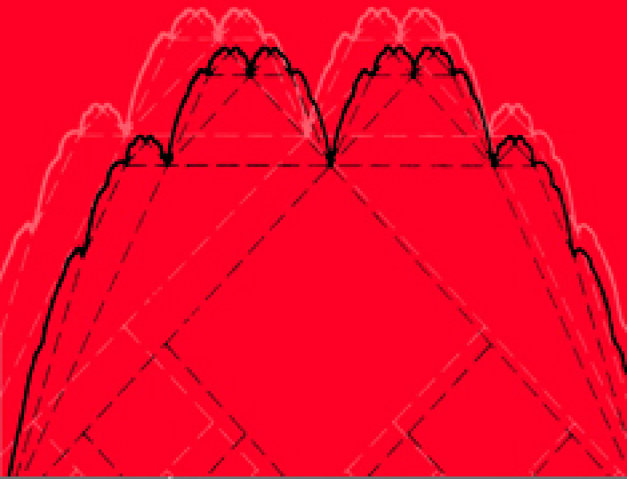
A First Course in Mathematical Analysis

David Brannan



CAMBRIDGE www.cambridge.org/9780521864398

This page intentionally left blank

A First Course in Mathematical Analysis

Mathematical Analysis (often called Advanced Calculus) is generally found by students to be one of their hardest courses in Mathematics. This text uses the so-called sequential approach to continuity, differentiability and integration to make it easier to understand the subject.

Topics that are generally glossed over in the standard Calculus courses are given careful study here. For example, what exactly is a 'continuous' function? And how exactly can one give a careful definition of 'integral'? This latter is often one of the mysterious points in a Calculus course – and it is quite tricky to give a rigorous treatment of integration!

The text has a large number of diagrams and helpful margin notes, and uses many graded examples and exercises, often with complete solutions, to guide students through the tricky points. It is suitable for self study or use in parallel with a standard university course on the subject.

A First Course in Mathematical Analysis

DAVID ALEXANDER BRANNAN

Published in association with The Open University

The Open University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521864398

First published in print format 2006

ISBN-13 978-0-511-34568-5 eBook (Adobe Reader)
ISBN-10 0-511-34568-2 eBook (Adobe Reader)

ISBN-13 978-0-521-86439-8 hardback
ISBN-10 0-521-86439-9 hardback

To my wife *Margaret*
and my sons *David*, *Joseph* and *Michael*

Contents

Preface *page* ix

1 Numbers 1
1.1 Real numbers 2
1.2 Inequalities 9
1.3 Proving inequalities 14
1.4 Least upper bounds and greatest lower bounds 22
1.5 Manipulating real numbers 30
1.6 Exercises 35

2 Sequences 37
2.1 Introducing sequences 38
2.2 Null sequences 43
2.3 Convergent sequences 52
2.4 Divergent sequences 61
2.5 The Monotone Convergence Theorem 68
2.6 Exercises 79

3 Series 83
3.1 Introducing series 84
3.2 Series with non-negative terms 92
3.3 Series with positive and negative terms 103
3.4 The exponential function $x \mapsto e^x$ 122
3.5 Exercises 127

4 Continuity 130
4.1 Continuous functions 131
4.2 Properties of continuous functions 143
4.3 Inverse functions 151
4.4 Defining exponential functions 161
4.5 Exercises 164

5 Limits and continuity 167
5.1 Limits of functions 168
5.2 Asymptotic behaviour of functions 176
5.3 Limits of functions – using ε and δ 181
5.4 Continuity – using ε and δ 193
5.5 Uniform continuity 200
5.6 Exercises 203

6 Differentiation 205
6.1 Differentiable functions 206
6.2 Rules for differentiation 216
6.3 Rolle's Theorem 228
6.4 The Mean Value Theorem 232
6.5 L'Hôpital's Rule 238
6.6 The Blancmange function 244
6.7 Exercises 252

7 Integration 255
7.1 The Riemann integral 256
7.2 Properties of integrals 272
7.3 Fundamental Theorem of Calculus 282
7.4 Inequalities for integrals and their applications 288
7.5 Stirling's Formula for $n!$ 303
7.6 Exercises 309

8 Power series 313
8.1 Taylor polynomials 314
8.2 Taylor's Theorem 320
8.3 Convergence of power series 329
8.4 Manipulating power series 338
8.5 Numerical estimates for π 346
8.6 Exercises 350

Appendix 1 Sets, functions and proofs 354

Appendix 2 Standard derivatives and primitives 359

Appendix 3 The first 1000 decimal places of $\sqrt{2}$, e and π 361

Appendix 4 Solutions to the problems 363
Chapter 1 363
Chapter 2 371
Chapter 3 382
Chapter 4 393
Chapter 5 402
Chapter 6 413
Chapter 7 426
Chapter 8 443

Index 457

Preface

Analysis is a central topic in Mathematics, many of whose branches use key analytic tools. Analysis also has important applications in Applied Mathematics, Physics and Engineering, where a good appreciation of the underlying ideas of Analysis is necessary for a modern graduate.

Changes in the school curriculum over the last few decades have resulted in many students finding Analysis very difficult. The author believes that Analysis nowadays has an unjustified reputation for being hard, caused by the traditional university approach of providing students with a highly polished exposition in lectures and associated textbooks that make it impossible for the average learner to grasp the core ideas. Many students end up agreeing with the German poet and philosopher Goethe who wrote that 'Mathematicians are like Frenchmen: whatever you say to them, they translate into their own language, and forthwith it is something entirely different!'

Johann Wolfgang von Goethe (1749–1832) is said to have studied all areas of science of his day except mathematics – for which he had no aptitude.

Since 1971, the Open University in United Kingdom has taught Mathematics to students in their own homes via specially written correspondence texts, and has traditionally given Analysis a central position in its curriculum. Its philosophy is to provide clear and complete explanations of topics, and to teach these in a way that students can understand without much external help. As a result, students should be able to learn, and to enjoy learning, the key concepts of the subject in an uncluttered way. This book arises from correspondence texts for its course *Introduction to Pure Mathematics*, that has now been studied successfully by over ten thousand students.

This book is therefore different from most Mathematics textbooks! It adopts a student-friendly approach, being designed for study by a student on their own OR in parallel with a course that uses as set text either this text or another text. But this is the text that the student is likely to use to learn the subject from. The author hopes that readers will gain enormous pleasure from the subject's beauty and that this will encourage them to undertake further study of Mathematics!

Once a student has grasped the principal notions of *limit* and *continuous function* in terms of inequalities involving the three symbols ε, X and δ, they will quickly understand the unity of areas of Analysis such as limits, continuity, differentiability and integrability. Then they will thoroughly enjoy the beauty of some of the arguments used to prove key theorems – whether their proofs are short or long.

Calculus is the initial study of limits, continuity, differentiation and integration, where functions are assumed to be well-behaved. Thus all functions continuous on an interval are assumed to be differentiable at most points in the interval, and so on. However, Mathematics is not that simple! For example, there exist functions that are continuous everywhere on $\mathbb{R}$, but differentiable nowhere on $\mathbb{R}$; this discovery by Karl Weierstrass in 1872 caused a sensation in the mathematical community. In *Analysis* (sometimes called *Advanced Calculus*) we make no assumptions about the behaviour of functions – and the result is that we sometimes come across real surprises!

The book has two principal features in its approach that make it stand out from among other Analysis texts.

Firstly, this book uses the ‘sequential approach’ to Analysis. All too often students starting on the subject find that they cannot grasp the significance of both ε and δ simultaneously. This means that the whole underlying idea about what is happening is lost, and the student takes a very long time to master the topic – or, in many cases in fact, never masters the topic and acquires a strong dislike of it. In the sequential approach they proceed at a more leisurely pace to understand the notion of limit using ε and X – to handle convergent sequences – before coming across the other symbol δ, used in conjunction with ε to handle continuous functions. This approach avoids the conventional student horror at the perceived ‘difficulty of Analysis’. Also, it avoids the necessity to re-prove broadly similar results in a range of settings – for example, results on the sum of two sequences, of two series, of two continuous functions and of two differentiable functions.

Secondly, this book makes great efforts to teach the ε–δ approach too. After students have had a first pass at convergence of sequences and series and at continuity using ‘the sequential approach’, they then meet ‘the ε–δ approach’, explained carefully and motivated by a clear ‘ε–δ game’ discussion. This makes the new approach seem very natural, and this is motivated by using each approach in later work in the appropriate situation. By the end of the book, students should have a good facility at using both the sequential approach and the ε–δ approach to proofs in Analysis, and should be better prepared for later study of Analysis than students who have acquired only a weak understanding of the conventional approach.

Outline of the content of the book

In Chapter 1, we define *real numbers* to be decimals. Rather than give a heavy discussion of *least upper bound* and *greatest lower bound*, we give an introduction to these matters sufficient for our purposes, and the full discussion is postponed to Chapter 7, where it is more timely. We also study *inequalities*, and their properties and proofs.

In Chapter 2, we define *convergent sequences* and examine their properties, basing the discussion on the notion of *null sequence*, which simplifies matters considerably. We also look at *divergent sequences*, sequences defined by *recurrence formulas* and particular sequences which converge to π and e.

In Chapter 3, we define *convergent infinite series*, and establish a number of tests for determining whether a given series is convergent or divergent. We demonstrate the equivalence of the two definitions of the *exponential function* $x \mapsto e^x$, and prove that the number e is irrational.

We define e^x as $\lim\limits_{n\to\infty}\left(1+\frac{x}{n}\right)^n$ and as $\sum\limits_{n=0}^{\infty}\frac{x^n}{n!}$.

In Chapter 4, we define carefully what we mean by a *continuous function*, in terms of sequences, and establish the key properties of continuous functions. We also give a rigorous definition of the *exponential function* $x \mapsto a^x$.

In Chapter 5, we define the *limit of a function* as x tends to c or as x tends to ∞ in terms of the convergence of sequences. Then we introduce the ε–δ definitions of limit and continuity, and check that these are equivalent to the earlier definitions in terms of sequences. We also look briefly at *uniform continuity*.

In Chapter 6, we define what we mean by a *differentiable function*, using *difference quotients* $Q(h)$; this enables us to use our earlier results on limits to prove corresponding results for differentiable functions. We establish some interesting properties of differentiable functions. Finally, we construct the *Blancmange function* that is continuous everywhere on $\mathbb{R}$, but differentiable nowhere on $\mathbb{R}$.

In Chapter 7, we give a careful definition of what we mean by an *integrable function*, and establish a number of related criteria for establishing whether a given function is integrable or not. Our integral is the so-called *Riemann integral*, defined in terms of upper and lower Riemann sums. We check the standard properties of integrals and verify a number of standard approaches for calculating definite integrals. Then we give a number of applications of integrals to limits of certain sequences and series and prove *Stirling's Formula*.

Stirling's Formula says that, for large n, $n!$ is 'roughly' $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, in a sense that we explain.

Finally, in Chapter 8, we study the convergence and properties of *power series*. The chapter ends with a marvellous proof of the irrationality of the number π that uses a whole range of the techniques that have been met in the previous chapters.

For completeness and for students' convenience, we give a brief guide to our notation for sets and functions, together with a brief indication of the logic involved in proofs in Mathematics (in particular, the Principle of Mathematical Induction) in Appendix 1. Appendix 2 contains a list of standard derivatives and primitives and Appendix 3 the first 1000 decimal places in the values of the numbers $\sqrt{2}$, π and e. Appendix 4 contains full solutions to all the problems set during each chapter.

Solutions are not given to the exercises at the end of each chapter, however. Lecturers/instructors may wish to use these exercises in homework assignments.

Study guide

This book assumes that students have a fair understanding of Calculus. The assumptions on technical background are deliberately kept slight, however, so that students can concentrate on the newer aspects of the subject 'Analysis'.

Most students will have met some of the material in the early chapters previously. Although this means that they can therefore proceed fairly quickly through some sections, it does NOT mean that those sections can be ignored – each section contains important ideas that are used later on and most include something new or have a different emphasis.

Most chapters are divided into five or six sections (each often further divided into sub-sections); sections are numbered using two digits (such as 'Section 3.2') and sub-sections using three digits (such as 'Sub-section 3.2.4'). Generally a section is considered to be about one evening's work for an average student.

Chapter 7 on Integration is arguably the highlight of the book. However, it contains some rather complicated mathematical arguments and proofs.

Therefore, when reading Chapter 7, it is important not to get bogged down in details, but to keep progressing through the key ideas, and to return later on to reading the things that were left out at the earlier reading. Most students will require three or four passes at this chapter before having a good idea of most of it.

We use wide pages with a large number of margin notes in which we place teaching comments and some diagrams to aid in the understanding of particular points in arguments. We also provide advice on which proofs to omit on a first study of the topic; it is important for the student NOT to get bogged down in a technical discussion or a proof until they have a good idea of the message contained in the result and the situations in which it can be used. Therefore clear encouragement is given on which portions of the text to leave till later, or to simply skim on a first reading.

It is important to read the margin notes!

This signposting benefits students greatly in the author's experience.

The end of the proof of a Theorem is indicated by a solid symbol '■' and the end of the solution of a worked Example by a hollow symbol '□'. There are many worked examples within the text to explain the concepts being taught, together with a good stock of problems to reinforce the teaching. The solutions are a key part of the teaching, and tackling them on your own and then reading our version of the solution is a key part of the learning process.

Tackling the problems is a good use of your time, not something to skimp.

No one can learn Mathematics by simply reading – it is a 'hands on' activity. The reader should not be afraid to draw pictures to illustrate what seems to be happening to a sequence or a function, to get a feeling for their behaviour. A wise old man once said that 'A picture is worth a thousand words!'. A good picture may even suggest a method of proof. However, at the same time it is important not to regard a picture on its own as a proof of anything; it may illustrate just one situation that can arise and miss many other possibilities!

