An Introduction to Analysis

R. C. Vaughan

Pennsylvania State University

19th April 2024

Contents

Preface

Thus book is based on courses given for nearly fifty years starting in 1972 at Imperial College London and Penn State University, and is quite close to the undergraduate course the author took on the subject in 1963 from Professor T. Estermann with the problem class conducted by Professor C. A. Rogers. For many students (including the author) analysis is the hardest course they take - the chasm between calculus and analysis can be a massive jump. This author prefers to avoid as much jargon as possible and generally avoids clouding the issue with constant reference to concepts from topology and metric spaces. The intent is to *keep it simple*.

The book contains typically enough material for about thirty six hours of presentations and nine to twelve hours of problem solving and tutorials. All the exercises have been used at least once for homework or the basis of examination questions.

One word of warning. This is a subject which demands proofs, and it would be wise to also have some facility with constructing simple proofs in good English. If one wishes to understand the reasons for a particular phenomenon this can often only be seen by understanding why the proof works.

Chapter 1

Introduction

ch:one

1.1 Motivation

sec:one1

The great power of modern mathematics lies in the axiomatic approach. The original model for this is Euclid's axiomatisation of geometry about 300BC. That is the establishment of a few simple basic statements (axioms) from which all propositions are deduced by basic rules of logical deduction. The wisdom of Euclids original choice is demonstrated by the observation that in the intervening 2300 years nobody has found anything self contradictory in the vast panoply of geometric theorems which have been established in Euclidean geometry. In other words we can have essentially absolute reassurance of the results.

However, Euclidean geometry has its limitations. In the 19th century it was observed that there are different geometries which lie outside Euclidean geometry. Nevertheless they can be described by adjusting the axiom which deals with the concept of parallel lines.

One of the great deficiencies of the ancient world was a good way of describing numbers. They had some understanding of positive whole numbers, but, at least in Europe, the systems for describing them which was derived from the Etruscans was similar to, and eventually evolved into, the Roman numeral system. We know how clumsy that is for doing arithmetic, and it is no surprise to learn that there was no simple way of dealing even with quite simple fractions. Euclid in his elements needs to understand the "length" of a given line segment. Whilst there was language in commerce for the use of simple fractions, normally with denominator 12 (the duodecimal system), for a general fraction he had to resort to the idea of "proportion". In other words given a particular unit length he understood how to produce line segments whose length is twice, thrice, and so on, the unit length. He also understood how to product a line segment whose length l satisfies

$$l : 1 :: m : n$$

where m and n are positive whole numbers, and which in modern notation is simply

$$l = \frac{l}{1} = \frac{m}{n}.$$

All very well and good, but it had already been discovered by the Pythagorean school that not all lengths could be described in this way. That is, there was no rational length whose square was 2. Yet they could construct such lengths from right angled triangles.

thm:one1 **Theorem 1.1.** *There is no rational number whose square is 2. In modern notation the equation*

$$\sqrt{2} = \frac{m}{n}$$

with m and n whole numbers is impossible.

Proof. We argue by contradiction. We can certainly suppose that m and n are positive, and we can remove common factors so that m and n have no common prime factors. Moreover we have

$$2n^2 = m^2.$$

The prime number 2 is a factor of the left hand side, so it must also be a factor of m^2, and hence of m. Write $q = m/2$, so that q is also a positive whole number and

$$2n^2 = 2^2q^2, \quad n^2 = 2q^2.$$

Now repeating the argument we have that 2 is also a factor of n. That is we just showed that m and n do have a common prime factor contradicting our basic assumption. $\square$

The problem for the Pythagoreans was that this seemed to imply that $\sqrt{2}$ does not exist, and gave a paradox against Pythagorus' theorem. Our problem is to resolve this.

Of course we are all familiar with the fact that we can get good approximations to $\sqrt{2}$

$$\begin{aligned} (1.4)^2 &= 1.96 & (1.5)^2 &= 2.25 \\ (1.41)^2 &= 1.9881 & (1.42)^2 &= 2.0164 \\ (1.414)^2 &= 1.999396 & (1.415)^2 &= 2.002225 \\ (1.4142)^2 &= 1.99996164 & (1.4143)^2 &= 2.0002449 \end{aligned}$$

Well it looks as though we should consider $\sqrt{2}$ as the result of some kind of limiting process.

