instance, in the set $(0,1] \cup \{2\}$, the element $2 \in \mathbb{R}$ is an isolated point.

Definition B.3.7

Let $f : D \to \mathbb{R}$ be a function, let $a \in \overline{D}$, and let $\ell \in \mathbb{R}$. We say ℓ is a **limit** of $f(x)$ as x **approaches** a if

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in D,\ 0 < |x-a| < \delta \ \Rightarrow\ |f(x) - \ell| < \varepsilon$$

In other words, for values of $x \in D$ near a (but not equal to a), the values of $f(x)$ become arbitrarily close to ℓ.

We write '$f(x) \to \ell$ as $x \to a$' (LaTeX code: `\to`) to denote the assertion that ℓ is a limit of $f(x)$ as x approaches a.

Example B.3.8

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x$ for all $x \in \mathbb{R}$. Then $f(x) \to 0$ as $x \to 0$. To see this, let $\varepsilon > 0$, and define $\delta = \varepsilon > 0$. Then for all $x \in \mathbb{R}$, if $0 < |x-a| < \delta = \varepsilon$, then

$$|f(x) - f(a)| = |x-a| < \varepsilon$$

as required.

Exercise B.3.9

Let $f : D \to \mathbb{R}$ be a function, let $a \in D$ and let $\ell \in \mathbb{R}$. Fix some sequence (x_n) of elements of D, not eventually constant, such that $(x_n) \to a$. Prove that if $f(x) \to \ell$ as $x \to a$, then the sequence $(f(x_n))$ converges to ℓ.

The next exercise tells us that limits of functions are unique, provided that they exist. Its proof looks much like the analogous result we proved for sequences in Theorem 9.2.41.

Exercise B.3.10

Let $f : D \to \mathbb{R}$ be a function, let $a \in D$, and $\ell_1, \ell_2 \in \mathbb{R}$. Prove that if $f(x) \to \ell_1$ as $x \to a$, and $f(x) \to \ell_2$ as $x \to a$, then $\ell_1 = \ell_2$.

When the domain D of a function $f : D \to \mathbb{R}$ is unbounded, we might also be interested in finding out how the values of $f(x)$ behave as $x \in D$ gets (positively or negatively) larger and larger.

Definition B.3.11

Let $f : D \to \mathbb{R}$ be a function and let $\ell \in \mathbb{R}$. If D is unbounded above—that is, for all $p \in \mathbb{R}$, there exists $x \in D$ with $x > p$—then we say ℓ is a **limit** of $f(x)$ as x **increases without bound** if

$$\forall \varepsilon > 0,\ \exists p \in \mathbb{R},\ \forall x \in D,\ x > p \ \Rightarrow\ |f(x) - \ell| < \varepsilon$$

We write '$f(x) \to \ell$ as $x \to \infty$' (LaTeX code: `\infty`) to denote the assertion that ℓ is a limit of $f(x)$ as x increases without bound.

Likewise, if D is unbounded below—that is, for all $p \in \mathbb{R}$, there exists $x \in D$ with $x < p$—then we say ℓ is a **limit** of $f(x)$ as x **decreases without bound** if

$$\forall \varepsilon > 0,\ \exists p \in \mathbb{R},\ \forall x \in D,\ x < p \ \Rightarrow\ |f(x) - \ell| < \varepsilon$$

We write '$f(x) \to \ell$ as $x \to -\infty$' to denote the assertion that ℓ is a limit of $f(x)$ as x decreases without bound.

Example B.3.12

Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $f(x) = \dfrac{x}{1+|x|}$ for all $x \in \mathbb{R}$. Then:

- $f(x) \to 1$ as $x \to \infty$. To see this, let $\varepsilon > 0$, and define $p = \max\{1, \frac{1}{\varepsilon}\}$. Then for all $x > p$, we have $x > 0$, so that $f(x) = \dfrac{x}{1+x}$, and $x > \frac{1}{\varepsilon} - 1$. Hence:

$$\left|\frac{x}{1+x} - 1\right| = \left|\frac{-1}{1+x}\right| = \frac{1}{1+x} < \frac{1}{1+(\frac{1}{\varepsilon} - 1)} = \varepsilon$$

  as required.

- $f(x) \to -1$ as $x \to \infty$. To see this, let $\varepsilon > 0$ and define $p = \min\{-1, \frac{-1}{\varepsilon}\}$. Then for all $x < p$, we have $x < 0$, so that $f(x) = \frac{1}{1-x}$, and $x < \frac{-1}{\varepsilon} + 1$. Hence:

$$\left|\frac{x}{1-x} - (-1)\right| = \left|\frac{1}{1-x}\right| = \frac{1}{1-x} < \frac{1}{1-(-\frac{1}{\varepsilon} + 1)} = \varepsilon$$

  as required.

So $f(x) \to 1$ as $x \to \infty$ and $f(x) \to -1$ as $x \to -\infty$.

Exercise B.3.13

Let $f : D \to \mathbb{R}$ be a function and $\ell_1, \ell_2 \in \mathbb{R}$. Prove that if D is unbounded above, and if $f(x) \to \ell_1$

as $x \to \infty$ and $f(x) \to \ell_2$ as $x \to \infty$, then $\ell_1 = \ell_2$. Prove the analogous result for limits as $x \to -\infty$ in the case when D is unbounded below.

The results of Exercises B.3.10 and B.3.13 justify the following definition.

Definition B.3.14

Let $f : D \to \mathbb{R}$ and let $a \in [-\infty, \infty]$. Assuming the limits in question are well-defined and exist, we write $\lim_{x \to a} f(x)$ to denote the unique real number $\ell \in \mathbb{R}$ such that $f(x) \to \ell$ as $x \to a$.

Appendix C

Hints and solutions

Section C.1

Hints for selected exercises

Chapter 0

Hint for Exercise 0.13

Set $n = \lfloor x \rfloor$ and verify that $n+1$ satisfies the definition of 'ceiling of x', using the fact that n is the floor of x in your proof.

Hint for Exercise 0.14

Set $n = \lfloor -x \rfloor$ and show that $-n$ satisfies the definition of being the ceiling of x; use a similar argument for the other part of the exercise.

Hint for Exercise 0.17

Use the definition of rational numbers to express x and y as fractions of integers, and then do the algebra to express each of $x+y$, $x-y$, xy and $\dfrac{x}{y}$ (when $y \neq 0$) as fractions of integers.

Hint for Exercise 0.24

Putting the plain English definition of dyadic rationals in terms of variables, a dyadic rational is a rational number x that can be expressed as $x = \dfrac{a}{2^n}$ for some $a \in \mathbb{Z}$ and some $n \in \mathbb{N}$. Think about how to express this using the two variants of set-builder notation described earlier.

Hint for Exercise 0.27

Unpack the definitions of A and B carefully using the definition of set-builder notation from earlier in the chapter.

There are several different ways to express C in the desired form. One way would be to think about the properties of the elements of C as they relate to elements of $[n]$ for some $n \in \mathbb{N}$. Another way would be to think about how arithmetic operations could be applied to elements of some set of the form $[n]$ to obtain the elements of C.

Hint for Exercise 0.30

For the intervals with endpoints, be careful to note whether each endpoint is an element of the interval; use a round bracket if it is and a square bracket if it isn't.

Chapter 0 exercises

Section 1.1

Hint for Exercise 1.1.19

Mimic the proof of Proposition 1.1.18.

Hint for Exercise 1.1.32

Suppose $n = d_r \cdot 10^r + \cdots + d_1 \cdot 10 + d_0$ and let $s = d_r + \cdots + d_1 + d_0$. Start by proving that $3 \mid n - s$.

Hint for Exercise 1.1.46
Use the law of excluded middle according to whether the proposition '$\sqrt{2}^{\sqrt{2}}$ is rational' is true or false.

Section 1.2

Hint for Exercise 1.2.15
Consider the sum of $x+y$ and $x-y$.

Hint for Exercise 1.2.22
Look carefully at the definition of divisibility (Definition 0.36).

Section 1.3

Hint for Exercise 1.3.9
Note that you may need to use the law of excluded middle (Axiom 1.1.42) and the principle of explosion (Axiom 1.1.47).

Hint for Exercise 1.3.22
Express this statement as $\forall n \in \mathbb{Z}$, (n is even) $\Leftrightarrow$ (n^2 is even), and note that the negation of 'x is even' is 'x is odd'.

Hint for Exercise 1.3.32
Find a rational number all of whose representations as a ratio of two integers have an even denominator.