It is important NOT to become discouraged if a topic seems difficult. It took mathematicians hundreds of years to develop Analysis to its current polished state, so it may take the reader a few hours at several sittings to really grasp the more complex or subtle ideas.

Acknowledgements

The material in the Open University course on which this book is based was contributed to in some way by many colleagues, including Phil Rippon, Robin Wilson, Andrew Brown, Hossein Zand, Joan Aldous, Ian Harrison, Alan Best, Alison Cadle and Roberta Cheriyan. Its eventual appearance in book form owes much to Lynne Barber.

Without the forebearance of my family, the writing of the book would have been impossible.

1 Numbers

In this book we study the properties of real functions defined on intervals of the real line (possibly the whole real line) and whose image also lies on the real line. In other words, they map $\mathbb{R}$ into $\mathbb{R}$. Our work will be from a very precise point of view in order to establish many of the properties of such functions which seem intuitively obvious; in the process we will discover that some apparently true properties are in fact not necessarily true!

The types of functions that we shall examine include:

exponential functions, such as $x \mapsto a^x$, where $a, x \in \mathbb{R}$,

trigonometric functions, such as $x \mapsto \sin x$, where $x \in \mathbb{R}$,

root functions, such as $x \mapsto \sqrt{x}$, where $x \geq 0$.

The types of behaviour that we shall examine include *continuity*, *differentiability* and *integrability* – and we shall discover that functions with these properties can be used in a number of surprising applications.

However, to put our study of such functions on a secure foundation, we need first to clarify our ideas of the *real numbers* themselves and their properties. In particular, we need to devote some time to the manipulation of *inequalities*, which play a key role throughout the book.

For example, what exactly *is* the number $\sqrt{2}$?

In Section 1.1, we start by revising the properties of rational numbers and their *decimal representation*. Then we introduce the real numbers as infinite decimals, and describe the difficulties involved in doing arithmetic with such decimals.

In Section 1.2, we revise the rules for manipulating inequalities and show how to find the *solution set* of an inequality involving a real number, x, by applying the rules. We also explain how to deal with inequalities which involve *modulus* signs.

In Section 1.3, we describe various techniques for proving inequalities, including the very important technique of *Mathematical Induction*.

The concept of a *least upper bound*, which is of great importance in Analysis, is introduced in Section 1.4, and we discuss the *Least Upper Bound Property* of $\mathbb{R}$.

Finally, in Section 1.5, we describe how least upper bounds can be used to define arithmetical operations in $\mathbb{R}$.

You may omit this section at a first reading.

Even though you may be familiar with much of this material we recommend that you read through it, as we give the system of real numbers a more careful treatment than you may have met before. The material on inequalities and least upper bounds is particularly important for later on.

In later chapters we shall define exactly what the numbers π and e are, and find various ways of calculating them. But, first, we examine numbers in general.

1.1 Real numbers

We start our study of the real numbers with the rational numbers, and investigate their decimal representations, then we proceed to the irrational numbers.

1.1.1 Rational numbers

We assume that you are familiar with the set of **natural numbers**

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

Note that 0 is *not* a natural number.

and with the set of **integers**

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

The set of **rational numbers** consists of all fractions (or ratios of integers)

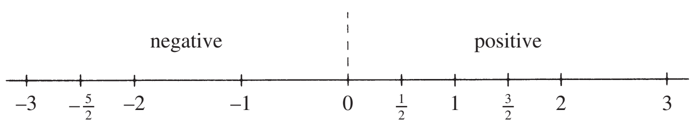
$$\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\right\}.$$

Remember that each rational number has many different representations as a ratio of integers; for example

$$\frac{1}{3} = \frac{2}{6} = \frac{10}{30} = \ldots.$$

We also assume that you are familiar with the usual arithmetical operations of addition, subtraction, multiplication and division of rational numbers.

It is often convenient to represent rational numbers geometrically as points on a **number line**. We begin by drawing a line and marking on it points corresponding to the integers 0 and 1. If the distance between 0 and 1 is taken as a unit of length, then the rationals can be arranged on the line with positive rationals to the right of 0 and negative rationals to the left.



For example, the rational $\frac{3}{2}$ is placed at the point which is one-half of the way from 0 to 3.

This means that rationals have a natural order on the number-line. For example, $\frac{19}{22}$ lies to the left of $\frac{7}{8}$ because

$$\frac{19}{22} = \frac{76}{88} \quad \text{and} \quad \frac{7}{8} = \frac{77}{88}.$$

If a lies to the left of b on the number-line, then

a is *less than* b or b is *greater than* a,

and we write

$$a < b \quad \text{or} \quad b > a.$$

For example

$$\frac{19}{22} < \frac{7}{8} \quad \text{or} \quad \frac{7}{8} > \frac{19}{22}.$$

We write $a \le b$, or $b \ge a$, if either $a < b$ or $a = b$.

Problem 1 Arrange the following rational numbers in order:

$$0, 1, -1, \tfrac{17}{20}, -\tfrac{17}{20}, \tfrac{45}{53}, -\tfrac{45}{53}.$$

Problem 2 Show that between any two distinct rational numbers there is another rational number.

1.1.2 Decimal representation of rational numbers

The decimal system enables us to represent all the natural numbers using only the ten integers

$$0, 1, 2, 3, 4, 5, 6, 7, 8 \text{ and } 9,$$

which are called *digits*. We now remind you of the basic facts about the representation of *rational* numbers by decimals.

> **Definition** A **decimal** is an expression of the form
>
> $$\pm a_0 \cdot a_1 a_2 a_3 \ldots,$$
>
> where a_0 is a non-negative integer and $a_1, a_2, a_3, \ldots$ are digits.

For example
$0.8500\ldots,$
$13.1212\ldots,$
$-1.111\ldots.$

If only a finite number of the digits $a_1, a_2, a_3, \ldots$ are non-zero, then the decimal is called **terminating** or **finite**, and we usually omit the tail of 0s.

For example, $0.8500\ldots = 0.85.$

Terminating decimals are used to represent rational numbers in the following way

$$\pm a_0 \cdot a_1 a_2 a_3 \ldots a_n = \pm\left(a_0 + \frac{a_1}{10^1} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n}\right).$$

For example
$$0.85 = 0 + \frac{8}{10^1} + \frac{5}{10^2} = \frac{85}{100} = \frac{17}{20}.$$

It can be shown that any fraction whose denominator contains only powers of 2 and/or 5 (such as $20 = 2^2 \times 5$) can be represented by such a terminating decimal, which can be found by long division.

However, if we apply long division to many other rationals, then the process of long division never terminates and we obtain a **non-terminating** or **infinite** decimal. For example, applying long division to $\frac{1}{3}$ gives $0.333\ldots$, and for $\frac{19}{22}$ we obtain $0.86363\ldots$.

Problem 3 Apply long division to $\frac{1}{7}$ and $\frac{2}{13}$ to find the corresponding decimals.

These non-terminating decimals, which are obtained by applying the long division process, have a certain common property. All of them are **recurring**; that is, they have a recurring block of digits, and so can be written in shorthand form, as follows:

$$\begin{aligned} 0.333\ldots &= 0.\overline{3}, \\ 0.142857142857\ldots &= 0.\overline{142857}\ldots, \\ 0.86363\ldots &= 0.8\overline{63}. \end{aligned}$$

Another commonly used notation is
$0.\dot{3}$ or $0.\dot{1}4285\dot{7}$.

It is not hard to show, whenever we apply the long division process to a fraction $\frac{p}{q}$, that the resulting decimal is recurring. To see why we notice that there are only q possible remainders at each stage of the division, and so one of these remainders must eventually recur. If the remainder 0 occurs, then the resulting decimal is, of course, terminating; that is, it ends in recurring 0s.

Non-terminating recurring decimals which arise from the long division of fractions are used to represent the corresponding rational numbers. This representation is not quite so straight-forward as for terminating decimals, however. For example, the statement

$$\frac{1}{3} = 0.\overline{3} = \frac{3}{10^1} + \frac{3}{10^2} + \frac{3}{10^3} + \cdots$$

can be made precise only when we have introduced the idea of the *sum of a convergent infinite series*. For the moment, when we write the statement $\frac{1}{3} = 0.\overline{3}$ we mean simply that the decimal $0.\overline{3}$ arises from $\frac{1}{3}$ by the long division process.

We return to this topic in Chapter 3.

The following example illustrates one way of finding the rational number with a given decimal representation.

Example 1 Find the rational number (expressed as a fraction) whose decimal representation is $0.8\overline{63}$.

Solution First we find the fraction x such that $x = 0.\overline{63}$.

If we multiply both sides of this equation by 10^2 (because the recurring block has length 2), then we obtain

$$100x = 63.\overline{63} = 63 + x.$$

Hence

$$99x = 63 \Rightarrow x = \frac{63}{99} = \frac{7}{11}.$$

Thus

$$0.8\overline{63} = \frac{8}{10} + \frac{x}{10} = \frac{8}{10} + \frac{7}{110} = \frac{95}{110} = \frac{19}{22}.$$ □

The key idea in the above solution is that multiplication of a decimal by 10^k moves the decimal point k places to the right.

Problem 4 Using the above method, find the fractions whose decimal representations are:

(a) $0.\overline{231}$; (b) $2.2\overline{81}$.

The decimal representation of rational numbers has the advantage that it enables us to decide immediately which of two distinct positive rational numbers is the greater. We need only examine their decimal representations and notice the first place at which the digits differ. For example, to order $\frac{7}{8}$ and $\frac{19}{22}$ we write

$$\frac{7}{8} = 0.875\ldots \quad \text{and} \quad \frac{19}{22} = 0.86363\ldots,$$

and so

$$0.8\overset{\downarrow}{6}363\ldots < 0.8\overset{\downarrow}{7}5 \Rightarrow \frac{19}{22} < \frac{7}{8}.$$

Problem 5 Find the first two digits after the decimal point in the decimal representations of $\frac{17}{20}$ and $\frac{45}{53}$, and hence determine which of these two rational numbers is the greater.

Warning Decimals which end in recurring 9s sometimes arise as alternative representations for terminating decimals. For example

$$1 = 0.\overline{9} = 0.999\ldots \quad \text{and} \quad 1.35 = 1.34\overline{9} = 1.34999\ldots.$$

Whenever possible, we avoid using the form of a decimal which ends in recurring 9s.

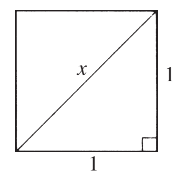
You may find this rather alarming, but it is important to realise that this is a matter of *convention*. We wish to allow the decimal 0.999... to represent a number x, so x must be less than or equal to 1 and greater than each of the numbers

$$0.9, 0.99, 0.999, \ldots.$$

The *only* rational with these properties is 1.

1.1.3 Irrational numbers

One of the surprising mathematical discoveries made by the Ancient Greeks was that the system of rational numbers is not adequate to describe all the magnitudes that occur in geometry. For example, consider the diagonal of a square of side 1. What is its length? If the length is x, then, by Pythagoras' Theorem, x must satisfy the equation $x^2 = 2$. However, there is no rational number which satisfies this equation.



$$x^2 = 1^2 + 1^2 = 2$$

Theorem 1 There is no rational number x such that $x^2 = 2$.

Proof Suppose that such a rational number x exists. Then we can write $x = \frac{p}{q}$. By cancelling, if necessary, we may assume that p and q have no common factor. The equation $x^2 = 2$ now becomes

This is a proof by contradiction.

$$\frac{p^2}{q^2} = 2, \quad \text{so} \quad p^2 = 2q^2.$$

Now, the square of an odd number is odd, and so p cannot be odd. Hence p is even, and so we can write $p = 2r$, say. Our equation now becomes

For

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

$$(2r)^2 = 2q^2, \quad \text{so} \quad q^2 = 2r^2.$$

Reasoning as before, we find that q is also even.

Since p and q are both even, they have a common factor 2, which contradicts our earlier statement that p and q have no common factors.

Arguing from our original assumption that x exists, we have obtained two contradictory statements. Thus, our original assumption must be false. In other words, no such x exists. ■

Problem 6 By imitating the above proof, show that there is no rational number x such that $x^3 = 2$.

Since we want equations such as $x^2 = 2$ and $x^3 = 2$ to have solutions, we must introduce new numbers which are not rational numbers. We denote these new numbers by $\sqrt{2}$ and $\sqrt[3]{2}$, respectively; thus $\left(\sqrt{2}\right)^2 = 2$ and $\left(\sqrt[3]{2}\right)^3 = 2$. Of course, we must introduce many other new numbers, such as $\sqrt{3}, \sqrt[5]{11}$, and so on. Indeed, it can be shown that, if m, n are natural numbers and $x^m = n$ has no integer solution, then $\sqrt[m]{n}$ cannot be rational. A number which is not rational is called **irrational**.

The case $m = 2$ is treated in Exercise 5 for this section in Section 1.6.

There are many other mathematical quantities which cannot be described exactly by rational numbers. For example, the number π which denotes the area of a disc of radius 1 (or half the length of the perimeter of such a disc) is irrational, as is the number e.

Lambert proved that π is irrational in 1770.

It is natural to ask whether irrational numbers, such as $\sqrt{2}$ and π, can be represented as decimals. Using your calculator, you can check that $(1.41421356)^2$ is very close to 2, and so 1.41421356 is a very good approximate value for $\sqrt{2}$. But is there a decimal which represents $\sqrt{2}$ *exactly*? If such a decimal exists, then it cannot be recurring, because all the recurring decimals correspond to rational numbers.