Here is another famous paradox. *In a race, the quickest runner (Achilles) can never overtake the slowest (tortoise), since the pursuer must first reach the point whence the pursued started, so that the slower must always hold a lead - Aristotle c400BC.* If one sets this up as a mathematical problem one ends up by summing convergent infinite series whose limiting values will tell one the time and distance at which Achilles overtakes the tortoise.

During the 18th and early 19th centuries there was controversy in some explanations of the differential calculus because it looked as though the derivative was defined as

$$\frac{0}{0}.$$

Another problem was that it seemed that some functions could be drawn but might have places where they could not be differentiated.

ex:one1 **Example 1.1.** *Which of the following functions is differentiable at 0?*

$$f(x) = \begin{cases} x\sin(1/x) & (x \neq 0), \\ 0 & (x = 0), \end{cases} \tag{1.1}$$ eq:one1

$$g(x) = \begin{cases} x^2\sin(1/x) & (x \neq 0) \\ 0 & (x = 0). \end{cases} \tag{1.2}$$ eq:one2

1.1.1 Exercises

subsec:one1

1. Prove that there is no rational number r with $r^2 = 3$.
2. Prove that there is no rational number y with $y^2 = 180$.
3. Prove that if k is a positive whole number which is not a perfect square, then there is no rational number z with $z^2 = k$.
4. Prove that $\sqrt[3]{5}$ is irrational.
5. Try sketching the function

$$f(x) = \begin{cases} 0 & x \text{ is rational}, \\ 1 & x \text{ is irrational}. \end{cases}$$

1.2 Sets

sec:one2

Here is the dictionary definition of a set.

def:one1 **Definition 1.1.** *A set is a collection of objects called elements.*

Like most dictionary definitions it does not help very much without further insight. If one is not careful it can lead to further paradoxes and difficulties. In order to avoid this we will be concerned solely with sets of numbers or mathematical objects which are defined in a similar way, such as ordered k-tuples of numbers.

When x is an element of the set $\mathcal{S}$ we write

$$x \in \mathcal{S} \tag{1.3}$$ eq:one3

The symbol $\in$ is a variant of the Greek epsilon, ϵ or ε but should not be confused with them and one should try to distinguish them when writing them.

Sets can be defined in various ways.

1. By listing the elements.

$$\begin{aligned}
\mathcal{S} &= \{1, 3, \pi, 7/2, \sqrt{17}\}, && && (1.4)\\
\mathbb{N} &= \{1, 2, 3, 4, 5, 6, \ldots\} && \text{The natural numbers,} && (1.5)\\
\mathbb{Z} &= \{\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\} && \text{The integers,} && (1.6)\\
\mathbb{Q} &= \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\right\} && \text{The rational numbers.} && (1.7)
\end{aligned}$$

eq:one5 eq:one6 eq:one7

2. By some kind of defining formula.

$$\begin{aligned}
\mathcal{T} &= \{x : 1 < x < 2\}, && (1.8)\\
\mathcal{U} &= \{(x, y) : x^2 + y^2 = 1\}. && (1.9)
\end{aligned}$$

3. By combining sets. We will look at this in more detail later.

4. There is one very special set, the *empty* set, usually denoted by

$$\emptyset$$

which is the set which has *NO* elements.

Example 1.2.

$$\{x : x^2 < 0\} = \emptyset.$$

There is an important logical observation. Since the set has no elements its elements can have any property. For example they can be both positive and negative!

An important concept is that of a subset.

def:one2 **Definition 1.2.** *We say that $\mathcal{S}$ is a subset of $\mathcal{T}$ when every element of $\mathcal{S}$ is also an element of $\mathcal{T}$, and we write*

$$\mathcal{S} \subset \mathcal{T}.$$

In this course we will include the possibility that $\mathcal{S} = \mathcal{T}$. Increasingly it is common to use $\subseteq$ in place of $\subset$ and to use the latter to mean that $\mathcal{S}$ is a subset with $\mathcal{S} \neq \mathcal{T}$, i.e. $\mathcal{S}$ is a *proper* subset of $\mathcal{T}$.