Hint for Exercise 1.3.38
Start by expressing $\Leftrightarrow$ in terms of $\Rightarrow$ and $\wedge$, as in Definition 1.1.28.

Chapter 1 exercises

Section 2.1

Hint for Exercise 2.1.1
Use the set-builder notation defining R to determine whether $X \in R$ is true or false when $X = R$.

Hint for Exercise 2.1.3
Compare the strategies for proving universally quantified statements and proving implications, and for proving existentially quantified statements and proving conjunctions.

Hint for Exercise 2.1.5
It may be true or false that $R \in \mathscr{U}$.

Hint for Exercise 2.1.17

Resist the temptation to draw a graph and claim that you're done: use a proof by double containment, unpacking what it means to be an element of each set using its definition.

Hint for Exercise 2.1.18
If you feel like you're not proving anything, you're probably doing it correctly.

Hint for Exercise 2.1.24
Recall from the beginning of Section 2.1 that $\forall x \in X,\, p(x)$ is equivalent to $\forall x,\, (x \in X \Rightarrow p(x))$ and $\exists x \in X,\, p(x)$ is equivalent to $\exists x,\, (x \in X \wedge p(x))$. What can be said about the truth value of $x \in E$ when E is empty?

Section 2.2

Hint for Exercise 2.2.37
You need to find a family of subsets of $\mathbb{N}$ such that (i) any two of the subsets have infinitely many elements in common, but (ii) given any natural number, you can find one of the subsets that it is *not* an element of.

Chapter 2 exercises

Hint for Question 2.E.13
The temptation is to write a long string of equations, but it is far less painful to prove this by double containment, splitting into cases where needed. An even less painful approach is to make a cunning use of truth tables.

Hint for Question 2.E.14
The hint for Question 2.E.13 applies here too.

Section 3.1

Hint for Exercise 3.1.16
What is the value of $i(z)$ if $z \in X \cap Y$?

Hint for Exercise 3.1.22
Look closely at Definition 3.1.17.

Section 3.2

Hint for Exercise 3.2.13
Recall Definition 3.1.31.

Hint for Exercise 3.2.15
If Z were a subset of Y, then we could easily define an injection $i : Z \to Y$ by $i(z) = z$ for all $z \in Z$. Are there any subsets of Y that are associated with a function whose codomain is Y?

Hint for Exercise 3.2.16
Note that $B \in \mathscr{P}(X)$. Prove that there does not exist $a \in X$ such that $B = f(a)$.

Hint for Exercise 3.2.19
Verify directly that id_X is injective and surjective.

Hint for Exercise 3.2.28
This can be proved in a single sentence; if you find yourself writing a long proof, then there is an easier way.

Hint for Exercise 3.2.31
The proof is almost identical to Exercise 3.2.28.

Hint for Exercise 3.2.39
For part (c), don't try to write a formula for the inverse of h; instead, use the fundamental theorem of arithmetic.

Hint for Exercise 3.2.45
Use Exercise 3.2.40.

Hint for Exercise 3.2.47
Define $h : X \times Y \to A \times B$ by $h(x,y) = (f(x), g(y))$ for all $x \in X$ and all $y \in Y$; find an inverse for h in terms of the inverses of f and g.

Chapter 3 exercises

Hint for Question 3.E.2
The function $X \to \mathscr{P}(X)$ must send each $a \in X$ to the (unique) $U \in \mathscr{P}(X)$ satisfying $a \in U \wedge \exists! x \in X, x \in U$. Think about what the latter logical formula tells you about what elements U has, and doesn't have. This will help with the proof of existence and uniqueness.

Hint for Question 3.E.3
Given a function $f : X \to \varnothing$, we must have $f(a) \in \varnothing$ for each $a \in X$.

Hint for Question 3.E.8
Consider $f(x) + f(-x)$ and $f(x) - f(-x)$ for $x \in \mathbb{R}$.

Hint for Question 3.E.9
Fix $n \in \mathbb{N}$ and consider how θ_n interacts with functions $f : [1] \to [n]$.

Hint for Question 3.E.15
Begin by observing that each $h \in f[\mathbb{R}^{\mathbb{R}}]$ is an even function, in the sense of Definition 3.E.4.

Hint for Question 3.E.20
This problem is very fiddly. First prove that conditions (i)–(iii) are satisfied when $A = f[X]$ and p and i are chosen appropriately. Then condition (iii) in each case (for A and for $f[X]$) defines functions $v : A \to f[X]$ and $w : f[X] \to A$, and gives uniqueness of v. You can prove that these functions are mutually inverse using the 'uniqueness' part of condition (iii).

Hint for Question 3.E.23

Prove it by double containment, and carefully unpack the definitions of what it means to be an element of $p[S]$ and what it means to be an element of $\mathrm{Gr}(g \circ f)$.

Hint for Question 3.E.24

Avoid the temptation to prove either part of this question by contradiction. For (a), a short proof is available directly from the definitions of 'injection' and 'surjection'. For (b), find as simple a counterexample as you can.

Hint for Question 3.E.27

Start by proving that $\binom{m}{2} < \binom{m+1}{2}$ for all $m \geqslant 1$. Deduce that, for all $n \in \mathbb{N}$, there is a unique natural number k such that $\binom{k+1}{2} \leqslant n < \binom{k+2}{2}$. Can you see what this has to do with the function f?

Hint for Question 3.E.28

Consider the set of fixed points of e—that is, elements $x \in X$ such that $e(x) = x$.

Hint for Question 3.E.47

First prove that f is bijective if and only if $ad \neq bc$.

Section 4.1

Section 4.2

Hint for Exercise 4.2.12

To define the bijection, think about what the elements of the two sets look like: The elements of $\prod_{k \in [n+1]} X_k$ look like $(a_0, a_1, \ldots, a_{n-1}, a_n)$, where $a_k \in X_k$ for each $k \in [n+1]$. On the other hand, the elements of $\left(\prod_{k \in [n]} X_k\right) \times X_n$ look like $((a_0, a_1, \ldots, a_{n-1}), a_n)$.

Hint for Exercise 4.2.19

Note that $(-1)^i \binom{n}{i} = \binom{n}{i} \cdot (-1)^i \cdot 1^{n-i}$.

Section 4.3

Hint for Exercise 4.3.10

Observe that if n dubloons can be obtained using only 3 and 5 dubloon coins, then so can $n+3$. See how you might use this fact to exploit strong induction with multiple base cases.

Hint for Exercise 4.3.15

Prove first that if $a \in \mathbb{Z}$ and a^2 is divisible by 3, then a is divisible by 3.

Hint for Exercise 4.3.22

Start by proving that if $\sqrt{n} = \dfrac{a}{b}$ for some $a, b \in \mathbb{Z}$ with $b \neq 0$, then $a(\sqrt{n} - \lfloor\sqrt{n}\rfloor) \in \mathbb{Z}$ and $b(\sqrt{n} - \lfloor\sqrt{n}\rfloor) \in \mathbb{Z}$. Then think about what this has to do with infinite descent.

Chapter 4 exercises

Hint for Question 4.E.6

For example, if there are exactly 3 elements of X making $p(x)$ true, then that means that there is some $a \in X$ such that $p(a)$ is true, and there are exactly two elements $x \in X$ *other than* a making $p(x)$ true.

Hint for Question 4.E.8

The number of trailing zeros in the base-b expansion of a natural number n is the greatest natural number r such that b^r divides n. How many times does 10 go into 41!? How many times does 2 go into 41!?

Hint for Question 4.E.11

Proving this by induction on x only demonstrates that it is true for *integers* $x \geqslant -1$, not *real numbers* $x \geqslant -1$. Try proving a more general fact by induction on a different variable, and deducing this as a special case. (Do you *really* think there is anything special about the natural number 123 456 789?)

Hint for Question 4.E.14

To find the value of a in part (a), try substituting some small values of x and y into the equation $f(x+y) = f(x) + f(y)$. For part (b), use the fact that $n + (-n) = 0$ for all $n \in \mathbb{N}$. For part (c), consider the number $bf(\frac{a}{b})$ when $a, b \in \mathbb{Z}$ and $b \neq 0$.

Hint for Question 4.E.15

Like in Question 4.E.14, prove this first for $x \in \mathbb{N}$ (by induction), then for $x \in \mathbb{Z}$, and finally for $x \in \mathbb{Q}$.

Hint for Question 4.E.17

Observe that $7^{n+1} - 2 \cdot 4^{n+1} + 1 = 4(7^n - 2 \cdot 4^n + 1) + 3(7^n - 1)$ for all $n \in \mathbb{N}$. You may need to do a separate proof by induction in your induction step.