In fact
$(1.41421356)^2$
$= 1.9999999932878736.$

In fact, it *is* possible to represent all the irrational numbers mentioned so far by non-recurring decimals. For example, there are non-recurring decimals such that

$$\sqrt{2} = 1.41421356\ldots \quad \text{and} \quad \pi = 3.14159265\ldots.$$

We prove that $\sqrt{2}$ has a decimal representation in Section 1.5.

It is also natural to ask whether non-recurring decimals, such as

$$0.101001000100001\ldots \quad \text{and} \quad 0.123456789101112\ldots,$$

represent irrational numbers. In fact, a decimal corresponds to a rational number if and only if it is recurring; so a non-recurring decimal must correspond to an irrational number.

We may summarise this as:

recurring decimal $\Leftrightarrow$ *rational number*

non-recurring decimal $\Leftrightarrow$ *irrational number*

1.1.4 The real number system

Taken together, the rational numbers (recurring decimals) and irrational numbers (non-recurring decimals) form the set of **real numbers**, denoted by $\mathbb{R}$.

As with rational numbers, we can determine which of two real numbers is greater by comparing their decimals and noticing the first pair of corresponding digits which differ. For example

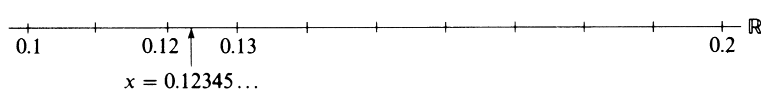
When comparing decimals in this way, we do not allow either decimal to end in recurring 9s.

$$0.1\overset{\downarrow}{0}100100010000\ldots < 0.1\overset{\downarrow}{2}3456789101112\ \ldots.$$

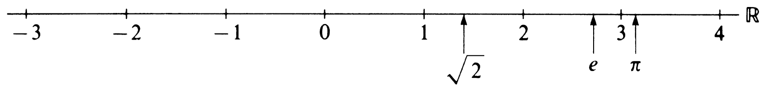
We now associate with each irrational number a point on the number-line. For example, the irrational number $x = 0.123456789101112\ldots$ satisfies each of the inequalities

$$\begin{gathered} 0.1 < x < 0.2 \\ 0.12 < x < 0.13 \\ 0.123 < x < 0.124 \\ \vdots \end{gathered}$$

We assume that there is a point on the number-line corresponding to x, which lies to the right of each of the (rational) numbers 0.1, 0.12, 0.123 ..., and to the left of each of the (rational) numbers 0.2, 0.13, 0.124, ...



As usual, negative real numbers correspond to points lying to the left of 0; and the 'number-line', complete with both rational and irrational points, is called the **real line**.



There is thus a one–one correspondence between the points on the real line and the set $\mathbb{R}$ of real numbers (or decimals).

We now state several properties of $\mathbb{R}$, with which you will be already familiar, although you may not have met their names before. These properties are used frequently in Analysis, and we do not always refer to them explicitly by name.

Order Properties of $\mathbb{R}$

1. **Trichotomy Property** If $a, b \in \mathbb{R}$, then *exactly one* of the following inequalities holds
$$a < b \quad \text{or} \quad a = b \quad \text{or} \quad a > b.$$
2. **Transitive Property** If $a, b, c \in \mathbb{R}$, then
$$a < b \quad \text{and} \quad b < c \Rightarrow a < c.$$
3. **Archimedean Property** If $a \in \mathbb{R}$, then there is a positive integer n such that
$$n > a.$$
4. **Density Property** If $a, b \in \mathbb{R}$ and $a < b$, then there is a rational number x and an irrational number y such that
$$a < x < b \quad \text{and} \quad a < y < b.$$

The first three of these properties are almost self-evident, but the Density Property is not so obvious.

Remark

The Archimedean Property is sometimes expressed in the following equivalent way: for any positive real number a, there is a positive integer n such that $\frac{1}{n} < a$.

The following example illustrates how we can prove the Density Property.

Example 2 Find a rational number x and an irrational number y satisfying
$$a < x < b \quad \text{and} \quad a < y < b,$$
where $a = 0.12\overline{3}$ and $b = 0.12345\ldots$.

Solution The two decimals
$$a = 0.123\overset{\downarrow}{3}\ldots \quad \text{and} \quad b = 0.123\overset{\downarrow}{4}5\ldots$$
differ first at the fourth digit. If we truncate b after this digit, we obtain the rational number $x = 0.1234$, which satisfies the requirement that $a < x < b$.

To find an irrational number y between a and b, we attach to x a (sufficiently small) non-recurring tail such as $010010001\ldots$ to give $y = 0.1234|010010001\ldots$. It is then clear that y is irrational (because its decimal is non-recurring) and that $a < y < b$. □

Problem 7 Find a rational number x and an irrational number y such that $a < x < b$ and $a < y < b$, where $a = 0.\overline{3}$ and $b = 0.3401$.

Theorem 2 Density Property of $\mathbb{R}$
If $a, b \in \mathbb{R}$ and $a < b$, then there is a rational number x and an irrational number y such that

$$a < x < b \quad \text{and} \quad a < y < b.$$

You may omit the following proof on a first reading.

Proof For simplicity, we assume that $a, b \geq 0$. So, let a and b have decimal representations

$$a = a_0 \cdot a_1 a_2 a_3 \ldots \quad \text{and} \quad b = b_0 \cdot b_1 b_2 b_3 \ldots,$$

Here a_0, b_0 are non-negative integers, and $a_1, b_1, a_2, b_2, \ldots$ are digits.

where we arrange that a does not end in recurring 9s, whereas b does not terminate (this latter can be arranged by replacing a terminating representation by an equivalent representation that ends in recurring 9s).

Since $a < b$, there must be some integer n such that

$$a_0 = b_0,\ a_1 = b_1,\ \ldots,\ a_{n-1} = b_{n-1},\ \text{but } a_n < b_n.$$

Then $x = a_0 \cdot a_1 a_2 a_3 \ldots a_{n-1} b_n$ is rational, and $a < x < b$ as required.

Finally, since $x < b$, it follows that we can attach a sufficiently small non-recurring tail to x to obtain an irrational number y for which $a < y < b$. ■

Remark

One consequence of the Density Property is that between any two real numbers there are infinitely many rational numbers and infinitely many irrational numbers.

Problem 8 Prove that between any two real numbers a and b there are at least two distinct rational numbers.

A proof of the previous remark would involve ideas similar to those involved in tackling Problem 8.

1.1.5 Arithmetic in $\mathbb{R}$

We can do arithmetic with *recurring* decimals by first converting the decimals to fractions. However, it is not obvious how to perform arithmetical operations with *non-recurring* decimals.

Assuming that we can represent $\sqrt{2}$ and π by the non-recurring decimals

$$\sqrt{2} = 1.41421356\ldots \quad \text{and} \quad \pi = 3.14159265\ldots,$$

can we also represent the sum $\sqrt{2} + \pi$ and the product $\sqrt{2} \times \pi$ as decimals? Indeed, what do we mean by the operations of addition and multiplication when non-recurring decimals (irrationals) are involved, and do these operations satisfy the same properties as addition and multiplication of rationals?

It would take many pages to answer these questions fully. Therefore, we shall *assume* that it is possible to define all the usual arithmetical operations with decimals, and that they do satisfy the usual properties. For definiteness, we now list these properties.

Arithmetic in $\mathbb{R}$

Addition	*Multiplication*	
A1 If $a, b \in \mathbb{R}$, then $a + b \in \mathbb{R}$.	M1 If $a, b \in \mathbb{R}$, then $a \times b \in \mathbb{R}$.	CLOSURE
A2 If $a \in \mathbb{R}$, then $a + 0 = 0 + a = a$.	M2 If $a \in \mathbb{R}$, then $a \times 1 = 1 \times a = a$.	IDENTITY
A3 If $a \in \mathbb{R}$, then there is a number $-a \in \mathbb{R}$ such that $a + (-a) = (-a) + a = 0$.	M3 If $a \in \mathbb{R} - \{0\}$, then there is a number $a^{-1} \in \mathbb{R}$ such that $a \times a^{-1} = a^{-1} \times a = 1$.	INVERSES
A4 If $a, b, c \in \mathbb{R}$, then $(a + b) + c = a + (b + c)$.	M4 If $a, b, c \in \mathbb{R}$, then $(a \times b) \times c = a \times (b \times c)$.	ASSOCIATIVITY
A5 If $a, b \in \mathbb{R}$, then $a + b = b + a$.	M5 If $a, b \in \mathbb{R}$, then $a \times b = b \times a$.	COMMUTATIVITY
D If $a, b, c \in \mathbb{R}$, then $a \times (b + c) = a \times b + a \times c$.		DISTRIBUTATIVITY

To summarise the contents of this table:

- $\mathbb{R}$ is an Abelian group under the operation of addition $+$; (Properties A1–A5)
- $\mathbb{R} - \{0\}$ is an Abelian group under the operation of multiplication $\times$; (Properties M1–M5)
- These two group structures are linked by the Distributive Property. (Property D)

It follows from the above properties that we can perform addition, subtraction (where $a - b = a + (-b)$), multiplication and division (where $\frac{a}{b} = a \times b^{-1}$) in $\mathbb{R}$, and that these operations satisfy all the usual properties.

Any system satisfying the properties listed in the table is called a **field**. Both $\mathbb{Q}$ and $\mathbb{R}$ are fields.

Furthermore, we shall assume that the set $\mathbb{R}$ contains the nth roots and rational powers of positive real numbers, with their usual properties. In Section 1.5 we describe one way of justifying the existence of nth roots.

1.2 Inequalities

Much of Analysis is concerned with inequalities of various kinds; the aim of this section and the next section is to provide practice in the manipulation of inequalities.

1.2.1 Rearranging inequalities

The fundamental rule, upon which much manipulation of inequalities is based, is that the statement $a < b$ means exactly the same as the statement $b - a > 0$. This fact can be stated concisely in the following way:

Recall that the symbol '$\Leftrightarrow$' means 'if and only if' or 'implies and is implied by'.

Rule 1 For any $a, b \in \mathbb{R}$, $a < b \Leftrightarrow b - a > 0$.

Put another way, the inequalities $a < b$ and $b - a > 0$ are *equivalent*.

There are several other standard rules for rearranging a given inequality into an equivalent form. Each of these can be deduced from our first rule above. For example, we obtain an equivalent inequality by adding the same expression to both sides.

Rule 2 For any $a, b, c \in \mathbb{R}$, $a < b \Leftrightarrow a + c < b + c$.

Another way to rearrange an inequality is to multiply both sides by a non-zero expression, making sure to *reverse* the inequality if the expression is negative.

Rule 3
- For any $a, b \in \mathbb{R}$ and any $c > 0$, $a < b \Leftrightarrow ac < bc$;
- For any $a, b \in \mathbb{R}$ and any $c < 0$, $a < b \Leftrightarrow ac > bc$.

For example

$2 < 3 \Leftrightarrow 20 < 30\ (c = 10)$,

$2 < 3 \Leftrightarrow -20 > -30$

$(c = -10)$.

Sometimes the most effective way to rearrange an inequality is to take reciprocals. However, in this case, both sides of the inequality should be positive, and the direction of the inequality has to be *reversed*.

Rule 4 (Reciprocal Rule)
For any positive $a, b \in \mathbb{R}$, $a < b \Leftrightarrow \frac{1}{a} > \frac{1}{b}$.

For example

$$2 < 4 \Leftrightarrow \frac{1}{2}(=0.5) > \frac{1}{4}(=0.25).$$

Some inequalities can be simplified only by taking powers. However, in order to do this, both sides must be non-negative and must be raised to a *positive* power.

Rule 5 (Power Rule)
For any non-negative $a, b \in \mathbb{R}$, and any $p > 0$, $a < b \Leftrightarrow a^p < b^p$.

For example

$$4 < 9 \Leftrightarrow 4^{\frac{1}{2}}(=2) < 9^{\frac{1}{2}}(=3).$$

We shall discuss the meaning of non-integer powers in Section 1.5.

For positive integers p, Rule 5 follows from the identity

$$b^p - a^p = (b - a)\left(b^{p-1} + b^{p-2}a + \cdots + ba^{p-2} + a^{p-1}\right);$$

thus, since the right-hand bracket is positive, we have

$$b - a > 0 \Leftrightarrow b^p - a^p > 0,$$

which is equivalent to our desired result.

For example,

$$b^3 - a^3 = (b - a)\times \left(b^2 + ba + a^2\right).$$

Remark

There are corresponding versions of Rules 1–5 in which the *strict* inequality $a < b$ is replaced by the *weak* inequality $a \leq b$.

Problem 1 State (without proof) the versions of Rules 1–5 for weak inequalities.

We shall give one more rule for rearranging inequalities in Sub-section 1.2.3.

1.2.2 Solving inequalities

Solving an inequality involving an unknown real number x means determining those values of x for which the given inequality holds; that is, finding the *solution set* of the inequality. We can often do this by rewriting the inequality in an equivalent, but simpler form, using the rules given in the last sub-section.

The *solution set* is the set of those values of x for which the inequality holds.

In this activity we frequently use the usual rules for the sign of a product, and the fact that the square of any real number is non-negative. Also, we need to remember the difference between the logical statements: 'implies', 'is implied by' and 'implies and is implied by'.

$\times$	$+$	$-$
$+$	$+$	$-$
$-$	$-$	$+$

Example 1 Solve the inequality $\frac{x+2}{x+4} > \frac{x-3}{2x-1}$.

Solution We rearrange this inequality to give a somewhat simpler inequality, using Rule 1

$$\begin{aligned}\frac{x+2}{x+4} > \frac{x-3}{2x-1} &\Leftrightarrow \frac{x+2}{x+4} - \frac{x-3}{2x-1} > 0\\ &\Leftrightarrow \frac{x^2+2x+10}{(x+4)(2x-1)} > 0\\ &\Leftrightarrow \frac{(x+1)^2+9}{(x+4)(2x-1)} > 0.\end{aligned}$$

It is a common strategy to bring all terms to one side.