Note that the empty set $\emptyset$ is a subset of *every* set!

ex:one2 **Example 1.3.** *The set* $\mathcal{T} = \{1, 3, \pi\}$ *has subsets*

$$\{1,3,\pi\},$$
$$\{1,3\},\ \{1,\pi\},\ \{3,\pi\},$$
$$\{1\},\ \{3\},\ \{\pi\},$$
$$\emptyset$$

Generally a finite set with k elements has 2^k subsets and $\binom{k}{j}$ subsets with exactly j elements.

As promised above we now look at various ways of combining sets. There are three ways commonly used to do this.

def:one3 **Definition 1.3.** *The union of two sets* $\mathcal{A}$ *and* $\mathcal{B}$ *is the set which contains all the elements of* $\mathcal{A}$ *and* $\mathcal{B}$

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

Note the use of the logical "or", not to be confused with "xor", i.e it includes x which are in both sets.

ex:one4 **Example 1.4.**

$$\mathcal{A} = \{1, 2, 3\},\ \mathcal{B} = \{2, 3, 4\},\ \mathcal{A} \cup \mathcal{B} = \{1, 2, 3, 4\}.$$

def:one4 **Definition 1.4.** *The intersection of two sets* $\mathcal{A}$ *and* $\mathcal{B}$ *is the set which contains the elements which are in both* $\mathcal{A}$ *and* $\mathcal{B}$*.*

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

ex:one5 **Example 1.5.** *In the above example*

$$\mathcal{A} \cap \mathcal{B} = \{2, 3\}.$$

ex:one6 **Example 1.6.** *Another example.*

$$\mathcal{U} = \{x : 0 < x < 1\},\ \mathcal{V} = \{1 \leq x \leq 2\},\ \mathcal{U} \cap \mathcal{V} = \emptyset.$$

def:one5 **Definition 1.5.** *The complement of* $\mathcal{B}$ *with respect to* $\mathcal{A}$ *is the set of* x *in* $\mathcal{A}$ *which are not in* $\mathcal{B}$*,*

$$\mathcal{A} \setminus \mathcal{B} = \{x : x \in \mathcal{A} \text{ and } x \notin \mathcal{B}\}.$$

ex:one7 **Example 1.7.** *Again in Example 1.4,*

$$\mathcal{A} \setminus \mathcal{B} = \{1\}.$$

These relationships form quite a complex algebra.

ex:one8 **Example 1.8.** *In general*

$$(\mathcal{C} \setminus \mathcal{D}) \cap (\mathcal{D} \setminus \mathcal{C}) = \emptyset.$$

ex:one9 **Example 1.9.** *In general*

$$\mathcal{C} \cap (\mathcal{D} \cup \mathcal{E}) = (\mathcal{C} \cap \mathcal{D}) \cup (\mathcal{C} \cap \mathcal{E}).$$

We now come to the need for proofs, since some of these relationships are not completely obvious. The recommended way of proving such relationships is by *truth tables.*

For each object x there are two possibilities for each set, x is in it, or x is not in it. To indicate which I will use a 0 or 1 respectively (think of it as the "characteristic or indicator function". Some people use F and T corresponding to it being false or true that the element is in the set.

Returning to Example 1.8.

$\mathcal{C}$	$\mathcal{D}$	$\mathcal{C} \setminus \mathcal{D}$	$\mathcal{D} \setminus \mathcal{C}$	$(\mathcal{C} \setminus \mathcal{D}) \cap (\mathcal{D} \setminus \mathcal{C})$	$\emptyset$
1	1	0	0	0	0
1	0	1	0	0	0
0	1	0	1	0	0
0	0	0	0	0	0

Here is Example 1.9.