Hint for Question 4.E.22

You will need to use the angle addition formulae

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta \quad \text{and} \quad \cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$$

Hint for Question 4.E.23

For (a) use integration by parts; you do not need induction. For parts (b) and (c), use part (a).

Section 5.1

Section 5.2

Hint for Exercise 5.2.36

Given a partition $\mathscr{U}$ of a set X, find a surjection $q : X \to \mathscr{U}$. Then prove that, for every surjection

$p : X \to A$, there is a unique partition $\mathscr{U}_p$ of X and a unique bijection $f : \mathscr{U}_p \to A$ such that, for all $U \in \mathscr{U}_p$, we have $p(x) = f(U)$ for all $x \in U$. The structure of the proof will be similar to that of Theorem 5.2.35.

Chapter 5 exercises

Section 6.1

Hint for Exercise 6.1.7
Part (b) has a proof by induction that looks much like the one in part (a). In the induction step, given a surjection $g : [m+1] \to [n]$, observe that we must have $n \geqslant 1$, and construct a surjection $g^- : [m+1] \setminus \{a\} \to [n-1]$ for some suitable $a \in [m+1]$. Then invoke Lemma 6.1.5, now using the fact that $[m] = [m+1] \setminus \{m+1\}$.

Hint for Exercise 6.1.16
Given $A \subseteq X$, find an injection $A \to X$ and apply Theorem 6.1.13(a).

Hint for Exercise 6.1.17
Recall that $X \cap Y \subseteq X$ for all sets X and Y.

Hint for Exercise 6.1.20
Apply Proposition 6.1.18 with $Y = U$.

Hint for Exercise 6.1.21
Prove by induction on $n \in \mathbb{N}$ that, for all $m, n \in \mathbb{N}$, there is a bijection $[mn] \to [m] \times [n]$.

Section 6.2

Hint for Exercise 6.2.6
The fundamental theorem of arithmetic (Theorem 7.2.12) implies that, for all $n \in \mathbb{N}$, we can express $n+1$ uniquely as a power of 2 multiplied by an odd number.

Hint for Exercise 6.2.8
Use Exercise 3.2.47, together with the definition of countably infinite sets, to construct a bijection $\mathbb{N} \times \mathbb{N} \to X \times Y$. Then apply Exercise 6.2.6 and Proposition 6.2.7.

Hint for Exercise 6.2.11
Intuitively, we should be able to define f by listing the values of g in increasing order and, for all $n \in \mathbb{N}$, defining $f(n)$ to be the element $x \in X$ for which $g(x)$ is the n^{th} (counting from 0) value in the list. Your task is to make this precise. Be sure to note where in your proof you use injectivity of g.

Hint for Exercise 6.2.12
For (ii)⇒(i), use the fact that the composite of two injections is injective. Likewise for (iii)⇒(i).

Hint for Exercise 6.2.14
Suppose $X = \mathbb{N}$. By Proposition 6.2.9, the set $\mathbb{N}^k$ is countable. By Theorem 6.2.10(c), it suffices to find an injection $\binom{\mathbb{N}}{k} \to \mathbb{N}^k$.

Hint for Exercise 6.2.25
We have already proved this when X is finite. When X is countably infinite, use Theorem 6.2.17. When X is uncountably infinite, find an injection $X \to \mathscr{P}(X)$ and find a way to apply Exercise 6.2.12.

Hint for Exercise 6.2.28
How many '÷' symbols can a string from Σ^* have if the string is to represent a rational number? Where in a word over Σ can a ÷ symbol appear?

Hint for Exercise 6.2.29
Prove by induction that Σ^n is countable for all $n \in \mathbb{N}$, and then apply Theorem 6.2.15.

Hint for Exercise 6.2.38
Is Σ *really* finite? Is D *really* well-defined?

Hint for Exercise 6.2.40
Start by proving that there is a finite description of the elements of $\mathbb{N}^*$ over a finite alphabet Φ. This defines an injection $\mathbb{N}^* \to \Phi^*$. Now given a countable alphabet Σ, use the injection $\Sigma \to \mathbb{N}$ given by Theorem 6.2.10 to construct an injection $\Sigma^* \to \mathbb{N}^*$, and therefore an injection $\Sigma^* \to \Phi^*$. Finally, use this to turn finite descriptions over Σ into finite descriptions over Φ.

Chapter 8 exercises

Section 7.1

Hint for Exercise 7.1.13
Remember that negative integers can be greatest common divisors too.

Hint for Exercise 7.1.15
Start by proving that d_1 and d_2 must divide each other.

Hint for Exercise 7.1.28
Example 7.1.25 would be a good starting point.

Hint for Exercise 7.1.42
This is essentially the same as Exercise 7.1.15.

Hint for Exercise 7.1.44
Define $m = \dfrac{ab}{\gcd(a,b)}$ and prove that m satisfies the definition of being a least common multiple of a and b (Definition 7.1.41). Then apply Exercise 7.1.42.

Section 7.2

Hint for Exercise 7.2.5
Use the factorial formula for binomial coefficients (Theorem 4.2.14).

Hint for Exercise 7.2.9
Assume $p = mn$ for some $m, n \in \mathbb{Z}$. Prove that m or n is a unit.

Hint for Exercise 7.2.23
What are the prime factors of $n! - 1$?

Section 7.3

Hint for Exercise 7.3.18
Consider the list $a^0, a^1, a^2, \ldots$. Since there are only finitely many remainders modulo n, we must have $a^i \equiv a^j \pmod{n}$ for some $0 \leqslant i < j$.

Hint for Exercise 7.3.25
First find the remainder of 244886 when divided by 12.

Hint for Exercise 7.3.28
Find a bijection $[p] \times C_n \to C_{pn}$, where $C_n = \{k \in [|n|] \mid k \perp n\}$. You will need to use the techniques of Section 6.1 in your proof.

Hint for Exercise 7.3.29
Start by proving that $k \in [pn]$ is *not* coprime to pn if and only if either $p \mid k$ or k is not coprime to n. You will need to use the techniques of Section 6.1 in your proof.

Hint for Exercise 7.3.34
Recall that $\varphi(100) = 40$—this was Example 7.3.30.

Hint for Exercise 7.3.38
You need to use the fact that p is prime at some point in your proof.

Hint for Exercise 7.3.39
Pair as many elements of $[p-1]$ as you can into multiplicative inverse pairs modulo p.

Hint for Exercise 7.3.49
This generalisation will be tricky! You may need to generalise the definitions and results about greatest common divisors and least common multiples that we have seen so far, including Bézout's lemma. You might want to try proving this first in the case that $n_i \perp n_j$ for all $i \neq j$.

Hint for Exercise 7.3.50
Observe that if $a, k \in \mathbb{Z}$ and $k \mid a$, then $k \mid a + k$.

Chapter 7 exercises

Hint for Question 7.E.4
Consider the linear Diophantine equation $(a+b)x + (c+d)y = 1$.

Hint for Question 7.E.11
Use the 'uniqueness' part of the fundamental theorem of arithmetic.

Hint for Question 7.E.12
Use appropriate restrictions of the functions from Question 7.E.11. Apply the 'existence' and

'uniqueness' parts of the fundamental theorem of arithmetic to prove surjectivity and injectivity, respectively.

Hint for Question 7.E.13

A sequence of k integers $(n_1, n_2, \ldots, n_k) \in \mathbb{Z}^k$ can be encoded as a sequence of $2k$ natural numbers $(i_1, |n_1|, i_2, |n_2|, \ldots, i_k, |n_k|) \in \mathbb{N}^{2k}$, where for each $j \in [k]$, either $i_j = 0$ or $i_j = 1$ according to the sign of n_j. For example we can encode $(10, -32, 81) \in \mathbb{Z}^3$ as $(0, 10, 1, 32, 0, 81) \in \mathbb{N}^6$.

Hint for Question 7.E.15

Start by proving that, for all $n \in \mathbb{N}$ and all $k \in \mathbb{N}$, there are exactly $\lfloor n/5^k \rfloor$ natural numbers $\leqslant n$ that are divisible by 5^k. Then consider how many zeros are contributed by each factor of $n!$ to the decimal expansion of $n!$.

Hint for Question 7.E.16

This will be similar to Question 7.E.15 in the end; to get started, consider the greatest prime factor of b.

Section 8.1

Hint for Exercise 8.1.19

Any function $f : X \to Y$ with finite domain can be specified by listing its values. For each $x \in X$, how many choices do you have for the value $f(x)$?

Hint for Exercise 8.1.24

The image (Definition 3.1.31) of an injection $[3] \to [4]$ must be a subset of $[4]$ of size three.