We bring everything to a common denominator.

Here we complete the square in the numerator, since we cannot factorise it.

Now, the numerator is always positive. The denominator vanishes when $x=-4$ or $x=\frac{1}{2}$. By examining separately the sign of the denominator when $x<-4$, $-4<x<\frac{1}{2}$ and $x>\frac{1}{2}$, we can deduce that the last fraction is positive precisely when $x<-4$ or $x>\frac{1}{2}$. Hence the solution set of the original inequality is

$$\left\{x : \frac{x+2}{x+4} > \frac{x-3}{2x-1}\right\} = (-\infty,-4)\cup\left(\tfrac{1}{2},\infty\right). \qquad \square$$

This is because the final displayed inequality is *equivalent* to the inequality we are solving. The logical implication symbols between the displayed inequalities were all 'implies and is implied by'.

Example 2 Solve the inequality $\frac{1}{2x^2+2} < \frac{1}{4}$.

Solution Since $2x^2+2>0$, we have

$$\begin{aligned}\frac{1}{2x^2+2} < \frac{1}{4} &\Leftrightarrow 2x^2+2 > 4 && \text{(by Rule 4)}\\ &\Leftrightarrow x^2+1 > 2 && \text{(by Rule 3)}\\ &\Leftrightarrow x^2-1 > 0 && \text{(by Rule 1)}\\ &\Leftrightarrow (x-1)(x+1) > 0.\end{aligned}$$

Here we factorise the left-hand side of the inequality to examine the signs of its factors.

This last inequality holds precisely when $x<-1$ or $x>1$. It follows that the solution set of the original inequality is

$$\left\{x : \frac{1}{2x^2+2} < \frac{1}{4}\right\} = (-\infty,-1)\cup(1,\infty). \qquad \square$$

Problem 2 Use each of the following expressions to write down an inequality with the given expression on its left-hand side which is equivalent to the inequality $x>2$:

(a) $x+3$; (b) $2-x$; (c) $5x+2$; (d) $\frac{-1}{5x+2}$.

Problem 3 Solve the following inequalities:

(a) $\frac{4x-x^2-7}{x^2-1} \geq 3$; (b) $2x^2 \geq (x+1)^2$.

Warning Great care is needed when solving inequalities which involve rational powers. In particular, when applying Rule 5 both sides of the inequality *must* be non-negative.

Example 3 Solve the inequality $\sqrt{2x+3} > x$.

Solution The expression $\sqrt{2x+3}$ is defined only when $2x+3 \geq 0$; that is, when $x \geq -\frac{3}{2}$. Hence we need only consider those x in $\left[\frac{-3}{2}, \infty\right)$.

We can obtain an equivalent inequality by squaring, provided that both $\sqrt{2x+3}$ and x are non-negative. Thus, for $x \geq 0$, we obtain

Note that, for the moment, we are examining only those x for which $x \geq 0$.

$$\begin{aligned}\sqrt{2x+3} > x &\Leftrightarrow 2x+3 > x^2 \qquad \text{(by Rule 5, with } p=2)\\ &\Leftrightarrow x^2 - 2x - 3 < 0\\ &\Leftrightarrow (x-3)(x+1) < 0.\end{aligned}$$

So the part of the solution set in $[0, \infty)$ is $[0, 3)$.

We now examine those x for which $-\frac{3}{2} \leq x < 0$. For such x, $\sqrt{2x+3} \geq 0$ and $x < 0$, so that $\sqrt{2x+3}\ (\geq 0)\ > x$, for all such x. It follows that all these x, namely the set $[-\frac{3}{2}, 0)$, belong to the solution set too.

Notice the use of the Transitive Property here.

Combining these results, the solution set of the original inequality is

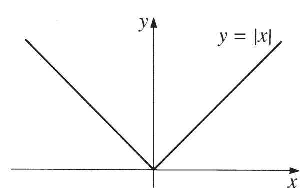
$$\begin{aligned}\left\{x\colon \sqrt{2x+3} > x\right\} &= \left[-\frac{3}{2}, 0\right) \cup [0,3)\\ &= \left[-\frac{3}{2}, 3\right).\end{aligned}$$ □

Problem 4 Solve the inequality $\sqrt{2x^2-2} > x$.

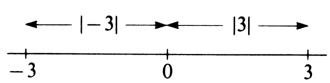
1.2.3 Inequalities involving modulus signs

We now turn our attention to inequalities involving the *modulus*, or *absolute value*, of a real number. Recall that, if $a \in \mathbb{R}$, then its modulus $|a|$ is defined by

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



It is often useful to think of $|a|$ as the distance along the real line from 0 to a.



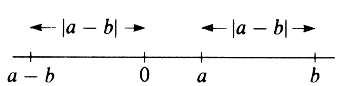
For example

$$|3| = |-3| = 3.$$

In the same way, $|a-b|$ is the distance along the real line from 0 to $a-b$, which is the same as the distance from a to b.

We sometimes write

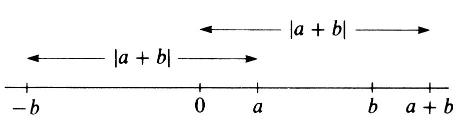
$$|a-b| = d(a,b).$$



Notice also that $|a+b| = |a-(-b)|$ is the distance from a to $-b$.

For example, the distance from -2 to 3 is

$$|(-2)-3| = |-5| = 5.$$



We now list some basic properties of the modulus, which follow immediately from the definition:

Properties of the modulus For any real numbers a and b:

1. $|a| \geq 0$, with equality if and only if $a = 0$;
2. $-|a| \leq a \leq |a|$;
3. $|a|^2 = a^2$;
4. $|a - b| = |b - a|$;
5. $|ab| = |a| \times |b|$.

For example, if $a = -2$, $b = 1$, then:

1. $|-2| > 0$;
2. $-|-2| \leq -2 \leq |-2|$;
3. $|-2|^2 = (-2)^2$;
4. $|(-2) - 1| = |1 - (-2)|$;
5. $|(-2) \times 1| = |-2| \times |1|$.

There is a basic rule for rearranging inequalities involving modulus signs:

Rule 6 For any real numbers a and b, where $b > 0$: $|a| < b \Leftrightarrow -b < a < b$.

Note that, in a similar way

$$|a| \leq b \Leftrightarrow -b \leq a \leq b.$$

Also, it is often possible, and sometimes easier, to use Rule 5 with $p = 2$ than to use Rule 6. The following example illustrates the use of both rules.

Example 4 Solve the inequality $|x - 2| < 1$.

Solution Using Rule 6, we obtain

$$\begin{aligned} |x-2| < 1 &\Leftrightarrow -1 < x - 2 < 1 \\ &\Leftrightarrow 1 < x < 3. \end{aligned}$$

We take $a = x - 2$, $b = 1$ in Rule 6.

So the solution set of the original inequality is

$$\{x: |x-2| < 1\} = (1, 3).$$

Alternatively, using Rule 5 (with $p = 2$), we obtain

$$\begin{aligned} |x-2| < 1 &\Leftrightarrow (x-2)^2 < 1 \\ &\Leftrightarrow x^2 - 4x + 3 < 0 \\ &\Leftrightarrow (x-1)(x-3) < 0. \end{aligned}$$

For $|x - 2|^2 = (x - 2)^2$.

Again, this shows that the required solution set is (1, 3). □

Example 5 Solve the inequality $|x - 2| \leq |x + 1|$.

Solution Using Rule 5 (with $p = 2$), we obtain

$$\begin{aligned} |x-2| \leq |x+1| &\Leftrightarrow (x-2)^2 \leq (x+1)^2 \\ &\Leftrightarrow x^2 - 4x + 4 \leq x^2 + 2x + 1 \\ &\Leftrightarrow 3 \leq 6x \\ &\Leftrightarrow \frac{1}{2} \leq x. \end{aligned}$$

So the solution set of the original inequality is

$$\{x: |x-2| \leq |x+1|\} = \left[\frac{1}{2}, \infty\right).$$ □

The inequalities in Examples 4 and 5 can easily be interpreted geometrically.

In Example 4, the inequality $|x - 2| < 1$ holds when the distance from x to 2 is strictly less than 1. So it holds for all points on either side of 2 at a distance less than 1 from 2 – namely, in the open interval (1, 3).

In Example 5, the inequality $|x-2| \leq |x+1|$ holds when the distance from x to 2 is less than or equal to the distance from x to -1, since $|x+1| = |x-(-1)|$. The mid-point of 2 and -1 (that is, the point x where the distance from x to 2 equals the distance from x to -1) is $\frac{1}{2}$. So the inequality holds when x lies in $[\frac{1}{2}, \infty)$.

Some good ideas when tackling problems involving inequalities of these types are:

- use your geometrical intuition, where possible, to give yourself an idea of the sets involved;
- test one or two values of x in your final solution set to see if they are valid – this often detects errors in manipulating inequality signs!

Problem 5 Solve the following inequalities:

(a) $|2x^2 - 13| < 5$; (b) $|x-1| \leq 2|x+1|$.

1.3 Proving inequalities

In this section we show you how to *prove* inequalities of various types. We shall use the rules for rearranging inequalities given in Section 1.2, and also use other rules which enable us to *deduce new inequalities from old*. We have already met the first rule in Section 1.1, where it was called the Transitive Property of $\mathbb{R}$.

Transitive Rule $a < b$ and $b < c \Rightarrow a < c$.

We use the Transitive Rule when we want to prove that $a < c$, and we know that $a < b$ and $b < c$.

The following rules are also useful:

Combination Rules

If $a < b$ and $c < d$, then:

Sum Rule $a + c < b + d$;

Product Rule $ac < bd$ (provided that $a, c \geq 0$).

For example, since $2 < 3$ and $4 < 5$, then

$$2 + 4 < 3 + 5;$$
$$2 \times 4 < 3 \times 5.$$

There are also weak and weak/strict versions of the Transitive Rule and Combination Rules, which we will ask you to work out and use as they arise.

For example, if $a < b$ and $c \leq d$, then

$$a + c < b + d$$

and

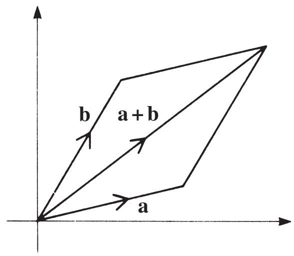
$$ac < bd, \text{provided } a, c > 0.$$

Remark

It is important to appreciate that the Transitive Rule and the Combination Rules have a different nature from Rules 1–6 in Section 1.2. Rules 1–6 tell you how to rearrange inequalities into equivalent forms, whereas the Transitive Rule and the Combination Rules enable you to deduce new inequalities which are *not* equivalent to the old ones.

1.3.1 The Triangle Inequality

If **a** and **b** are both vectors in $\mathbb{R}^2$, then the vector $\mathbf{a}+\mathbf{b}$ is obtained from the 'parallelogram construction':



Here we discuss $\mathbb{R}^2$, rather than $\mathbb{R}$, simply because the argument is then geometrically clearer.

In the 'parallelogram construction', you draw the vector **a** from the origin to some point, then the vector **b** from that point to a final point. The vector $\mathbf{a}+\mathbf{b}$ is then the vector from the origin to that final point.

By elementary geometry, the length of any side of a triangle is less than or equal to the sum of the lengths of the other two sides. In the special case of $\mathbb{R}$, when all the vectors lie on a line, this can be interpreted as the *Triangle Inequality*, which involves the absolute value of the real numbers a, b and $a+b$.

> **Triangle Inequality** If $a, b \in \mathbb{R}$, then:
> 1. $|a+b| \leq |a|+|b|$;
> 2. $|a-b| \geq \big||a|-|b|\big|$ (the 'reverse form' of the Triangle Inequality).

For example, with $a=-1$ and $b=3$:
1. $|-1+3| \leq |-1|+|3|$;
2. $|(-1)-3| \geq \big|\,|-1|-|3|\,\big|$.

Proof In order to prove part 1, we use Rule 5, with $p=2$

$$\begin{aligned}|a+b| \leq |a|+|b| &\Leftrightarrow (a+b)^2 \leq (|a|+|b|)^2 \\ &\Leftrightarrow a^2+2ab+b^2 \leq a^2+2|a||b|+b^2 \\ &\Leftrightarrow 2ab \leq 2|a||b|.\end{aligned}$$

Remember that $|a|^2 = a^2$.

The final inequality is certainly true for all $a, b \in \mathbb{R}$, and so the first inequality must also be true for all $a, b \in \mathbb{R}$. Hence we have proved part 1.

We prove part 2 by using the same method

$$\begin{aligned}|a-b| \geq \big||a|-|b|\big| &\Leftrightarrow (a-b)^2 \geq (|a|-|b|)^2 \\ &\Leftrightarrow a^2-2ab+b^2 \geq a^2-2|a||b|+b^2 \\ &\Leftrightarrow -2ab \geq -2|a||b| \\ &\Leftrightarrow 2ab \leq 2|a||b|,\end{aligned}$$

which is again true for all $a, b \in \mathbb{R}$. ■

Part 2 is sometimes called the 'cunning form' or 'backwards form' of the Triangle Inequality.

Remarks

1. Although we have used double-headed arrows here, the actual proof requires only the arrows going *from right to left*. For example, in the proof of part 1 the important implication is
$$|a+b| \leq |a|+|b| \Leftarrow 2ab \leq 2|a||b|.$$
2. Part 1 of the Triangle Inequality can also be proved by using Rule 6, as follows. By Rule 6
$$|a+b| \leq |a|+|b| \Leftrightarrow -(|a|+|b|) \leq a+b \leq (|a|+|b|). \quad (1)$$

Now, we know that $-|a| \le a \le |a|$ and $-|b| \le b \le |b|$; so, by the Sum Rule

$$-(|a| + |b|) \le a + b \le (|a| + |b|).$$

It follows that the left-hand inequality in (1) must also hold, as required.