$\mathcal{C}$	$\mathcal{D}$	$\mathcal{E}$	$\mathcal{D} \cup \mathcal{E}$	$\mathcal{C} \cap \mathcal{D}$	$\mathcal{C} \cap \mathcal{E}$	$\mathcal{C} \cap (\mathcal{D} \cup \mathcal{E})$	$(\mathcal{C} \cap \mathcal{D}) \cup (\mathcal{C} \cap \mathcal{E})$
1	1	1	1	1	1	1	1
1	1	0	1	1	0	1	1
1	0	1	1	0	1	1	1
0	1	1	1	0	0	0	0
0	0	1	1	0	0	0	0
0	1	0	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

1.2.1 Exercises

subsec:one2

1. Let $\mathcal{A} = \{1, 2, 3, 4, 5\}$, $\mathcal{B} = \{1, 3, 5, 7, 9\}$, $\mathcal{C} = \{2, 4, 6, 8, 10\}$. Find $\mathcal{A} \cup \mathcal{B}$, $\mathcal{B} \cup \mathcal{C}$, $\mathcal{B} \cap \mathcal{C}$, $\mathcal{C} \cap \mathcal{A}$.

2. Prove that $\mathcal{A} = (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \setminus \mathcal{B})$.

3. Prove that $\mathcal{A} \setminus (\mathcal{B} \cup \mathcal{C}) = (\mathcal{A} \setminus \mathcal{B}) \cap (\mathcal{A} \setminus \mathcal{C})$.

4. Prove that $\mathcal{A} \setminus (\mathcal{B} \cap \mathcal{C}) = (\mathcal{A} \setminus \mathcal{B}) \cup (\mathcal{A} \setminus \mathcal{C})$.

5. Prove that $(\mathcal{A} \setminus \mathcal{B}) \cup (\mathcal{B} \setminus \mathcal{A}) = (\mathcal{A} \cup \mathcal{B}) \setminus (\mathcal{A} \cap \mathcal{B})$.
6. Prove that $\mathcal{A} \cap \mathcal{B} = A \setminus (\mathcal{A} \setminus \mathcal{B}) = \mathcal{B} \setminus (\mathcal{B} \setminus \mathcal{A})$.
7. Prove that $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$.
8. Prove that $\mathcal{A} \cup (\mathcal{B} \cap \mathcal{C}) = (\mathcal{A} \cup \mathcal{B}) \cap (\mathcal{A} \cup \mathcal{C})$.
9. Prove that $\mathcal{A} \cap (\mathcal{B} \cup \mathcal{C}) = (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \cap \mathcal{C})$.
10. Prove that

$$\big((\mathcal{B} \cap \mathcal{C}) \cup (\mathcal{C} \cap \mathcal{A})\big) \cup (\mathcal{A} \cap \mathcal{B}) = \big((\mathcal{B} \cup \mathcal{C}) \cap (\mathcal{C} \cup \mathcal{A})\big) \cap (\mathcal{A} \cup \mathcal{B}).$$

11. Prove that $(\mathcal{C} \backslash \mathcal{A}) \cup (\mathcal{C} \backslash \mathcal{B}) = \mathcal{C} \backslash (\mathcal{A} \cap \mathcal{B})$.
12. Prove that $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$.
13. Prove that $(\mathcal{A} \cup \mathcal{B}) \cap \mathcal{C} = (\mathcal{A} \cap \mathcal{C}) \cup (\mathcal{B} \cap \mathcal{C})$.

1.3 The Integers and Rational Numbers

sec:one3

We have already introduced the standard notation $\mathbb{N}$ (1.5), $\mathbb{Z}$ (1.6) and $\mathbb{Q}$ (1.7) for the natural numbers, the integers and the rational numbers respectively. Our main interest is the set or real numbers $\mathbb{R}$, and since we expect that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ we will not dally for long on the other sets. Our intention is to simply introduce a collection of axioms which the elements of $\mathbb{R}$ have to satisfy. However because the properties of $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ impact those of $\mathbb{R}$ it is necessary to say something about how we might axiomatise these sets. We start with the simplest of these sets, $\mathbb{N}$.

def:one6 **Definition 1.6** (Peano axioms for $\mathbb{N}$)**. 1.** *There is an element of* $\mathbf{N}$ *denoted by* 1 *and an operation* $+$ *which combines* 1 *and any element* n *of* $\mathbb{N}$ *to give another element denoted by* $n+1$*, i.e. for every* $n \in \mathbb{N}$ *we have* $n+1 \in \mathbb{N}$*.*