Hint for Exercise 8.1.41

How many ways can you select $k+1$ animals from a set containing n cats and one dog?

Hint for Exercise 8.1.44

Find two procedures for counting the number of pairs (U, u), such that $U \subseteq [n]$ is a k-element subset and $u \in U$. Equivalently, count the number of ways of forming a committee of size k from a population of size n, and then appointing one member of the committee to be the chair.

Hint for Exercise 8.1.46

Find an expression for $(a+b+c)!$ in terms of $a!$, $b!$, $c!$ and $\binom{a+b+c}{a,b,c}$, following the pattern of Theorem 8.1.43.

Section 8.2

Hint for Exercise 8.2.1

You will need to use the recursive definition of binomial coefficients

Hint for Exercise 8.2.27

Consider selecting a committee from a population of size n, with a subcommittee of size exactly ℓ; toggle whether the oldest member of the population that is not on the subcommittee is or is not on the committee. If you prefer mathematical objects, consider the set of pairs (A, B), where $A \subseteq B \subseteq [n]$ and $|A| = \ell$; toggle the least element of $[n] \setminus A$ in the set B.

Chapter 8 exercises

Hint for Question 8.E.8
Prove that $|X/\!\sim| \cdot k = n$ using the multiplication principle: find a two-step procedure for specifying an element of X by first considering its $\sim$-equivalence class.

Hint for Question 8.E.17
Find a way to apply Lemma 8.2.21.

Section 9.1

Section 9.2

Hint for Exercise 9.2.28
If $(x_n) \to a \neq 0$, show that $|x_n - a|$ is eventually small enough that no x_n can be equal to zero after a certain point in the sequence. On the other hand, there are plenty of sequences, *all* of whose terms are nonzero, which converge to zero—find one!

Hint for Exercise 9.2.40
Divide the numerator and denominator by n^r and apply Theorem 9.2.34 and Example 9.2.39.

Hint for Exercise 9.2.47
You might want to begin by solving Question 4.E.10.

Hint for Exercise 9.2.61
In the definition of a Cauchy sequence, observe that $x_m - x_n = (x_m - a) - (x_n - a)$, and apply the triangle inequality (Theorem 9.1.9).

Section 9.3

Hint for Exercise 9.3.5
Begin by observing that $\binom{n}{3}^{-1} = \dfrac{3}{n-2} - \dfrac{6}{n-1} + \dfrac{3}{n}$.

Hint for Exercise 9.3.13
Proceed by contraposition: suppose $j \in \mathbb{N}$ is least such that $u_j \neq v_j$—without loss of generality $u_j < v_j$—then

$$0 = \sum_{i \geqslant 0} \frac{v_i}{b^i} - \sum_{i \geqslant 0} \frac{u_i}{b^i} = \frac{v_j - u_j}{b^j} + \sum_{i > j} \frac{v_i - u_i}{b^i}$$

Prove that this is nonsense. You will use condition (iii) in Theorem 9.3.12 somewhere in your proof.

Hint for Exercise 9.3.33
Read the hypotheses of Theorem 9.3.30 very carefully.

Hint for Exercise 9.3.37
We've already proved that such a series exists—go find it!

Hint for Exercise 9.3.47

This exercise looks harder than it is. Write out the definitions of $\sum_{i\in I} a_{f(i)}$ and $\sum_{j\in J} a_j$ and apply Theorem 9.3.43.

Hint for Exercise 9.3.51

We know that $\dfrac{1}{(1-x)^2} = \left(\dfrac{1}{1-x}\right)^2 = \left(\sum_{n\geqslant 0} x^n\right)^2$ by Theorem 9.3.8. Multiply out this sum and see what happens.

Chapter 9 exercises

Hint for Question 9.E.5

Consider what the Cauchy–Schwarz inequality has to say about the vectors (a,b) and (c,d). Note also that $a+2b = (a,b)\dot{(}1,2)$ and $c+2d = (c,d)\cdot(1,2)$.

Hint for Question 9.E.8

Begin by comparing the harmonic and arithmetic means of the numbers a, b and c.

Hint for Question 9.E.9

Begin by seeing what the QM–AM inequality says about the numbers a^2, b^2 and c^2. Then apply the Cauchy–Schwarz inequality, noting that $ab+bc+ca$ is the scalar product of two vectors.

Section 10.1

Hint for Exercise 10.1.6

Start by finding a bijection $(-1,1)\to(a,b)$ and apply Exercise 10.1.5.

Section 10.2

Hint for Exercise 10.2.11

A nice trick is to construct an injection $[0,1)\times[0,1)\to[0,1)$ using decimal expansions, and then invoke the Cantor–Schröder–Bernstein theorem.

Chapter 10 exercises

Hint for Question 10.E.22

Start by proving that any disc $D\subseteq\mathbb{R}^2$ contains a point whose coordinates are both rational.

Hint for Question 10.E.35

Think combinatorially. You should not try to generalise the recursive definition of factorials (Definition 4.1.14); and you *very much should not* try to generalise the informal definition '$n! = 1\times 2\times\cdots\times n$'.

Hint for Question 10.E.53

Don't waste too much time on this question—it has been proved to be independent of the usual axioms of mathematics (see Section B.1), meaning that neither an answer of 'true' nor an answer of 'false' can be proved (so neither answer leads to a contradiction). An answer of 'false' is known as the *continuum hypothesis*.

Section 11.1

Section 11.2

Chapter 11 exercises

Hint for Exercise 12.1.25

Use the characterisation of gcd and lcm in terms of prime factorisation.

Hint for Exercise 12.1.28

Use distributivity, together with the fact that $y = y \vee \bot$ and $y = y \wedge \top$.

Hint for Exercise 12.1.38

This can be proved by swapping $\wedge$ with $\vee$ and $\top$ with $\bot$ in the proof of (a). But there is a shorter proof which uses the result of part (a) together with Proposition 12.1.30.

Chapter 12 exercises

Section B.1

Hint for Exercise B.1.15

Let X be the set whose existence is asserted by the axiom of infinity, and take $p(x)$ to be the formula $x \neq x$ in the axiom of separation.

Section B.2

Hint for Exercise B.2.11

Define two relations $\leqslant$ and $\leqslant'$ on $\mathbb{N}$, one using the recursive definition and the other using the definition as a logical formula; then use induction to prove that $m \leqslant n \Leftrightarrow m \leqslant' n$ for all $m, n \in \mathbb{N}$.

Hint for Exercise B.2.28

Prove that x is an additive inverse for $-x$ (in the sense of Axiom B.2.20(F4)) and use uniqueness of additive inverses. Likewise for x^{-1}.

Section B.3

Section C.2

Solutions to in-chapter exercises

Chapter 0

Solution to Exercise 0.2

(a) 2 is an even number.

(b) 2 is an odd number.

(c) Every even integer greater than two can be expressed as the sum of two positive prime numbers. [This is *Goldbach's conjecture*, proposed in 1742 by Christian Goldbach. It a famous example of an extant *open problem* in mathematics. Will you be the one to solve it?]

Solution to Exercise 0.13

Set $n = \lfloor x \rfloor$. Then $n \leqslant x < n+1$ by definition of floor. Since x is not an integer and n is, we have $n < x$; but then $n < x < n+1$, and so $(n+1)-1 < x \leqslant n+1$. This inequality demonstrates that $n+1$ is the ceiling of x, and so $\lceil x \rceil = n+1 = \lfloor x \rfloor + 1$, as required.

Solution to Exercise 0.14

Set $n = \lfloor -x \rfloor$. By definition of floor we have $n \leqslant -x < n+1$. Multiplying through by -1 gives $-n \geqslant x > -(n+1)$, or equivalently $-n-1 < x \leqslant -n$. Thus $-n$ satisfies the definition of being the ceiling of x, so that $\lceil x \rceil = -n = -\lfloor -x \rfloor$. Hence $\lfloor -x \rfloor = -\lceil x \rceil$, as required.

Now set $m = \lceil -x \rceil$. By definition of ceiling we have $m-1 < -x \leqslant m$. Multiplying through by -1 and rearranging gives $-m \leqslant x < -m+1$. Thus $\lfloor x \rfloor = -m = -\lceil -x \rceil$, and so $\lceil -x \rceil = -\lfloor x \rfloor$, as required.