3. An obvious modification of the proof in Remark 2 shows that the following more general form of the Triangle Inequality also holds:

Triangle Inequality for *n* terms For any real numbers $a_1, a_2, \ldots, a_n$

$$|a_1 + a_2 + \cdots + a_n| \le |a_1| + |a_2| + \cdots + |a_n|.$$

The following example is a typical application of the Triangle Inequality.

Example 1 Use the Triangle Inequality to prove that:

(a) $|a| \le 1 \Rightarrow |3 + a^3| \le 4$; (b) $|b| < 1 \Rightarrow |3 - b| > 2$.

Solution

(a) Suppose that $|a| \le 1$. The Triangle Inequality then gives

$$\begin{aligned} \left|3 + a^3\right| &\le |3| + \left|a^3\right| \\ &= 3 + |a|^3 \\ &\le 3 + 1 \qquad (\text{since } |a| \le 1) \\ &= 4. \end{aligned}$$

Note the use of the Transitive Rule here.

(b) Suppose that $|b| < 1$. The 'reverse form' of the Triangle Inequality then gives

$$\begin{aligned} |3 - b| &\ge \big||3| - |b|\big| \\ &= \big|3 - |b|\big| \\ &\ge 3 - |b|. \end{aligned}$$

Now $|b| < 1$, so that $-|b| > -1$. Thus

$$\begin{aligned} 3 - |b| &> 3 - 1 \\ &= 2, \end{aligned}$$

Again, we use the Transitive Rule.

and we can then deduce from the previous chain of inequalities that $|3 - b| > 2$, as desired. □

Remarks

1. The results of Example 1 can also be stated in the form:
   (a) $|3 + a^3| \le 4$, for $|a| \le 1$;
   (b) $|3 - b| > 2$, for $|b| < 1$.
2. The reverse implications

$$\left|3 + a^3\right| \le 4 \Rightarrow |a| \le 1 \quad \text{and} \quad |3 - b| > 2 \Rightarrow |b| < 1$$

are FALSE. For example, try putting $a = -\frac{3}{2}$ and $b = -2$!

Problem 1 Use the Triangle Inequality to prove that:

(a) $|a| \le \frac{1}{2} \Rightarrow |a + 1| \le \frac{3}{2}$; (b) $|b| < \frac{1}{2} \Rightarrow \left|b^3 - 1\right| > \frac{7}{8}$.

1.3.2 Inequalities involving *n*

In Analysis we often need to prove inequalities involving an integer n. It is a common convention in mathematics that the symbol n is used to denote an integer (frequently a natural number).

It is often possible to deal with inequalities involving n by using the rearrangement rules given in Section 1.2. Here is such an example.

Example 2 Prove that $2n^2 \geq (n+1)^2$, for $n \geq 3$.

n	1	2	3	4
$2n^2$	2	8	18	32
$(n+1)^2$	4	9	16	25

Solution Rearranging this inequality into an equivalent form, we obtain

$$\begin{aligned} 2n^2 \geq (n+1)^2 &\Leftrightarrow 2n^2 - (n+1)^2 \geq 0 \\ &\Leftrightarrow n^2 - 2n - 1 \geq 0 \\ &\Leftrightarrow (n-1)^2 - 2 \geq 0 \quad \text{(by 'completing the square')} \\ &\Leftrightarrow (n-1)^2 \geq 2. \end{aligned}$$

This final inequality is clearly true for $n \geq 3$, and so the original inequality $2n^2 \geq (n+1)^2$ is true for $n \geq 3$. □

Remarks

1. In Problem 3 of Section 1.2, we asked you to solve the inequality $2x^2 \geq (x+1)^2$; its solution set was $(-\infty, 1-\sqrt{2}] \cup [1+\sqrt{2}, \infty)$. In Example 2, above, we found those natural numbers n lying in this solution set.
2. An alternative solution to Example 2 is as follows

$$\begin{aligned} 2n^2 \geq (n+1)^2 &\Leftrightarrow 2 \geq \left(\frac{n+1}{n}\right)^2 \quad \text{(by Rule 3)} \\ &\Leftrightarrow \sqrt{2} \geq 1 + \frac{1}{n} \quad \text{(by Rule 5, with } p = \tfrac{1}{2}); \end{aligned}$$

and this final inequality certainly holds for $n \geq 3$.

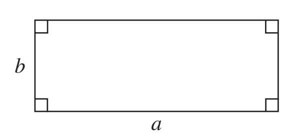
Problem 2 Prove that $\frac{3n}{n^2+2} < 1$, for $n > 2$.

1.3.3 More on inequalities

We now look at a number of inequalities and methods for proving inequalities that will be useful later on.

Example 3 Prove that $ab \leq \left(\frac{a+b}{2}\right)^2$, for $a, b \in \mathbb{R}$.

This has the following geometric interpretation: The area of a rectangle with sides of length a and b is less than or equal to the area of a square with sides of length $\frac{a+b}{2}$.



Solution We tackle this inequality using the various rearrangement rules and a chain of equivalent inequalities until we obtain an inequality that we know must be true

$$\begin{aligned} ab \leq \left(\frac{a+b}{2}\right)^2 &\Leftrightarrow ab \leq \frac{a^2+2ab+b^2}{4} \\ &\Leftrightarrow 4ab \leq a^2 + 2ab + b^2 \\ &\Leftrightarrow 0 \leq a^2 - 2ab + b^2 \\ &\Leftrightarrow 0 \leq (a-b)^2. \end{aligned}$$

This final inequality is certainly true, since all squares are non-negative. It follows that the original inequality $ab \le \left(\frac{a+b}{2}\right)^2$ is also true, for $a, b \in \mathbb{R}$. □

In the form $\sqrt{ab} \le \frac{a+b}{2}$ this inequality is sometimes called the *Arithmetic–Geometric Mean Inequality* for a, b.

Remark

A close examination of the above chain of equivalent statements shows that in fact $ab = \left(\frac{a+b}{2}\right)^2$ if and only if $a = b$.

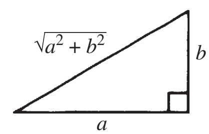
Problem 3 Prove that $\frac{a+b}{\sqrt{2}} \le \sqrt{a^2+b^2}$, for $a, b \in \mathbb{R}$.

Problem 4 Suppose that $a > \sqrt{2}$. Prove the following inequalities:
(a) $\frac{1}{2}\left(a + \frac{2}{a}\right) < a$; (b) $\left(\frac{1}{2}\left(a + \frac{2}{a}\right)\right)^2 > 2$.

Hint: In part (b), use the result of Example 3 and the subsequent remark.

Example 4 Prove that $\sqrt{a^2+b^2} \le a + b$, for $a, b \ge 0$.

This has the following geometric interpretation: The length of the hypotenuse of a right-angled triangle whose other sides are of lengths a and b is less than or equal to the sum of the lengths of those two sides.



Solution We tackle this inequality using the various rearrangement rules and a chain of equivalent inequalities until we obtain an inequality that we know must be true

$$\begin{aligned}\sqrt{a^2+b^2} \le a+b &\Leftrightarrow a^2+b^2 \le (a+b)^2 \\ &\Leftrightarrow a^2+b^2 \le a^2+2ab+b^2 \\ &\Leftrightarrow 0 \le 2ab.\end{aligned}$$

This final inequality is certainly true, since $a, b \ge 0$. It follows that the original inequality $\sqrt{a^2+b^2} \le a+b$ is also true, for $a, b \ge 0$. □

Problem 5 Use the result of Example 4 to prove that $\sqrt{c+d} \le \sqrt{c} + \sqrt{d}$, for $c, d \ge 0$.

Example 5 Prove that $\left|\sqrt{a} - \sqrt{b}\right| \le \sqrt{|a-b|}$, for $a, b \ge 0$.

Solution Notice first that interchanging the roles of a and b leaves the inequality unaltered. It follows that it is sufficient to prove the inequality under the assumption that $a \ge b$.

This will simplify the details of our chain of inequalities.

So, assume that $a \ge b$. Then we know that $\sqrt{a} \ge \sqrt{b}$ and $|a-b| = a-b$. Hence

$$\begin{aligned}\left|\sqrt{a} - \sqrt{b}\right| \le \sqrt{|a-b|} &\Leftrightarrow \sqrt{a} - \sqrt{b} \le \sqrt{a-b} \\ &\Leftrightarrow \sqrt{a} \le \sqrt{a-b} + \sqrt{b}.\end{aligned}$$

We avoid one modulus as a result of our simplifying assumption!

This final inequality is certainly true, and is obtained from the result of Problem 5 by simply substituting $a-b$ in place of c and b in place of d. It follows that the original inequality $\left|\sqrt{a} - \sqrt{b}\right| \le \sqrt{|a-b|}$ is also true, for $a, b \ge 0$. □

Never be ashamed to utilise every tool at your disposal! (Why do the same work twice?)

Problem 6 Prove that $\sqrt{a+b+c} \le \sqrt{a} + \sqrt{b} + \sqrt{c}$, for $a, b, c \ge 0$.

We often use the Binomial Theorem and the Principle of Mathematical Induction (see Appendix 1) to prove inequalities.

Example 6 Prove the following inequalities, for $n \ge 1$:
(a) $2^n \ge 1 + n$; (b) $2^{\frac{1}{n}} \le 1 + \frac{1}{n}$.

n	1	2	3	4
2^n	2	4	8	16
$1+n$	2	3	4	5

Solution

(a) By the Binomial Theorem for $n \geq 1$

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + x^n$$
$$\geq 1 + nx, \qquad \text{for } x \geq 0.$$

Then, if we substitute $x = 1$ in this last inequality, we get

$$2^n \geq 1 + n, \quad \text{for } n \geq 1.$$

n	1	2	3	4
$2^{\frac{1}{n}}$	2	1.41	1.26	1.19
$1+\frac{1}{n}$	2	1.5	1.33	1.25

We decrease the sum by omitting subsequent non-negative terms.

(b) We start by rewriting the required result in an equivalent form

$$2^{\frac{1}{n}} \leq 1 + \frac{1}{n} \Leftrightarrow 2 \leq \left(1 + \frac{1}{n}\right)^n \qquad \text{(by the Power Rule).}$$

Now, if we substitute $x = \frac{1}{n}$ in the Binomial Theorem for $(1+x)^n$, we get

$$\left(1+\frac{1}{n}\right)^n = 1 + n\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{1}{n}\right)^n$$
$$\geq 1 + 1 = 2.$$

We decrease the sum by omitting all but the first two terms.

Since the inequality $2 \leq \left(1+\frac{1}{n}\right)^n$ is true, it follows that the original inequality $2^{\frac{1}{n}} \leq 1 + \frac{1}{n}$, for $n \geq 1$, is also true, as required. □

Problem 7 Prove the inequality $\left(1+\frac{1}{n}\right)^n \geq \frac{5}{2} - \frac{1}{2n}$, for $n \geq 1$.

Hint: consider the first three terms in the binomial expansion.

Example 7 Prove that $2^n \geq n^2$, for $n \geq 4$.

n	1	2	3	4	5
2^n	2	4	8	16	32
n^2	1	4	9	16	25

Solution Let $P(n)$ be the statement

$$P(n):\ 2^n \geq n^2.$$

First we show that $P(4)$ is true: $2^4 \geq 4^2$.

STEP 1 Since $2^4 = 16$ and $4^2 = 16$, $P(4)$ is certainly true.

STEP 2 We now assume that $P(k)$ holds for some $k \geq 4$, and deduce that $P(k+1)$ is then true.

So, we are assuming that $2^k \geq k^2$. Multiplying this inequality by 2 we get

This assumption is just $P(k)$.

$$2^{k+1} \geq 2k^2,$$

so it is therefore sufficient for our purposes to prove that $2k^2 \geq (k+1)^2$. Now

Since $P(k+1)$ is: $2^{k+1} \geq (k+1)^2$.

$$2k^2 \geq (k+1)^2 \Leftrightarrow 2k^2 \geq k^2 + 2k + 1$$
$$\Leftrightarrow k^2 - 2k - 1 \geq 0 \quad \text{(by 'completing the square')}$$
$$\Leftrightarrow (k-1)^2 - 2 \geq 0.$$

This last inequality certainly holds for $k \geq 4$, and so $2^{k+1} \geq (k+1)^2$ also holds for $k \geq 4$.

In other words: $P(k)$ true for some $k \geq 4 \Rightarrow P(k+1)$ true.

It follows, by the Principle of Mathematical Induction, that $2^n \geq n^2$, for $n \geq 4$. □

Problem 8 Prove that $4^n > n^4$, for $n \geq 5$.

Three important inequalities in Analysis

Our first inequality, called *Bernoulli's Inequality*, will be of regular use in later chapters.

> **Theorem 1 Bernoulli's Inequality**
> For any real number $x \geq -1$ and any natural number n, $(1+x)^n \geq 1+nx$.

The value of this result will come from making suitable choices of x and n for particular purposes.

Remark

In part (a) of Example 6, you saw that $(1+x)^n \geq 1+nx$, for $x>0$ and n a natural number. Theorem 1 asserts that the same result holds under the *weaker* assumption that $x \geq -1$.

Proof Let $P(n)$ be the statement

$$P(n): (1+x)^n \geq 1+nx, \quad \text{for } x \geq -1.$$

We prove the result using Mathematical Induction.

STEP 1 First we show that $P(1)$ is true: $(1+x)^1 \geq 1+x$. This is obviously true.

STEP 2 We now assume that $P(k)$ holds for some $k \geq 1$, and prove that $P(k+1)$ is then true.