2. *For all* m*,* $n \in \mathbb{N}$*, we have* $m = n$ *if and only if* $m+1 = n+1$*.*

3. *For every* $n \in \mathbf{N}$ *we have* $n + 1 \neq 1$*.*

4. *If* $\mathcal{S}$ *is a set with the properties that (a)* $1 \in \mathcal{S}$ *and (b) whenever* $n \in \mathcal{S}$ *we have* $n+1 \in \mathcal{S}$*, then* $\mathcal{S} = \mathbb{N}$*.*

What this says is that we should think of $\mathbb{N}$ as being

$$\mathbb{N} = \{1, 1+1, 1+1+1, 1+1+1+1, \cdots\}.$$

Axiom **4** is the Principle of Induction.

We can now deduce that $m + n \in \mathbb{N}$ for any m, $n \in \mathbb{N}$. Given m let $\mathcal{S}$ denote the set of n for which $m+n \in \mathbb{N}$. Then by Axiom 1 $m+1 \in \mathbb{N}$, so $1 \in \mathcal{S}$. Suppose $n \in \mathcal{S}$.

Then $m+n \in \mathbb{N}$ and so by Axiom 1 once more we have $m+n+1 \in \mathbb{N}$ and so $n+1 \in \mathcal{S}$. Hence, by Axiom 4 we have $\mathcal{S} = \mathbb{N}$.

In this kind of way various other properties of $\mathbb{N}$ can be established. For example if m, $n \in \mathbb{N}$, then $m+n = n+m$.

We can also define multiplication by taking $n \times 1 = 1 \times n = n$ and $n \times (m+1) = (n \times m) + n$ and using induction. We can then combine addition and multiplication more generally to show that

$$l \times (m+n) = (l \times m) + (l \times n).$$

Later, when developing the ideas of limits we will need to look at the elements of $\mathbb{N}$ in more detail.

How about the integers? It would be good if we could just build on the above. We could introduce a symbol 0 to mean $n+0 = 0+n = n$, and then we could introduce an object $-n$ with the property that $n + (-n) = 0$. However, this begs the question, "why should this exist". To avoid this we follow a different route. One of the more powerful techniques we have is the ability to create more complex and richer systems out of simpler ones. Thus we can think about "extending $\mathbb{N}$ to give $\mathbb{Z}$, and there is a very nice way of doing this by the use of ordered pairs of natural numbers (m, n) and something called equivalence classes. I do not want to spend too much time on this, but briefly it goes like this. Two ordered pairs (k, l) and (m, n) are *equivalent* precisely when

$$k + n = l + m.$$

If we use $\mathcal{A}(m, n)$ to denote the set of ordered pairs equivalent to (m, n), then we can define addition and multiplication by

$$\mathcal{A}(k, l) + \mathcal{A}(m, n) = \mathcal{A}(k+n, l+m), \mathcal{A}(k, l) \times \mathcal{A}(m, n) = \mathcal{A}(km + ln, kn + lm)$$

and define negatives by

$$-\mathcal{A}(m, n) = \mathcal{A}(n, m).$$

Then we can check that these equivalence classes have all the properties that we expect of the integers and declare them to be the integers. In other words we found a way of *constructing* the integers from the natural numbers.

We can then use a similar procedure to construct the rational numbers by now looking at equivalence classes of ordered pairs (p, q) of integers p, q with $q \neq 0$. For example, let $\mathcal{B}(r, s)$ for $s \neq 0$ be the set of such ordered pairs (p, q) with $ps = rq$. Now we can define

$$\mathcal{B}(r, s) + \mathcal{B}(r', s') = \mathcal{B}(rs' + r's, ss'),\ \mathcal{B}(r, s) \times \mathcal{B}(r', s') = \mathcal{B}(rr', ss')$$

and again check that this results in the properties we expect of elements of $\mathbb{Q}$. Again I do not want to spend time checking this. The main problem at hand at this stage is dealing with the question of numbers such as $\sqrt{2}$ where something more profound is needed.