Solution to Exercise 0.17

Let a, b, c, d be integers, with $b \neq 0$ and $d \neq 0$, such that $x = \dfrac{a}{b}$ and $y = \dfrac{c}{d}$; we know a, b, c, d exist by the definition of what it means for x and y to be rational. Then:

- $x - y = \dfrac{a}{b} - \dfrac{c}{d} = \frac{ad-bc}{bd}$;
- $xy = \dfrac{a}{b} \times \dfrac{c}{d} = \dfrac{ac}{bd}$; and
- If $y \neq 0$, then $c \neq 0$ (since if $c = 0$ then $y = \dfrac{c}{d} = \dfrac{0}{d} = 0$), and so $\dfrac{x}{y} = \dfrac{a/b}{c/d} = \dfrac{ad}{bc}$.

The expressions on the right-hand side of each of the above chains of equations is a fraction of two integers, with a nonzero integer in the denominator, and so the numbers $x+y$, $x-y$, xy and $\dfrac{x}{y}$ are all rational.

Solution to Exercise 0.21

(a) $2 \in \mathbb{N}$ is true, since it is a nonnegative whole number and can be used for counting (for example, I have 2 eyes).

(b) $-2 \in \mathbb{N}$ is false, since $-2 < 0$ but all natural numbers are nonnegative.

(c) $-2 \in \mathbb{Q}$ is true, since $-2 = \dfrac{-2}{1}$, which is an integer divided by a nonzero integer.

(d) $\dfrac{2}{5} \in \mathbb{Z}$ is false, since $0 < \dfrac{2}{5} < 1$, meaning that $\dfrac{2}{5}$ is strictly between two consecutive integers.

(e) $\dfrac{\sqrt{8}+\sqrt{2}}{\sqrt{8}-\sqrt{2}} \in \mathbb{Q}$ is true: note that $\sqrt{8} = 2\sqrt{2}$, so cancelling the common factor of $\sqrt{2}$ from the numerator and denominator reveals that this number is equal to $\dfrac{2+1}{2-1} = \dfrac{3}{1}$, which is an integer divided by a nonzero integer.

(f) $\pi^\pi \in \mathbb{Q}\ldots$ OK, sorry, this is (kind of) a trick question. Whether or not π^π is rational is an open problem.

(g) $\sqrt{-7} \in \mathbb{R}$ is false: the square of every real number is nonnegative, but whatever $\sqrt{-7}$ actually *is* (a mystery to be revealed later in this chapter—see Definition 0.62), it presumably satisfies $\sqrt{-7}^2 = -7$, which is negative.

(h) $\mathbb{N} \in \mathbb{Z}$ might appear true at first sight, since every natural number is indeed an integer. However, $\mathbb{N}$ is *the set of all natural numbers*, which is a set, not an integer, and so the proposition $\mathbb{N} \in \mathbb{Z}$ is false.

Solution to Exercise 0.24

Letting X be the set of all dyadic rationals, we have equivalently:

$$X = \left\{ x \in \mathbb{Q} \;\middle|\; x = \frac{a}{2^n} \text{ for some } a \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\}$$

or

$$X = \left\{ \frac{a}{2^n} \;\middle|\; a \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\}$$

Note that $2^n \neq 0$ for all $n \in \mathbb{N}$ so we don't need to worry about the part of the definition of rational numbers that requires the denominator of the fraction to be nonzero.

These are not the only two possible answers to this question but they are perhaps the closest to the definition using the two variantes of set-builder notation described earlier in the section.

Solution to Exercise 0.27

(a) The elements of $[10]$ are the natural numbers that are less than 10, and those that are also powers of 2 are $2^0 = 1$, $2^1 = 2$, $2^2 = 4$ and $2^3 = 8$; therefore $A = \{1,2,4,8\}$.

(b) B is obtained by raising 2 to the powers given by the elements of $[10]$; thus

$$B = \{2^0, 2^1, \ldots, 2^9\} = \{1,2,4,8,16,32,64,128,256,512\}$$

(c) An integer is an element of C precisely if its absolute value is odd and less than 100; thus

$$C = \left\{ n \in \mathbb{Z} \;\middle|\; |n| \in [100] \text{ and } n \text{ is odd} \right\}$$

Another perspective is that by adding 99 to the elements of C and then dividing by 2 we obtain the numbers $0,1,2,\ldots,99$, which are the elements of $[100]$; thus we obtain the element of C by reversing these operations, i.e. taking the elements of $[100]$, multiplying them by 2 and then subtracting 99. Thus

$$C = \{2x - 99 \mid x \in [100]\}$$

Solution to Exercise 0.30

(a) $[-2,4)$

(b) $[-2,4]$

(c) $\{-2,-1,0,1,2,3,4\}$ — note that interval notation $[-2,4]$ is *not* appropriate in this case, since the non-integer real numbers between -2 and 5 are not elements of the set.

(d) $(-\infty,4)$

(e) $[-2,\infty)$

(f) You might be tepmted to write $(-\infty,\infty)$ for this one, and that is generally considered acceptable, but it can even more concisely be expressed simply as $\mathbb{R}$.

Section 1.1

Section 1.2

Section 1.3

Section 2.1

Solution to Exercise 2.1.1

By Definition 0.1, the proposition $R \in R$ must be either true or false.

By the set-builder notation defining R, given a set X, we have $X \in R$ if and only if $X \notin X$, and therefore we also have $X \notin R$ if and only if $X \in X$.

Setting $X = R$, it follows that:

- If $R \in R$ is true, then $R \notin R$; and
- If $R \in R$ is false, then $R \in R$.

Therefore the only way $R \in R$ can be true or false is if it is *both* true *and* false. But that can't happen... right?

Solution to Exercise 2.1.3

When proving a proposition of the form $\forall x \in X,\, p(x)$, we let $x \in X$ and then prove $p(x)$—thus, we introduce a variable x, we assume $x \in X$, and we prove $p(x)$. This is equivalent to introducing an element $x \in \mathscr{U}$ and proving $x \in X \Rightarrow p(x)$, which explains the implication operator in $\forall x, x \in X \Rightarrow p(x)$.

When proving a proposition of the form $\exists x \in X,\, p(x)$, we define a specific object a, we prove $a \in X$, and we prove $p(a)$. This is equivalent to defining $a \in \mathscr{U}$ and then proving $a \in X \wedge p(a)$, which explains the conjunction operator in $\exists x, x \in X \wedge p(x)$.

We can use maximal negation (as in Section 1.3) to verify de Morgan's laws:

$$\begin{aligned}\neg\forall x \in X,\, p(x) &\equiv \neg\forall x,\, (x \in X \Rightarrow p(x)) \\ &\equiv \exists x,\, \neg(x \in X \Rightarrow p(x)) && \text{by Theorem 1.3.29} \\ &\equiv \exists x,\, (x \in X \wedge (\neg p(x))) && \text{by Exercise 1.3.27} \\ &\equiv \exists x \in X,\, \neg p(x)\end{aligned}$$

and

$$\begin{aligned}\neg\exists x \in X,\, p(x) &\equiv \neg\exists x,\, (x \in X \wedge p(x)) \\ &\equiv \forall x,\, \neg(x \in X \wedge p(x)) && \text{by Theorem 1.3.29} \\ &\equiv \forall x,\, (x \in X \Rightarrow (\neg p(x))) && \text{by Exercise 1.3.28} \\ &\equiv \forall x \in X,\, \neg p(x)\end{aligned}$$

Solution to Exercise 2.1.5
Note that by the new characterisation of set-builder notation, we have $R \in R$ if and only if $R \in \mathscr{U} \wedge R \notin R$.

If $R \in \mathscr{U}$, this means that $R \in R$ if and only if $R \notin R$, which is a contradiction for the same reason as in Exercise 2.1.1.

It follows that $R \notin \mathscr{U}$; but then all this means is that $R \in R$ is false.

Solution to Exercise 2.1.10
($\Rightarrow$) Assume $[a,b) \subseteq (c,d]$.

- Since $a \in [a,b)$, we have $a \in (c,d]$ and so $a > c$ as required.
- Towards a contradiction, assume $b > d$. Then $\frac{b+d}{2} < \frac{b+b}{2} = b$, so $\frac{b+d}{2} \in [a,b)$, and so $\frac{b+d}{2} \in (c,d]$. This implies that $\frac{b+d}{2} \leqslant d$; however since $b > d$ we have $\frac{b+d}{2} > \frac{d+d}{2} = d$, contradicting $\frac{b+d}{2} \leqslant d$. So indeed $b \leqslant d$, as required.

($\Leftarrow$) Assume $a > c$ and $b \leqslant d$, and let $x \in [a,b)$, so that $x \in \mathbb{R}$ and $a \leqslant x < b$. We need to prove that $x \in (c,d]$. Well, combining these inequalities, we have

$$c < a \leqslant x < b \leqslant d$$

and so $c < x \leqslant d$, so that $x \in (c,d]$, as required.