So, we are assuming that $(1+x)^k \geq 1+kx$, for $x \geq -1$. Multiplying this inequality by $(1+x)$, we get

This assumption is $P(k)$. This multiplication is valid since $(1+x) \geq 0$.

$$\begin{aligned}(1+x)^{k+1} &\geq (1+x)(1+kx)\\ &= 1+(k+1)x+kx^2\\ &\geq 1+(k+1)x.\end{aligned}$$

We decrease the expression if we omit the final non-negative term.

Thus, we have $(1+x)^{k+1} \geq 1+(k+1)x$; in other words the statement $P(k+1)$ holds.

So, $P(k)$ true for some $k \geq 1 \Rightarrow P(k+1)$ true.

It follows, by the Principle of Mathematical Induction, that $(1+x)^n \geq 1+nx$, for $x \geq -1$, $n \geq 1$. □

Problem 9 By applying Bernoulli's Inequality with $x = -\frac{1}{(2n)}$, prove that $2^{\frac{1}{n}} \geq 1+\frac{1}{2n-1}$, for any natural number n.

You saw in part (b) of Example 6 that $2^{\frac{1}{n}} \leq 1+\frac{1}{n}$.

Our second inequality is of considerable use in various branches of Analysis. In Problem 3 you proved that $\frac{a+b}{\sqrt{2}} \leq \sqrt{a^2+b^2}$, for $a, b \in \mathbb{R}$. We can rewrite this inequality in the equivalent form $(a+b)^2 \leq 2(a^2+b^2)$ or $(a+b)^2 \leq (a^2+b^2)(1^2+1^2)$. The Cauchy–Schwarz Inequality is a generalisation of this result to $2n$ real numbers.

> **Theorem 2 Cauchy–Schwarz Inequality**
> For any real numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, we have
> $$\begin{aligned}&(a_1b_1+a_2b_2+\cdots+a_nb_n)^2\\ &\quad\leq \left(a_1^2+a_2^2+\cdots+a_n^2\right)\left(b_1^2+b_2^2+\cdots+b_n^2\right).\end{aligned}$$

We give the proof of Theorem 2 at the end of the sub-section.

Problem 10 Use Theorem 2 to prove that for any positive real numbers $a_1, a_2, \ldots, a_n$, then $(a_1 + a_2 + \cdots + a_n)\left(\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}\right) \geq n^2$.

For example, with $n = 3$

$$(1+2+3)\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}\right) = 6 \times \frac{11}{6} = 11 \geq 3^2.$$

Our final result also has many useful applications. In Example 3 you proved that $ab \leq \left(\frac{a+b}{2}\right)^2$, for $a, b \in \mathbb{R}$; it follows that, if a and b are positive, then $(ab)^{\frac{1}{2}} \leq \frac{a+b}{2}$. The Arithmetic Mean–Geometric Mean Inequality is a generalisation of this result for two real numbers to n real numbers.

Theorem 3 Arithmetic Mean–Geometric Mean Inequality
For any positive real numbers $a_1, a_2, \ldots, a_n$, we have

$$(a_1 a_2 \ldots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

We give the proof of Theorem 3 at the end of the sub-section.

Problem 11 Use Theorem 3 with the $n+1$ positive numbers $1, 1+\frac{1}{n}, 1+\frac{1}{n}, \ldots, 1+\frac{1}{n}$ to prove that, for any positive integer n

$$\left(1+\frac{1}{n}\right)^n \leq \left(1+\frac{1}{n+1}\right)^{n+1}.$$

For example, with $n = 3$

$$\left(1+\frac{1}{3}\right)^3 = \frac{64}{27} = 2.37\ldots < \left(1+\frac{1}{4}\right)^4 = \frac{625}{256} = 2.44\ldots$$

Proofs of Theorems 2 and 3

You may omit these proofs at a first reading.

Theorem 2 Cauchy–Schwarz Inequality
For any real numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, we have

$$(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n)^2 \leq \left(a_1^2 + a_2^2 + \cdots + a_n^2\right)\left(b_1^2 + b_2^2 + \cdots + b_n^2\right).$$

Proof If all the as are zero, the result is obvious; so we need only examine the case when not all the as are zero. It follows that, if we denote the sum $\sum_{k=1}^n a_k^2 = a_1^2 + a_2^2 + \cdots + a_n^2$ by A, then $A > 0$. Also, denote $\sum_{k=1}^n b_k^2 = b_1^2 + b_2^2 + \cdots + b_n^2$ by B and $\sum_{k=1}^n a_k b_k = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$ by C.

Note that we will use the Σ notation to keep the argument brief.

Now, for any real number λ, we have that $(\lambda a_k + b_k)^2 \geq 0$, so that

$$\sum_{k=1}^n \left(\lambda^2 a_k^2 + 2\lambda a_k b_k + b_k^2\right) = \sum_{k=1}^n (\lambda a_k + b_k)^2 \geq 0,$$

which we may rewrite in the form

$$\lambda^2 A + 2\lambda C + B \geq 0.$$

But this inequality is equivalent to the inequality

$$(\lambda A + C)^2 + AB \geq C^2, \text{ for any real number } \lambda.$$

Since A is non-zero, we may now choose $\lambda = -\frac{C}{A}$. It follows from the last inequality that $AB \geq C^2$, which is exactly what we had to prove. ■

A is non-zero, by assumption, so that $1/A$ makes sense.

Remark

If not all the as are zero, equality can only occur if $\sum_{k=1}^{n} (\lambda a_k + b_k)^2 = 0$; that is, if all the numbers a_k are proportional to all the numbers b_k, $1 \leq k \leq n$.

Theorem 3 Arithmetic Mean–Geometric Mean Inequality

For any positive real numbers $a_1, a_2, \ldots, a_n$, we have

$$(a_1 a_2 \ldots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}. \quad (2)$$

Proof Since the a_i are positive, we can rewrite (2) in the equivalent form

$$\frac{(a_1 a_2 \ldots a_n)^{\frac{1}{n}}}{(a_1 + a_2 + \cdots + a_n)/n} \leq 1. \quad (3)$$

We denote the typical term by a_i rather than a_k to avoid confusion with a different use of the letter k in the Mathematical Induction argument below.

Now, replacing each term a_i by λa_i for any non-zero number λ does not alter the left-hand side of the inequality (3). It follows that it is sufficient to prove the inequality (2) in the special case when the product of the terms a_i is 1. Hence it is sufficient to prove the following statement $P(n)$ for each natural number n:

P(n): For any positive real numbers a_i with $a_1 a_2 \ldots a_n = 1$, then $a_1 + a_2 + \cdots + a_n \geq n$.

We will prove this by Mathematical Induction.

First, the statement $P(1)$ is obviously true.

Next, we assume that $P(k)$ holds for some $k \geq 1$, and prove that $P(k+1)$ is then true.

Now, if all the terms $a_1, a_2, \ldots, a_{k+1}$ are equal to 1, the result $P(k+1)$ certainly holds. Otherwise, at least two of the terms differ from 1, say a_1 and a_2, such that $a_1 > 1$ and $a_2 < 1$. Hence

$$(a_1 - 1) \times (a_2 - 1) \leq 0,$$

The argument is exactly the same *whichever* two terms actually differ from 1.

which after some manipulation we may rewrite as

$$a_1 + a_2 \geq 1 + a_1 a_2. \quad (4)$$

You should check this yourself.

We are now ready to tackle $P(k+1)$. Then

$$\begin{aligned} a_1 + a_2 + \cdots + a_{k+1} &\geq 1 + a_1 a_2 + a_3 + a_4 + \cdots + a_{k+1} \\ &\geq k + 1, \end{aligned}$$

By (4).

since we may apply the assumption that $P(k)$ holds to the k quantities $a_1 a_2$, $a_3, a_4, \ldots, a_{k+1}$. This last inequality is simply the statement that $P(k+1)$ is indeed true.

That is $(a_1 a_2) a_3 a_4 \ldots a_{k+1} = 1$ $\Rightarrow (a_1 a_2) + a_3 + a_4 + \cdots + a_{k+1} \geq k$.

It follows by the Principle of Mathematical Induction that $P(n)$ holds for all natural numbers n, and so the inequality (2) must also hold. ∎

Remark

A careful examination of the proof of Theorem 3 shows that equality can only occur if all the terms a_i are equal.

1.4 Least upper bounds and greatest lower bounds

1.4.1 Upper and lower bounds

Any finite set $\{x_1, x_2, \ldots, x_n\}$ of real numbers obviously has a greatest element and a least element, but this property does not necessarily hold for infinite sets.

For example, the interval (0, 2] has greatest element 2, but neither of the sets $\mathbb{N} = \{1, 2, 3, \ldots\}$ nor [0, 2) has a greatest element. However the set [0, 2) is bounded above by 2, since all points of [0, 2) are less than or equal to 2.

> **Definitions** A set $E \subseteq \mathbb{R}$ is **bounded above** if there is a real number, M say, called an **upper bound** of E, such that
>
> $$x \le M, \quad \text{for all } x \in E.$$
>
> If the upper bound M belongs to E, then M is called the **maximum element** of E, denoted by max E.

Geometrically, the set E is bounded above by M if no point of E lies to the right of M on the real line.

For example, if $E = [0, 2)$, then the numbers 2, 3, 3.5 and 157.1 are all upper bounds of E, whereas the numbers 1.995, 1.5, 0 and -157.1 are not upper bounds of E. Although it seems obvious that [0, 2) has no maximum element, you may find it difficult to write down a formal proof. The following example shows you how to do this:

Example 1 Determine which of the following sets are bounded above, and which have a maximum element:

(a) $E_1 = [0, 2)$; (b) $E_2 = \{\frac{1}{n}: n = 1, 2, \ldots\}$; (c) $E_3 = \mathbb{N}$.

Solution

(a) The set E_1 is bounded above. For example, $M = 2$ is an upper bound of E_1, since

$$x \le 2, \quad \text{for all } x \in E_1.$$

However, E_1 has no maximum element. For each x in E_1, we have $x < 2$, and so there is some real number y such that

$$x < y < 2,$$

by the Density Property of $\mathbb{R}$.

Hence $y \in E_1$, and so x cannot be a maximum element.

For example, y can be of the form $1.99\ldots9$ or $y = \frac{1}{2}(x+2)$.

2 is not a maximum element, since $2 \notin E_1$.

(b) The set E_2 is bounded above. For example, $M = 1$ is an upper bound of E_2, since

$$\frac{1}{n} \le 1, \quad \text{for all } n = 1, 2, \ldots.$$

Also, since $1 \in E_2$

$$\max E_2 = 1.$$

(c) The set E_3 is not bounded above. For each real number M, there is a positive integer n such that $n > M$, by the Archimedean Property of $\mathbb{R}$.

Hence M cannot be an upper bound of E_3.

This also means that E_3 cannot have a maximum element. □

Problem 1 Sketch the following sets, and determine which are bounded above, and which have a maximum element:

(a) $E_1 = (-\infty, 1]$; (b) $E_2 = \{1 - \frac{1}{n} : n = 1, 2, \ldots\}$;
(c) $E_3 = \{n^2 : n = 1, 2, \ldots\}$.

Similarly, we define *lower bounds*. For example, the interval (0, 2) is bounded below by 0, since

$$0 \le x, \quad \text{for all } x \in (0, 2).$$

However, 0 does not belong to (0, 2), and so 0 is not a minimum element of (0, 2). In fact, (0, 2) has no minimum element.

Definitions A set $E \subseteq \mathbb{R}$ is **bounded below** if there is a real number, m say, called a **lower bound** of E, such that

$$m \le x, \quad \text{for all } x \in E.$$

If the lower bound m belongs to E, then m is called the **minimum element** of E, denoted by min E.

Geometrically, the set E is bounded below by m if no point of E lies to the left of m on the real line.

Problem 2 Determine which of the following sets are bounded below, and which have a minimum element:

(a) $E_1 = (-\infty, 1]$; (b) $E_2 = \{1 - \frac{1}{n}: n = 1, 2, \ldots\}$;
(c) $E_3 = \{n^2: n = 1, 2, \ldots\}$.

The following terminology is also useful:

Definition A set $E \subseteq \mathbb{R}$ is **bounded** if it is bounded above *and* bounded below.

For example, the set $E_2 = \{1 - \frac{1}{n}: n = 1, 2, \ldots\}$ is bounded, but the sets $E_1 = (-\infty, 1]$ and $E_3 = \{n^2: n = 1, 2, \ldots\}$ are not bounded.

Similar terminology applies to functions.

Definitions A function f defined on an interval $I \subseteq \mathbb{R}$ is said to:

- be **bounded above** by M if $f(x) \le M$, for all $x \in I$; M is an **upper bound** of f;
- be **bounded below** by m if $f(x) \ge m$, for all $x \in I$; m is a **lower bound** of f;
- have a **maximum** (or **maximum value**) M if M is an upper bound of f and $f(x) = M$, for at least one $x \in I$;
- have a **minimum** (or **minimum value**) m if m is a lower bound of f and $f(x) = m$, for at least one $x \in I$.

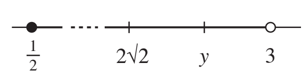
Strictly speaking, M and m are the upper bound and lower bound *of the image set* $\{f(x): x \in I\}$.

Example 2 Let f be the function defined by $f(x) = x^2$, $x \in \left[\frac{1}{2}, 3\right)$. Determine whether f is bounded above or below, and any maximum or minimum value of f.

Solution First, f is increasing on the interval $\left[\frac{1}{2}, 3\right)$, so that since $\frac{1}{2} \le x < 3$ it follows that $\frac{1}{4} \le f(x) < 9$. Hence f is bounded above and bounded below.