1.3.1 Exercises

subsec:one3

1. Prove that for each $n \in \mathbb{N}$ we have

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

2. (i) Prove that if $1 \leq m \leq n-1$, then

$$\frac{n!}{(m-1)!(n-m+1)!} + \frac{n!}{m!(n-m)!} = \frac{(n+1)!}{m!(n+1-m)!}.$$

(ii) Let $x \in \mathbb{R}$. Prove by induction on n that $(1+x)^n = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} x^m$. (This is the binomial theorem for positive integral index.)

3. Prove that, for every $n \in \mathbb{N}$, $15^n - 8$ is always a multiple of 7.

1.4 Notes

sec:one6

§1.3 For background and history on the Peano axioms see

`https://en.wikipedia.org/wiki/Peano_axioms`

Chapter 2

The Real Numbers

ch:two

2.1 Ordered Fields

sec:two4

We proceed by first listing a collection of axioms which apply more generally than just to $\mathbb{R}$. Indeed they will hold for $\mathbb{Q}$ also. Since there are quite a number we will divide them into two groups, the Arithmetic axioms and the Order axioms. Later we will have to decide what distinguishes $\mathbb{R}$ from $\mathbb{Q}$ and what extra axioms might be required.

def:two7 **Definition 2.1** (Arithmetic axioms for an ordered field)**.** *An ordered field $\mathcal{F}$ has $\mathbb{N}$ as a subset and the following hold for all $a, b, c \in \mathcal{F}$.*

Closure *There are two ways of combining two elements, $+$ and $.$ (or $\times$) such that $a + b$ and $a.b$ are in $\mathcal{F}$.*

Commutative axiom

$$a + b = b + a, \quad ab = ba$$

Associative axiom

$$(a + b) + c = a + (b + c), \quad (ab)c = a(bc).$$

Distributive axiom

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc.$$

Identities *There are elements 0 and 1 such that for every a we have*

$$a + 0 = a = 0 + a, \quad a.1 = 1.a = a.$$

Additive inverse *Given a there is an element $(-a) \in \mathcal{F}$ such that*

$$a + (-a) = (-a) + a = 0$$

Multiplicative inverse *Given $a \neq 0$ there is an $a^{-1} \in \mathcal{F}$ such that*

$$aa^{-1} = a^{-1}a = 1.$$

From these axioms we could deduce all the usual arithmetical properties of numbers. It would take far too long and be far too tedious to do so. Here are some examples.

ex:two10 **Example 2.1.** *If $x + y = x + z$, then $y = z$.*

Proof. We have

$$\begin{aligned} y &= 0 + y && \text{identity} \\ &= ((-x) + x) + y && \text{inverse} \\ &= (-x) + (x + y) && \text{associative} \\ &= (-x) + (x + z) && \text{hypothesis} \\ &= ((-x) + x) + z && \text{associative} \\ &= 0 + z && \text{inverse} \\ &= z && \text{identity.} \end{aligned}$$

□

ex:two11 **Example 2.2.** *Prove that for every $a \in \mathcal{F}$ we have $a.0 = 0$.*

Proof. We have

$$\begin{aligned} 0 + a.a &= a.a \\ &= (0 + a).a \\ &= 0.a + a.a \end{aligned}$$

where we have successively used the identity and distributive axioms. The conclusion then follows from the previous example. □

ex:two12 **Example 2.3.** *Prove that for every $x \in \mathcal{F}$ we have $(-x)^2 = x^2$.*

Proof. We have

$$\begin{aligned} (-x)^2 &= (-x)^2 + 0 \\ &= (-x)^2 + x.0 \\ &= (-x)^2 + x((-x) + x) \\ &= (-x)^2 + (x(-x) + x^2) \\ &= ((-x)^2 + x(-x)) + x^2 \\ &= ((-x) + x)(-x) + x^2 \\ &= 0.(-x) + x^2 \\ &= 0 + x^2 \\ &= x^2 \end{aligned}$$

where we have successively used the identity, the previous example, inverse, distributive, associative, distributive. identity, previous example, indentity axioms. □