Solution to Exercise 2.1.17
Write $A = \{x \in \mathbb{R} \mid x^2 < x\}$; we prove that $A = (0,1)$ by double containment.

- ($\subseteq$) Let $x \in A$. Then $x \in \mathbb{R}$ and $x^2 < x$, so that $0 < x - x^2 = x(1-x)$. Now, if $x \leqslant 0$ then $1 - x > 0$, and so $x(1-x) \leqslant 0$; and if $x \geqslant 1$ then $x > 0$ and $1 - x \leqslant 0$, and so $x(1-x) \leqslant 0$. Thus we must have $0 < x < 1$, and so $x \in (0,1)$, as required.
- ($\supseteq$) Let $x \in (0,1)$. Then $x \in \mathbb{R}$ and $0 < x < 1$. Since $x > 0$, the inequality $x < 1$ is preserved when we multiply through by x, and so $x^2 \leqslant x$. Hence $x \in A$, as required.

By double containment, we have $A = (0,1)$.

Solution to Exercise 2.1.18
Proving $\{0,1\} = \{1,0\}$:

- $(\subseteq)$ Let $x \in \{0,1\}$. Then either $x = 0 \in \{1,0\}$ or $x = 1 \in \{1,0\}$; in both cases we see that $x \in \{1,0\}$.
- $(\supseteq)$ Identical to $(\subseteq)$, but swap the roles of 0 and 1.

Hence $\{0,1\} = \{1,0\}$ by double containment.

Proving $\{0,0\} = \{0\}$:

- $(\subseteq)$ Let $x \in \{0,0\}$. Then $x = 0$ or $x = 0\ldots$ so $x = 0 \in \{0\}$.
- $(\supseteq)$ Let $x \in \{0\}$. Then $x = 0 \in \{0,0\}$.

Hence $\{0,0\} = \{0\}$ by double containment.

These proofs are a little bit silly, but they demonstrate that in list notation for sets, we can change the order that the elements are listed and we can repeat elements in the list without changing the set.

Section 2.2

Section 3.1

Section 3.2

Solution to Exercise 3.2.19

- (**Injectivity**) Let $a, b \in X$ and assume that $\mathrm{id}_X(a) = \mathrm{id}_X(b)$. Then $a = b$ by definition of id_X, as required.
- (**Surjectivity**) Let $x \in X$. Then $x = \mathrm{id}_X(x)$ by definition of id_X, as required.

Since id_X is both injective and surjective, it is bijective.

Section 4.1

Section 4.2

Section 4.3

Section 5.1

Section 5.2

Section 6.1

Solution to Exercise 6.1.3
Let $n \in \mathbb{N}$. The identity function $\mathrm{id}_{[n]} : [n] \to [n]$ is a bijection by Exercise 3.2.19, so is an enumeration of $[n]$.

Solution to Exercise 6.1.16
Let $i : A \to X$ be the *inclusion function*, defined by $i(x) = x$ for all $x \in A$. For all $a, b \in A$, if $i(a) = i(b)$, then $a = b$ by definition of i, so that i is injective. Since X is finite and $i : A \to X$ is an injection, it follows from Theorem 6.1.13(a).

Section 6.2

Section 7.1

Section 7.2

Section 7.3

Section 8.1

Section 8.2

Section 9.1

Section 9.2

Section 9.3

Section 10.1

Section 10.2

Section 11.1

Section 11.2

Section B.1

Section B.2

Section B.3

Appendix D

Typesetting mathematics with LaTeX

Being able to type up your mathematical writing is a beneficial skill to have, and is one that is expected of anyone working in a mathematical field. Unfortunately, most Office-style WYSIWYG ('what you see is what you get') text editors are not designed for this task—it becomes quickly cumbersome to deal with complicated mathematical notation, and fast alternations between notation and prose.

LaTeX is a markup language that allows you to input both regular text and mathematical notation, using code to input all symbols and configure the document's formatting and structure. What follows is a brisk introduction to LaTeX, that should suffice for the purposes of this book.

The word LaTeX is pronounced like 'LAY-tek' or 'LAH-tek', with a hard or soft 'k' sound—the 'X' is meant to resemble the Greek letter *chi* (χ), so is pronounced by some people as such. It doesn't really matter how you say it, but do be warned that if you pronounce it like 'LAY-teks' then people will think you're talking about something somewhat different.

Finding the software

You can use LaTeX by installing it on your computer, or by using a web-based editor on an internet browser. There are advantages and disadvantages to both.

- On a web-based editor, everything is set up and ready to go, but you have less control over how everything compiles and you need an internet connection at all times when editing your files.
- On a computer, you have more control over how your files compile, you have access to all the logs and auxiliary files (I won't go into this) and you don't need an internet connection—but it can be harder to use different packages, you're at the mercy of your own machine's limitations, and it's more technically involved.

  Installing LaTeX on a computer is slightly more complicated. In order to make LaTeX documents on your computer, you need both a *compiler*, for turning the code into a readable document, and an *editor*, for writing the code and facilitating the compilation process.

There are many online and computer-based options available; you are encouraged to research the options, and I make no recommendations for or against any particular LaTeX implementations. As a starting point, though, I present here the LaTeX implementations used for writing this book:

- **Computer implementation:**
  - ◇ **Compiler:** *TeX Live* (`https://www.tug.org/texlive/`);
  - ◇ **Editor:** *TeXstudio* (`https://www.texstudio.org/`);
- **Web-based implementation:** *Overleaf* (`https://www.overleaf.com/`).

TeX Live and TeXstudio are both free, open-source and cross-platform. Overleaf is a proprietary service with both free and paid subscriptions.

Getting started

When you have settled on an online LaTeX editor or installed LaTeX on your computer, you can start editing. The files containing the LaTeX code are plain-text files with the file extension `.tex` and consists of two components: a *header* and a *body*. Worrying about the details of what goes in the header and what goes in the body is not recommended if you are new to LaTeX so, with this in mind, a template can be downloaded from the book's website:

`https://infinitedescent.org/latex/`

The code is replicated at the end of this section, on page 551.

Text mode and math mode

Before we get into the nitty-gritty, I should mention the difference between 'text mode' and 'math mode'.

- **Text mode** is the default mode: the stuff you type will appear as text, and this is the mode you should use when writing anything that isn't mathematical notation.
- You should use **math mode** when you're typing anything which is mathematical notation, including variables, numbers, fractions, square roots, powers, sums, products, binomial coefficients, and so on.

To enter math mode, enclose whatever mathematical notation you are writing with dollar signs ($). For example, if I type `$E=mc^{2}$` then LaTeX shows $E = mc^2$. Sometimes it is convenient to put longer expressions on their own line, in which case you can enclose it between `\[` and `\]`; for example, if I type

```
\[ a^{2}+b^{2}+c^{2}=ab+bc+ca \]
```

then LaTeX displays

$$a^2 + b^2 + c^2 = ab + bc + ca$$

on a line all of its own.

If you need to type text inside math mode (enclosed by $ signs), you can do that using `\text{...}`, for example:

TeX code

```
\[ \sum_{i=1}^n i = \frac{n(n+1)}{2} \text{ for all } n \in \mathbb{N} \]
```

Output

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \text{ for all } n \in \mathbb{N}$$

Note the spaces before and after 'for all'; had I left those out of the code, they would not appear because LaTeX ignores spacing in math mode. You can force a space by putting a backslash before a space, for example `$a b$` gives ab but `$a\ b$` gives $a\ b$.

All mathematical notation should be in math mode, including single variables. Notice the difference between the following two lines:

If a and b are both even then so is a+b.

If a and b are both even then so is $a+b$.

While the first is written entirely in text mode, the second is written using math mode for the variables and + sign. Although the differences may not seem big, spread over a whole document it is much clearer when math mode is used (as in the second example). Sometimes ambiguities appear, and in any case LaTeX does a much better job at displaying mathematical notation when it is typed in math mode.

Table of mathematical symbols

The following table is a quick reference for the most commonly-used symbols in this book. A complete index of notation can be found at the end of the book.