Next, since $f\left(\frac{1}{2}\right) = \frac{1}{4}$ and $\frac{1}{4}$ is a lower bound for f on the interval $\left[\frac{1}{2}, 3\right)$, it follows that f has a minimum value of $\frac{1}{4}$ on this interval.

Finally, 9 is an upper bound for f on the interval $\left[\frac{1}{2}, 3\right)$ but there is no point x in $\left[\frac{1}{2}, 3\right)$ for which $f(x) = 9$. So 9 cannot be a maximum of f on the interval. However, if y is any number in (8, 9) there is a number $x > \sqrt{y}$ in $\left(\sqrt{y}, 3\right) \subset \left(2\sqrt{2}, 3\right) \subset \left[\frac{1}{2}, 3\right)$ such that $f(x) = x^2 > y$, so that no number in (8, 9) will serve



as a maximum of f on the interval. It follows that f has no maximum value on $\left[\frac{1}{2}, 3\right)$. □

Problem 3 Let f be the function defined by $f(x) = \frac{1}{x^2}$, $x \in [-3, -2)$. Determine whether f is bounded above or below, and any maximum or minimum value of f.

1.4.2 Least upper bounds and greatest lower bounds

We have seen that the interval [0, 2] has a maximum element 2, but [0, 2) has no maximum element. However, the number 2 is 'rather like' a maximum element of [0, 2), because 2 is an upper bound of [0, 2) and any number less than 2 is not an upper bound of [0, 2). In other words, 2 is the *least* upper bound of [0, 2).

Definition A real number M is the **least upper bound**, or **supremum**, of a set $E \subseteq \mathbb{R}$ if:

1. M is an upper bound of E;
2. if $M' < M$, then M' is not an upper bound of E.

In this case, we write $M = \sup E$.

Part 1 says that M is *an* upper bound.
Part 2 says that *no smaller number* can be an upper bound.

If E has a maximum element, max E, then $\sup E = \max E$. For example, the closed interval [0, 2] has least upper bound 2. We can think of the least upper bound of a set, when it exists, as a kind of 'generalised maximum element'.

If a set does not have a maximum element, but is bounded above, then we may be able to guess the value of its least upper bound. As in the case $E = [0, 2)$, there may be an obvious 'missing point' at the upper end of the set. However it is important to *prove* that your guess is correct. We now show you how to do this.

Example 3 Prove that the least upper bound of [0, 2) is 2.

Solution We know that $M = 2$ is an upper bound of [0, 2), because

$$x \leq 2, \quad \text{for all } x \in [0, 2).$$

To show that 2 is the *least* upper bound, we must prove that each number $M' < 2$ is *not* an upper bound of [0, 2). To do this, we must find an element x in [0, 2) which is greater than M'. But, if $M' < 2$, then there is a real number x such that

$$M' < x < 2$$

and also

$$0 < x < 2.$$

For example, x can be of the form 1.99 . . . 9 for a suitably large number of digits, or it can be $\frac{1}{2}(M' + 2)$ since $M' < \frac{1}{2}(M' + 2) < 2$.

Since $x \in [0, 2)$, the number M' cannot be an upper bound of [0, 2). Hence $M = 2$ is the least upper bound, or supremum, of [0, 2). □

Although the conclusion of Example 3 may seem painfully obvious, we have written out the solution in detail because it illustrates the strategy for determining the least upper bound of a set, if it has one.

> **Strategy** Given a subset E of $\mathbb{R}$, to show that M is the least upper bound, or supremum, of E, check that:
> 1. $x \le M$, for *all* $x \in E$;
> 2. if $M' < M$, then there is *some* $x \in E$ such that $x > M'$.

GUESS the value of M, then CHECK parts 1 and 2.

Notice that, if M is an upper bound of E and $M \in E$, then part 2 is automatically satisfied, and so $M = \sup E = \max E$.

Example 4 Determine the least upper bound of $E = \{1 - \frac{1}{n^2} : n = 1, 2, \ldots\}$.

Solution We guess that the least upper bound of E is $M = 1$. Certainly, 1 is an upper bound of E, since

$$1 - \frac{1}{n^2} \le 1, \quad \text{for } n = 1, 2, \ldots.$$

To check part 2 of the strategy, we need to show that, if $M' < 1$, then there is some natural number n such that

$$1 - \frac{1}{n^2} > M'. \tag{1}$$

However

$$\begin{aligned} 1 - \frac{1}{n^2} > M' &\Leftrightarrow 1 - M' > \frac{1}{n^2} \\ &\Leftrightarrow \frac{1}{1 - M'} < n^2 \qquad (\text{since } 1 - M' > 0) \\ &\Leftrightarrow \sqrt{\frac{1}{1 - M'}} < n \qquad \left(\text{since } \frac{1}{1 - M'} > 0 \text{ and } n > 0\right). \end{aligned}$$

We can certainly choose n so that this final inequality holds, by the Archimedean Property of $\mathbb{R}$, and so we can choose n so that inequality (1) holds.

Hence 1 is the least upper bound of E. □

That is, $1 = \sup E$.

Remark

Although we used double-headed arrows in this solution, the actual proof required only the implications going from *right* to *left*. In other words, the proof uses only the fact that

$$1 - \frac{1}{n^2} > M' \Leftarrow \sqrt{\frac{1}{1 - M'}} < n.$$

Problem 4 Determine sup E, if it exists, for each of the following sets:

(a) $E_1 = (-\infty, 1]$; (b) $E_2 = \{1 - \frac{1}{n} : n = 1, 2, \ldots\}$;
(c) $E_3 = \{n^2 : n = 1, 2, \ldots\}$.

Similarly, we define the notion of a *greatest lower bound*.

> **Definition** A real number m is the **greatest lower bound**, or **infimum**, of a set $E \subseteq \mathbb{R}$ if:
> 1. m is a lower bound of E;
> 2. if $m' > m$, then m' is not a lower bound of E.
>
> In this case, we write $m = \inf E$.

Part 1 says that m is *a* lower bound.
Part 2 says that *no larger number* can be a lower bound.

If E has a minimum element, $\min E$, then $\inf E = \min E$. For example, the closed interval [0, 2] has greatest lower bound 0. We can think of the greatest lower bound of a set, when it exists, as a kind of 'generalised minimum element'.

The strategy for establishing that a number is the greatest lower bound of a set is very similar to that for proving that a number is the least upper bound of a set.

Strategy Given a subset E of $\mathbb{R}$, to show that m is the greatest lower bound, or infimum, of E, check that:

1. $x \geq m$, for *all* $x \in E$;
2. if $m' > m$, then there is *some* $x \in E$ such that $x < m'$.

GUESS the value of m, then CHECK parts 1 and 2.

Notice that, if m is a lower bound of E and $m \in E$, then part 2 is automatically satisfied, and so $m = \inf E = \min E$.

Problem 5 Determine $\inf E$, if it exists, for each of the following sets:

(a) $E_1 = (1, 5]$; (b) $E_2 = \{\frac{1}{n^2} : n = 1, 2, \ldots\}$.

Remarks

1. For any subset E of $\mathbb{R}$, $\inf E \leq \sup E$. This follows from the fact that, for any $x \in E$, we have $\inf E \leq x \leq \sup E$.
2. For any bounded interval I of $\mathbb{R}$, let a be its left end-point and b its right end-point. Then $\inf I = a$ and $\sup I = b$.

Least upper bounds and greatest lower bounds of functions

Similar terminology applies to bounds for functions.

Definitions Let f be a function defined on an interval $I \subseteq \mathbb{R}$. Then:

- A real number M is the **least upper bound**, or **supremum**, of f on I if:
  1. M is an upper bound of $f(I)$;
  2. if $M' < M$, then M' is not an upper bound of $f(I)$.

  In this case, we write $M = \sup f$ or $\sup\limits_I f$ or $\sup\{f(x) : x \in I\}$ or $\sup\limits_{x \in I} f(x)$.
- A real number m is the **greatest lower bound**, or **infimum**, of f on I if:
  1. m is a lower bound of $f(I)$;
  2. if $m' > m$, then m' is not a lower bound of $f(I)$.

  In this case, we write $m = \inf f$ or $\inf\limits_I f$ or $\inf\{f(x) : x \in I\}$ or $\inf\limits_{x \in I} f(x)$.

There are similar definitions for the least upper bound and the greatest lower bound of f on a general set S in $\mathbb{R}$.

These are really the definitions for the least upper bound or the greatest lower bound of the set $\{f(x): x \in I\}$.

Notice, for instance, that:

- if M is an upper bound for f on I, then $\sup\limits_I f \leq M$,
- if m is a lower bound for f on I, then $\inf\limits_I f \geq m$.

For $\sup\limits_I f$ is the *least* upper bound, and $\inf\limits_I f$ is the *greatest* lower bound, for f on I.

The strategies for *proving* that M is the least upper bound or m the greatest lower bound of f on I are similar to the corresponding strategies for the least upper bound or the greatest lower bound of a set E.

> **Strategies** Let f be a function defined on an interval $I \subseteq \mathbb{R}$. Then:
>
> - To show that m is the **greatest lower bound**, or **infimum**, of f on I, check that:
>   1. $f(x) \geq m$, for *all* $x \in I$;
>   2. if $m' > m$, then there is *some* $x \in I$ such that $f(x) < m'$.
> - To show that M is the **least upper bound**, or *supremum*, of f on I, check that:
>   1. $f(x) \leq M$, for *all* $x \in I$;
>   2. if $M' < M$, then there is *some* $x \in I$ such that $f(x) > M'$.

Example 5 Let f be the function defined by $f(x) = x^2, x \in \left[\frac{1}{2}, 3\right)$. Determine the least upper bound and the greatest lower bound of f on $\left[\frac{1}{2}, 3\right)$.

Solution We have already seen that 9 is an upper bound for f on $\left[\frac{1}{2}, 3\right)$, and that no smaller number will serve as an upper bound. It follows that 9 must be the least upper bound of f on $\left[\frac{1}{2}, 3\right)$.

You saw this in Example 2.

Similarly, we have already seen that $\frac{1}{4}$ is the minimum value of f on $\left[\frac{1}{2}, 3\right)$; it follows that $\frac{1}{4}$ is the greatest lower bound of f on $\left[\frac{1}{2}, 3\right)$, and this is actually attained (at the point $\frac{1}{2}$). □

Example 2.

Problem 6 Let f be the function defined by $f(x) = \frac{1}{x^2}$, $x \in [1, 4)$. Determine the least upper bound and the greatest lower bound of f on $[1, 4)$.

Remark

For any interval I of $\mathbb{R}$, $\inf_I f \leq \sup_I f$. This follows from the fact that, for any $x \in I$, we have $\inf_I f \leq f(x) \leq \sup_I f$.

The least upper bound and the greatest lower bound of a function on an interval will be particularly significant in our later work on continuity and integrability of functions.

Chapters 4 and 7.

1.4.3 The Least Upper Bound Property

In the examples in the previous sub-section, it was easy to guess the values of sup E and inf E. At times, however, we shall meet sets for which these values are not so easy to determine. For example, if

$$E = \left\{ \left(1 + \frac{1}{n}\right)^n : n = 1,\ 2, \ldots \right\},$$

We will study this set closely in Section 2.5.

then it can be shown that E is bounded above by 3, but it is not easy to guess the least upper bound of E.

In such circumstances, it is reassuring to know that sup E does exist, even though it may be difficult to find. This existence is guaranteed by the following fundamental result.

The Least Upper Bound Property of $\mathbb{R}$ Let E be a non-empty subset of $\mathbb{R}$. If E is bounded above, then E has a least upper bound.

The Least Upper Bound Property of $\mathbb{R}$ is an example of an *existence theorem*, one which asserts that a real number exists having a certain property. Analysis contains many such results which depend on the Least Upper Bound Property of $\mathbb{R}$. While these results are often very general, and their proofs elegant, they do not always provide the most efficient methods of calculating good approximate values for the numbers in question.

We leave the proof of the Least Upper Bound Property of $\mathbb{R}$ to the next subsection. However, the Property itself is intuitively obvious. If the set E lies entirely to the left of some number M, then you can imagine moving M steadily to the left until you meet E. At this point, sup E has been reached.

The Least Upper Bound Property of $\mathbb{R}$ can be used to show that $\mathbb{R}$ does include decimals which represent irrational numbers such as $\sqrt{2}$, as we claimed in Section 1.1.

In Sections 1.2 and 1.3 we have taken for granted the existence of rational powers and their properties, without giving formal definitions. How can we supply these definitions? For example, how can we define $\sqrt{2}$ as a decimal?

Consider the set

$$E = \left\{x \in \mathbb{Q} : x > 0, x^2 < 2\right\}.$$

This is the set of positive rational numbers whose squares are less than 2. Intuitively, $\sqrt{2}$ lies on the number line to the right of the numbers in E, but 'only just'! In fact, we should expect $\sqrt{2}$ to be the least upper bound of E. Certainly E has a least upper bound, by the Least Upper Bound Property, because E is bounded above, by 1.5 for example. Thus it seems likely that sup E is the decimal representation of $\sqrt{2}$. But how can we *prove* that $(\sup E)^2 = 2$?

We shall prove this in Section 1.5, once we have described how to do arithmetic with real numbers (decimals).

Finally, note that there is a corresponding result about lower bounds.

The Greatest Lower Bound Property of $\mathbb{R}$ Let E be a non-empty subset of $\mathbb{R}$. If E is bounded below, then E has a greatest lower bound.

1.4.4 Proof of the Least Upper Bound Property

You may omit this proof at a first reading.

We know that E is a non-empty set, and we shall assume for simplicity that E contains at least one positive number. We also know that E is bounded above. The following procedure gives us the successive digits in a particular decimal, which we then prove to be the least upper bound of E.