Logic		
conjunction, disjunction	$\wedge$, $\vee$	`\wedge`, `\vee`
negation	$\neg$	`\neg`
implication, biconditional	$\Rightarrow$, $\Leftrightarrow$	`\Rightarrow`, `\Leftrightarrow`
exclusive disjunction	$\oplus$	`\oplus`
true, false (in truth table)	✓, $\times$	`\checkmark`, `\times`
quantifiers (universal, existential)	$\forall$, $\exists$	`\forall`, `\exists`
Set theory		
element, subset	$\in$, $\subseteq$	`\in`, `\subseteq`
not equal, proper subset	$\neq$, $\subsetneqq$	`\ne`, `\subsetneqq`
intersection, (indexed)	$\cap$, $\bigcap_{i=1}^{n}$	`\cap`, `\bigcap_{i=1}^{n}`
union, (indexed)	$\cup$, $\bigcup_{i=1}^{n}$	`\cup`, `\bigcup_{i=1}^{n}`
relative complement, complement	$X \setminus Y$, X^c	`\setminus`, `X^c`
product, (indexed)	$\times$, $\prod_{i=1}^{n}$	`\times`, `\prod_{i=1}^{n}`
implied lists	$\{1, \dots, n\}$	`\{ 1, \dots, n \}`
indexed sets	$\{x_i \mid i \in I\}$	`\{ x_i \mid i \in I \}`
set-builder notation	$\{x \mid p(x)\}$	`\{ x \mid p(x) \}`
empty, universal set	$\varnothing$, $\mathscr{U}$	`\varnothing`, `\mathcal{U}`
number sets	$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$	`\mathbb{N}`, `\mathbb{Z}`, etc.
Numbers and combinatorics		
multiplication	$m \times n$, $m \cdot n$	`\times`, `\cdot`
fractions, exponents	$\frac{m}{n}$, m^n	`\frac{m}{n}`, `m^{n}`
order relations	$\leqslant$, $\geqslant$	`\le`, `\ge`
divisibility, (non-)	$m \mid n$, $m \nmid n$	`\mid`, `\nmid`
binomial coefficient	$\binom{n}{k}$	`\binom{n}{k}`
indexed sum, product	$\sum_{i=1}^{n} a_i$, $\prod_{i=1}^{n} a_i$	`\sum_{i=1}^{n} a_i`, `\prod`
modular arithmetic	$a \equiv b \bmod n$	`a \equiv b \bmod{n}`
Functions and relations		
functions	$f : X \to Y$	`f : X \to Y`
composition	$g \circ f$	`\circ`
isomorphism	$\cong$	`\cong`
equivalence relations	$\sim$, $\approx$	`\sim`, `\approx`
Structured sets		
order relation	$\preccurlyeq$, $\prec$	`\preceq`, `\prec`
group operations	$\cdot$, $\star$, $\circ$	`\cdot`, `\star`, `\circ`

Organisation and formatting

When typing up solutions to problem, organisation can be the difference between a masterpiece and an unreadable heap of notation. Here are some tips to help you organise your work:

Sections and paragraphs

You can split your work up into sections, subsections, subsubsections, and even subsubsubsections. To do this, use `\section{Section title}` or `\section*{Section title}`; the former includes a section number, and the latter omits it. To start a new paragraph, simply make two new lines in the code.

Bulleted and enumerated lists

Sometimes it is useful to use bullet points or give an enumerated list. For example, in these notes, I separate the base case from the induction step in proofs by induction by using bullet points.

For a bulleted list you can use the `itemize` environment:

TeX code

```
\begin{itemize}
\item Something here\dots
\item You can also make a list inside
another list:
  \begin{itemize}
  \item Like this.
  \item Isn't it fun?
  \end{itemize}
\item Well, not that fun.
\end{itemize}
```

Output

- Something here...
- You can also make a list inside another list:
  - ◇ Like this.
  - ◇ Isn't it fun?
- Well, not that fun.

For an enumerated list, you can use the `enumerate` environment. You can play around with different methods of enumeration, which you specify in square brackets `[...]`; this book most frequently uses `(i)`, `(a)` and `(1)`:

TeX code

```
\begin{enumerate}[(a)]
\item Here's the first thing;
\item Here's the second thing;
\item Here's the third thing.
\end{enumerate}
```

Output

(a) Here's the first thing;

(b) Here's the second thing;

(c) Here's the third thing.

Definitions, results and proofs

If you use the provided templates, you can make definitions, and state and prove results, using the following environments:

`definition`, `example`, `proposition`, `theorem`, `lemma`, `corollary`, `proof`

They are given a number, such as **Definition 3** or **Theorem 2.11**, depending on how your document is set up.

Here's an example of a theorem appearing in the third section of a document, in which five definitions, results or examples come before it:

TeX code

```
\begin{theorem}
Let $a, b \in \mathbb{Z}$. Then
$a^2+b^2 \ge 0$.
\end{theorem}

\begin{proof}
Exercise to the reader.
\end{proof}
```

Output

Theorem 3.6. Let $a, b \in \mathbb{Z}$. Then $a^2 + b^2 \geqslant 0$.
Proof. Exercise to the reader.

Note that the box (□) designating the end of the proof is inserted automatically when you close the `proof` environment.

Labels

As you change the contents of a document, the numbering of the definitions, examples and results might change. To refer to a specific result, instead of typing the number and having to change it each time the number changes, you can use the `\label`) and `\ref`) commands.

An example of this in action is as follows:

TeX code

```
\begin{definition}
\label{defDivides}
Say $a$ \textbf{divides} $b$ if there
exists $k \in \mathbb{Z}$ such that
$ka = b$.
\end{definition}

We will use Definition
\ref{defDivides} for absolutely
nothing.
```

Output

Definition 2.11. Say a **divides** b if there exists $k \in \mathbb{Z}$ such that $ka = b$.

We will use Definition 2.11 for absolutely nothing.

Formatting

In text mode. To put the icing on the cake, you might want to make some words **bold** or *italicised*. This is simple: for bold text type `\textbf{text here}` and for italic text type `\textit{text here}`. In TeXstudio and Overleaf you can press `Ctrl+B` and `Ctrl+I` to avoid having to type all this out. Other useful fonts include `monospace` (`\texttt{text here}`), sans-serif (`\textsf{text here}`) and underlined (`\underline{text here}`).

In math mode. There are also various fonts or font styles that you can use inside math mode, including:

- Roman (i.e. not italic): AaBbCc, `\mathrm{AaBbCc}`;
- Bold: **AaBbCc**, `\mathbf{AaBbCc}`;
- Sans-serif: AaBbCc, `\mathsf{AaBbCc}`;
- Blackboard bold: $\mathbb{ABCDE}$, `\mathbb{ABCDE}` — only capital letters;
- Fraktur: $\mathfrak{AaBbCc}$, `\mathfrak{AaBbCc}`;
- Calligraphic: $\mathcal{ABCDE}$, `\mathcal{ABCDE}` — only capital letters;

Tables

Tables can be created using the `tabular` environment. You can specify how columns are aligned and separated as an argument to the command `\begin{tabular}`: write `l` or `c` or `r` to specify that a column should be aligned left, centre or right, respectively. If you want columns to be separated by a single or double line, enter a single or double bar (`|` or `||`), respectively.

Columns are then separated by ampersands (`\&`) and you can move to a new row by entering a double-backslash (`\\`). To insert a horizontal line between two rows, simply enter `\hline`.

Here's an example:

TeX code

```
\begin{tabular}{c|ccc}
$\times$ & 1 & 2 & 3 \\ \hline
1 & 1 & 2 & 3 \\
2 & 2 & 4 & 6 \\
3 & 3 & 6 & 9
\end{tabular}
```

Output

$\times$	1	2	3
1	1	2	3
2	2	4	6
3	3	6	9

Aligned equations

Occasionally a proof may require you to demonstrate that two terms are equal by proving a sequence of intermediate equations. This can be done using the `align*` environment, which behaves much like the `tabular` environment.

New lines are introduced by inserting a double-backslash (`\\`), and alignment points are introduced with an ampersand (`&`). For example:

TeX code

```
\begin{align*}
(n+1)! - n!
  &= (n+1)n! - n! \\
  &= n \cdot n! + n! - n! \\
  &= n \cdot n!
\end{align*}
```

Output

$$\begin{aligned}(n+1)!-n! &= (n+1)n!-n! \\ &= n\cdot n!+n!-n! \\ &= n\cdot n!\end{aligned}$$

Note that the `align*` environment automatically enters into math mode, so no dollar signs (`$`) are needed.