Procedure to find $a = a_0 \cdot a_1 a_2 \ldots = \sup E$

Choose in succession:

- the greatest integer a_0 such that a_0 is not an upper bound of E;
- the greatest digit a_1 such that $a_0 \cdot a_1$ is not an upper bound of E;
- the greatest digit a_2 such that $a_0 \cdot a_1 a_2$ is not an upper bound of E;

  $\vdots$
- the greatest digit a_n such that $a_0 \cdot a_1 a_2 \ldots a_n$ is not an upper bound of E;

  $\vdots$

For example, if

$$E = \{x \in \mathbb{Q} : x > 0, x^2 < 2\},$$

then

$a_0 = 1$, since $1^2 < 2 < 2^2$;

$a_0 \cdot a_1 = 1.4$, since $1.4^2 < 2 < 1.5^2$;

$a_0 \cdot a_1 a_2 = 1.41$, since $1.41^2 < 2 < 1.42^2$;

$\vdots$

Thus, at the nth stage we choose the digit a_n so that:

- $a_0 \cdot a_1 a_2 \ldots a_n$ *is not* an upper bound of E;
- $a_0 \cdot a_1 a_2 \ldots a_n + \frac{1}{10^n}$ *is* an upper bound of E.

We now prove that the least upper bound of E is $a = a_0 \cdot a_1 a_2 \ldots$.

First, we have to prove that a is an upper bound of E. To do this, we prove that, if $x > a$, then $x \notin E$ (this is equivalent to proving, that, if $x \in E$, then $x \leq a$). We begin by representing x as a non-terminating decimal $x = x_0 \cdot x_1 x_2 \ldots$. Since $x > a$, there is an integer n such that

$$a < x_0 \cdot x_1 x_2 \ldots x_n.$$

Hence

$$x_0 \cdot x_1 x_2 \ldots x_n \geq a_0 \cdot a_1 a_2 \ldots a_n + \frac{1}{10^n},$$

and so, by our choice of a_n, $x = x_0 \cdot x_1\ x_2 \ldots x_n$ is an upper bound of E. Since $x > x_0 \cdot x_1 x_2 \ldots x_n$, we have that $x \notin E$, as required.

Next, we have to show that, if $x < a$, then x is *not* an upper bound of E. Since $x < a$, there is an integer n such that

$$x < a_0 \cdot a_1 a_2 \ldots a_n,$$

and so x is *not* an upper bound of E, by our choice of a_n.

Thus we have proved that a is the least upper bound of E. □

Remark

Notice that this proof does not use any arithmetical properties of the real numbers but only their order properties, together with the arithmetical properties of *rational numbers*. In the next section, we use the Least Upper Bound Property to define some of the arithmetical operations on $\mathbb{R}$.

1.5 Manipulating real numbers

1.5.1 Arithmetic in $\mathbb{R}$

At the end of Section 1.1 we discussed the decimals

$$\sqrt{2} = 1.41421356\ldots \quad \text{and} \quad \pi = 3.14159265\ldots,$$

Here we are assuming that $\sqrt{2}$ and π *can* be represented as decimals.

and asked whether it is possible to add and multiply these numbers to obtain another real number. We now explain how this can be done, using the Least Upper Bound Property of $\mathbb{R}$.

A natural way to obtain a sequence of approximations to the sum $\sqrt{2} + \pi$ is to truncate each of the above decimals, and form the sums of the truncations.

If each of the decimals is truncated at the same decimal place, this gives a sequence of approximations which is increasing:

$\sqrt{2}$	π	$\sqrt{2}+\pi$
1	3	4
1.4	3.1	4.5
1.41	3.14	4.55
1.414	3.141	4.555
1.4142	3.1415	4.5557
$\vdots$	$\vdots$	$\vdots$

Intuitively, we should expect that the sum $\sqrt{2}+\pi$ is greater than each of the numbers in the right-hand column, but 'only just'! To accord with our intuition, therefore, we *define* the sum $\sqrt{2}+\pi$ to be the least upper bound of the set of numbers in the right-hand column; that is

$$\sqrt{2}+\pi = \sup\{4,\ 4.5,\ 4.55,\ 4.555,\ 4.5557, \ldots\}.$$

To be sure that this definition makes sense, we need to show that this set is bounded above. But all the truncations of $\sqrt{2}$ are less than 1.5, and all the truncations of π are less than, say, 4. Hence, all the sums in the right-hand column are less than $1.5+4=5.5$. So, by the Least Upper Bound Property, the set of numbers in the right-hand column *does* have a least upper bound, and we *can* define $\sqrt{2}+\pi$ in this way.

This method can be used to define the sum of any pair of positive real numbers.

Let us check that this method of adding decimals gives the correct answer when we use it in a familiar case. Consider the simple calculation

We do not expect *you* to use this method to add decimals!

$$\frac{1}{3}+\frac{2}{3} = 0.333\ldots + 0.666\ldots.$$

Truncating each of these decimals and forming the sums, we obtain the set

$$\{0,\ 0.9,\ 0.99,\ 0.999, \ldots\}.$$

The supremum of this set is, of course, the number $0.999\ldots = 1$, which is the correct answer.

Similarly, we can define the *product* of any two positive real numbers. For example, to define $\sqrt{2}\times\pi$, we can form the sequence of products of their truncations:

$\sqrt{2}$	π	$\sqrt{2}\times\pi$
1	3	3
1.4	3.1	4.34
1.41	3.14	4.4274
1.414	3.141	4.441374
1.4142	3.1415	4.4427093
$\vdots$	$\vdots$	$\vdots$

As before, we define $\sqrt{2} \times \pi$ to be the least upper bound of the set of numbers in the right-hand column.

Similar ideas can be used to define the operations of subtraction and division.

We omit the details.

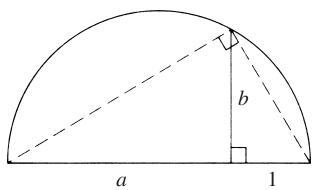
Thus we can define arithmetic with real numbers in terms of the familiar arithmetic with rationals, using the Least Upper Bound Property of $\mathbb{R}$. Moreover, it can be proved that these operations in $\mathbb{R}$ satisfy all the usual properties of a field.

These properties were listed in Sub-section 1.1.5.

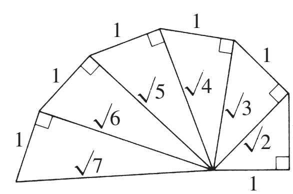
1.5.2 The existence of roots

Just as we usually take for granted the basic arithmetical operations with real numbers, so we usually assume that, given any positive real number a, there is a unique positive real number $b = \sqrt{a}$ such that $b^2 = a$. We now discuss the justification for this assumption.

First, here is a geometrical justification. Given line segments of lengths 1 and a, we can construct a semi-circle with diameter $a + 1$ as shown.



For each positive integer n, we can also construct $\sqrt{n}$ as follows:



Using similar triangles, we see that

$$\frac{a}{b} = \frac{b}{1},$$

and so

$$b^2 = a.$$

This shows that there should be a positive real number b such that $b^2 = a$, so that the length of the vertical line segment in the figure can be described exactly by the expression $\sqrt{a}$. But does $b = \sqrt{a}$ exist *exactly* as a real number? In fact it does, and a more general result is true.

> **Theorem 1** For each positive real number a and each integer $n > 1$, there is a unique positive real number b such that
> $$b^n = a.$$

We shall prove Theorem 1 in Sub-section 4.3.3.

We call this number b the nth root of a, and we write $b = \sqrt[n]{a}$. We also define $\sqrt[n]{0} = 0$, since $0^n = 0$, and if n is odd we define $\sqrt[n]{(-a)} = -\sqrt[n]{a}$, since $(-\sqrt[n]{a})^n = -a$ if n is odd.

For example, $\sqrt[3]{(-8)} = -2$.

Let us illustrate Theorem 1 with the special case $a = 2$ and $n = 2$. In this case, Theorem 1 asserts the existence of a real number b such that $b^2 = 2$. In other words, it asserts the existence of a decimal b which can be used to *define* $\sqrt{2}$ precisely.

Here is a direct proof of Theorem 1 in this special case. We choose the numbers 1, 1.4, 1.41, 1.414, ... to satisfy the inequalities

$$\begin{gathered}1^2 < 2 < 2^2 \\ (1.4)^2 < 2 < (1.5)^2 \\ (1.41)^2 < 2 < (1.42)^2 \\ (1.414)^2 < 2 < (1.415)^2 \\ \vdots\end{gathered} \qquad (1)$$

This process gives an infinite decimal

$$b = 1.414\ldots,$$

and we claim that

$$b^2 = (1.414\ldots)^2 = 2.$$

This can be proved using our method of multiplying decimals:

Notice that $b = 1.414\ldots$ is the decimal that we obtained as the least upper bound of the set $\{x \in \mathbb{Q} : x > 0,\ x^2 < 2\}$ in Sub-section 1.5.1.

b	b	b^2
1	1	1
1.4	1.4	1.96
1.41	1.41	1.9881
1.414	1.414	1.999396
⋮	⋮	⋮

We have to prove that the least upper bound of the set E of numbers in the right-hand column is 2, in other words that

$$\sup E = \sup\{1,\ (1.4)^2,\ (1.41)^2,\ (1.414)^2, \ldots\} = 2.$$

To do this, we employ the strategy given in Sub-section 1.4.2.

First, we check that $M = 2$ is an upper bound of E. This follows from the left-hand inequalities in (1).

Next, we check that, if $M' < 2$, then there is a number in E which is greater than M'. To prove this, put

$$x_0 = 1,\ x_1 = 1.4,\ x_2 = 1.41,\ x_3 = 1.414, \ldots.$$

Then, by the right-hand inequalities in (1), we have that

$$\left(x_n + \frac{1}{10^n}\right)^2 > 2.$$

For example, if $n = 1$, then $\left(1.4 + \frac{1}{10}\right)^2 = 1.5^2 > 2.$

Also

$$\begin{aligned}\left(x_n + \frac{1}{10^n}\right)^2 - x_n^2 &= \frac{1}{10^n}\left(2x_n + \frac{1}{10^n}\right) \\ &< \frac{1}{10^n}(2 \times 2 + 1) = \frac{5}{10^n},\end{aligned}$$

and so

$$\begin{aligned}x_n^2 &> \left(x_n + \frac{1}{10^n}\right)^2 - \frac{5}{10^n} \\ &> 2 - \frac{5}{10^n} \\ &= 1.\underbrace{99\ldots95}_{n \text{ digits}}.\end{aligned}$$

For example, if $n = 2$, then $(1.41)^2 > 1.95.$

So, if $M' < 2$, then we *can* choose n so large that $x_n^2 > M'$ (while still having $x_n \in E$). This proves that the least upper bound of E is 2, and so $(1.414\ldots)^2 = 2$.
Thus we can define

$$\sqrt{2} = 1.414\ldots,$$

which justifies our earlier claim that $\sqrt{2}$ can be represented exactly by a decimal.

1.5.3 Rational powers

Having discussed nth roots, we are now in a position to define the expression a^x, where a is positive and x is rational.

Definition If $a > 0$, $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, then

$$a^{\frac{m}{n}} = \left(\sqrt[n]{a}\right)^m.$$

For example, for $a > 0$, with $m = 1$ we have $a^{\frac{1}{n}} = \sqrt[n]{a}$, and with $m = 2$ and $n = 3$ we have $a^{\frac{2}{3}} = (\sqrt[3]{a})^2$.

This notation is particularly useful, because *rational powers* (or *rational exponents*) satisfy the following *exponent laws* (whose proofs depend on Theorem 1):

Exponent Laws

- If $a, b > 0$ and $x \in \mathbb{Q}$, then $a^x b^x = (ab)^x$.
- If $a > 0$ and $x, y \in \mathbb{Q}$, then $a^x\, a^y = a^{x+y}$.
- If $a > 0$ and $x, y \in \mathbb{Q}$, then $(a^x)^y = a^{xy}$.

For example

$$2^{\frac{1}{2}} \times 3^{\frac{1}{2}} = 6^{\frac{1}{2}},$$
$$2^{\frac{1}{2}} \times 2^{\frac{1}{3}} = 2^{\frac{5}{6}},$$
$$\left(2^{\frac{1}{2}}\right)^{\frac{1}{3}} = 2^{\frac{1}{6}}.$$

If x and y are *integers*, these laws actually hold for all non-zero real numbers a and b. However, if x and y are not integers, then we must have $a, b > 0$. For example, $(-1)^{\frac{1}{2}}$ is not defined as a real number.

However, if a is a negative real number, then $a^{\frac{m}{n}}$ can be defined whenever $m \in \mathbb{Z}$, $n \in \mathbb{N}$ and $\frac{m}{n}$ is reduced to its lowest terms with n odd, as follows

$$a^{\frac{m}{n}} = \left(\sqrt[n]{a}\right)^m.$$

This extends our above definition of $a^{\frac{m}{n}}$; for instance, it defines $a^{\frac{1}{n}}$ whenever $n \in \mathbb{N}$ and n is odd. For example, $(-8)^{\frac{2}{3}} = 4$.

Finally, you may have wondered why we did not mention that each positive number has two nth roots when n is even. For example, $2^2 = (-2)^2 = 4$. We shall adopt the convention that, for $a > 0$, $\sqrt[n]{a}$ and $a^{\frac{1}{n}}$ always refer to the *positive* nth root of a. If we wish to refer to both roots (for example, when solving equations), we write $\pm\sqrt[n]{a}$.

1.5.4 Real powers

We conclude this section by briefly discussing the meaning of a^x when $a > 0$ and x is an arbitrary real number. We have defined this expression when x is rational, but the same definition does not work if x is irrational. However, it is common practice to write down expressions such as $\sqrt{2}^{\sqrt{2}}$, and even to apply the Exponent Laws to give equalities such as