Entering more ampersands will create more columns, whose alignment alternates (right, left, right, left, and so on). For example, to add annotations to each line, you can enter a double ampersand (`&&`). For example:

TeX code

```
\begin{align*}
(n+1)! - n!
  & = (n+1)n! - n! && \text{by recursive def of factorials} \\
  & = n \cdot n! + n! - n! && \text{by distributivity} \\
  & = n \cdot n! && \text{by cancellation}
\end{align*}
```

Output

$$\begin{aligned}(n+1)!-n! &= (n+1)n!-n! && \text{by recursive def of factorials}\\ &= n\cdot n!+n!-n! && \text{by distributivity}\\ &= n\cdot n! && \text{by cancellation}\end{aligned}$$

Note again that, because the `align*` environment automatically enters math mode, any annotations must be made within the `\text{...}` command.

Graphics

Images can then be inserted using the `\includegraphics` command. The format is `\includegraphics[parameters]{filename}` where `parameters` denotes information telling LaTeX how large you want the image to be, and `filename` is the name of the image file, which includes the path relative to the main `.tex` file. For example if *donkey.png* is stored in a directory called *images*, you would enter '`images/donkey.png`' instead of '`donkey.png`'.

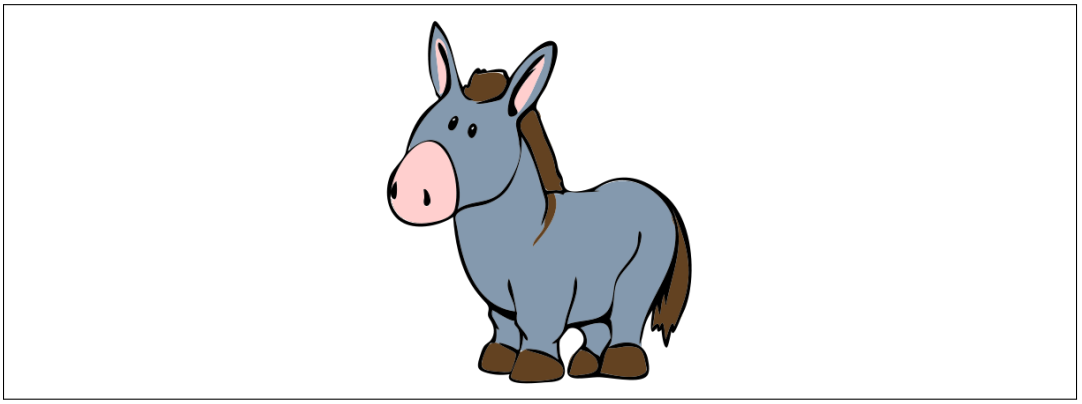
The simplest way to control the size of the image is to enter `[width=k\textwidth]`, where `k` is a scaling factor between 0 and 1.

For example:

TeX code

```
\begin{center}
\includegraphics[width=0.3\textwidth]{media/donkey.png}
\end{center}
```

Output



More advanced techniques

I should take a moment to emphasise that what really matters is your ability to communicate mathematical arguments clearly and correctly. The LaTeX tools discussed so far in this section are more than sufficient for our purposes.

However, if you are interested in pushing your LaTeX skills further or there is a feature you're unsure about how to implement, then I recommend browsing or searching one of the following websites:

- `http://tex.stackexchange.com` — Q&A website about LaTeX
- `https://en.wikibooks.org/wiki/LaTeX` — online LaTeX manual

Practice page

Try to recreate the following page, remembering to use `\label` and `\ref` to refer to enumerated items (such as 'Proposition 1.2').

Squarefree integers

Carl Friedrich Gauss, Wednesday 14th September 1831

Introduction

When you've written this page, you will be unstoppable, at least as far as typesetting mathematics is concerned. You will need to implement:

- Text mode stuff: sections, paragraphs, text formatting, labels and references, lists;
- Math mode stuff: definitions and results, aligned equations, etc.

So let's get on with it!

1 Squarefree integers

1.1 Definition and an elementary result

Definition 1.1. An integer a is **squarefree** if it is divisible by no perfect square other than 1. That is, if n^2 divides a then $n^2 = 1$.

Proposition 1.2. A non-zero non-unit a is squarefree if and only if

$$a = p_1 \times p_2 \times \cdots \times p_n$$

for distinct primes $p_1, p_2, \ldots, p_n$.

Proof. We leave the proof as an exercise to the reader. □

1.2 Some examples

Example 1.3. Some concrete examples include:

(i) 5610 is squarefree by Proposition 1.2, since

$$\begin{aligned} 5610 &= 10 \times 561 \\ &= (2 \times 5) \times (11 \times 17) \end{aligned}$$

(ii) 12 is not squarefree since $4 \mid 12$ and $4 = 2^2$.

1

Template file

What follows is a template `.tex` file to get you started; it can be downloaded from https://infinitedescent.org/latex/.

```
\documentclass[11pt]{article}

% Edit the following to change the title, author name and date
\title{A \LaTeX{} document}
\author{Firstnametina McLastnamerson}
\date{Someday 0th Jantember 3000}

% Packages
\usepackage{amsmath}
\usepackage{amsfonts}
\usepackage{amssymb}
\usepackage{amsthm}
\usepackage{enumerate}
\usepackage{geometry}
\usepackage{graphicx}
\usepackage{hyperref}
\usepackage{xcolor}

% Page setup
\setlength{\parskip}{10pt}
\setlength{\parindent}{0pt}
\geometry{
    paper={letterpaper}, % Change to 'a4paper' for A4 size
    marginratio={1:1},
    margin={1.25in}
}

% Theorem environments
\theoremstyle{definition}
\newtheorem{theorem}{Theorem}
\newtheorem{lemma}[theorem]{Lemma}
\newtheorem{corollary}[theorem]{Corollary}
\newtheorem{proposition}[theorem]{Proposition}
\newtheorem{definition}[theorem]{Definition}
\newtheorem{example}[theorem]{Example}

\begin{document}
\maketitle

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Start of document body %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Hello world! Did you know that $3^2 + 4^2 = 5^2$?

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% End of document body   %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
\end{document}
```

Indices

Index of topics

Index of vocabulary

Index of notation

Index of LaTeX commands

Licence

This book, in both physical and electronic form, is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Public License. The licence is replicated below from the Creative Commons website:

https://creativecommons.org/licenses/by-sa/4.0/legalcode

Please read the content of this licence carefully if you intend to share or adapt the material in this book.

If you have questions or would like to request permissions not granted by this licence, then contact the author (Clive Newstead, cnewstead@infinitedescent.org).

Creative Commons Attribution-ShareAlike 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution-ShareAlike 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 — Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.

b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.

c. **BY-SA Compatible License** means a license listed at creativecommons.org/compatiblelicenses, approved by Creative Commons as essentially the equivalent of this Public License.

d. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.

e. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.

f. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.

g. **License Elements** means the license attributes listed in the name of a Creative Commons Public License. The License Elements of this Public License are Attribution and ShareAlike.

h. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.

i. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.

j. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.

k. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.

l. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.

m. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. **License grant.**

1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:
   A. reproduce and Share the Licensed Material, in whole or in part; and
   B. produce, reproduce, and Share Adapted Material.
2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.
3. Term. The term of this Public License is specified in Section 6(a).
4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures. For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.
5. Downstream recipients.

A. *Offer from the Licensor – Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.*

B. *Additional offer from the Licensor – Adapted Material. Every recipient of Adapted Material from You automatically receives an offer from the Licensor to exercise the Licensed Rights in the Adapted Material under the conditions of the Adapter's License You apply.*

C. *No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.*

6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. **Other rights.**

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.
2. Patent and trademark rights are not licensed under this Public License.
3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

Section 3 – License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. **Attribution.**

1. If You Share the Licensed Material (including in modified form), You must:
   A. retain the following if it is supplied by the Licensor with the Licensed Material: identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated); a copyright notice;
      i. a notice that refers to this Public License;
      ii. a notice that refers to the disclaimer of warranties;
      iii. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;
   B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and
   C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.
2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.
3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.

b. **ShareAlike.**

In addition to the conditions in Section 3(a), if You Share Adapted Material You produce, the following conditions also apply.

1. The Adapter's License You apply must be a Creative Commons license with the same License Elements, this version or later, or a BY-SA Compatible License.
2. You must include the text of, or the URI or hyperlink to, the Adapter's License You apply. You may satisfy this condition in any reasonable manner based on the medium, means, and context in which You Share Adapted Material.
3. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, Adapted Material that restrict exercise of the rights granted under the Adapter's License You apply.

Section 4 – Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;
b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material, including for purposes of Section 3(b); and
c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors, whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.

b. To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.

c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 – Term and Termination.

a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.

b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:

   1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or

   2. upon express reinstatement by the Licensor.

   For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.

d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 – Other Terms and Conditions.

a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.

b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 – Interpretation.

a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License.

b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.

c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.

d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.