Henceforward, apart perhaps from the odd exercise or exam question we will assume that any arithmetical operation we are used to is allowed.

def:two8 **Definition 2.2** (Order axioms for an ordered field)**.** *In an ordered field $\mathcal{F}$ there is a relationship $<$ between all elements which satisfies the following axioms.*

O1 *For every a and b in $\mathcal{F}$ exactly one of the following holds.*

$$a < b, a = b, b < a$$

O2 *If $a, b, c \in \mathcal{F}$, $a < b$ and $b < c$, then $a < c$.*

O3 *If $a, b, c \in \mathcal{F}$ and $a < b$, then $a + c < b + c$.*

O4 *If $a, b, c \in \mathcal{F}$, $a < b$ and $0 < c$, then $ac < bc$.*

We can now define more symbols

def:two9 **Definition 2.3.** *The symbol $a \leq b$ means $a < b$ or $a = b$.*

The symbol $a > b$ means $b < a$.

The symbol $a \geq b$ means $b \leq a$.

By **O1** every element a of $\mathcal{F}$ satisfies exactly one of

$$a < 0, a = 0, 0 < a.$$

The elements with $0 < a$ are called the positive numbers, and those with $a < 0$ are the negative numbers. These two sets, together with the set

$$\{0\}$$

partition $\mathcal{F}$ into three disjoint sets.

ex:two13 **Example 2.4.** *Show that if $0 < x$, then $-x < 0$, and that if $x < 0$, then $0 < -x$.*

Proof. By **O3** with $a = 0$, $b = x$, $c = -x$ we have

$$-x = 0 + (-x) < x + (-x) = 0,$$

the last equality by the definition of $-x$.

The second part is left as an exercise. □

ex:two14 **Example 2.5.** *Show that if $x \neq 0$, then $0 < x^2$.*

Remark. From this it follows that for *any* x we have $0 \leq x^2$.

Proof. There are two cases. 1. If $0 < x$, then by **O4** with $a = 0$, $b = c = x$ we have

$$0 = 0.x < x.x = x^2.$$

2. If $x < 0$, then by Example 2.4, $0 < -x$ and so by part 1. we have

$$0 < (-x)^2 = x^2.$$

□

We have not said anything about multiplication of inequalities by negative numbers. There is good reason for this because in the analogue of the inequality **O4** the *order* is *flipped*! This is one of the most common sources of mistakes in mathematics. However, we do not need a new axiom we can deduce the correct conclusion from the axioms we already have.

thm:two2 **Theorem 2.1.** *Suppose that $a < b$ and $c < 0$. Then*

$$bc < ac.$$

Proof. By Example 2.4 we have $0 < -c$. Hence, by **O4**,

$$-ac = a(-c) < b(-c) = -bc.$$

Now we add $ac + bc$ to both sides. Thus, by **O3**,

$$\begin{aligned} bc &= bc + 0 = bc + (ac + (-ac)) = (bc + ac) + (-ac) \\ &< (bc + ac) + (-bc) = (ac + bc) + (-bc) = ac + (bc + (-bc)) = ac + 0 \\ &= ac \end{aligned}$$

□

Another important consequence is the following theorem

thm:two3 **Theorem 2.2.** *We have*

$$0 < 1.$$

Proof. We have $1 \neq 0$. Hence $1 < 0$ or $0 < 1$. But then in either case $0 < 1^2 = 1$. □

ex:two15 **Example 2.6.** *Show that if $0 < a$, then $0 < a^{-1}$.*

Proof. We have $1 = a.a^{-1}$ so $a^{-1} \neq 0$ since otherwise we would have $a.a^{-1} = a.0 = 0$. Hence $0 < (a^{-1})^2$. By **O4**, since $0 < a$ we have

$$0 = a.0 < a.(a^{-1})^2 = (a.a^{-1})a^{-1} = 1.a^{-1} = a^{-1}.$$

□

ex:two16 **Example 2.7.** *Suppose that x and y are positive. Prove that $x < y$ if and only if $x^2 < y^2$.*

Proof. Note, we have *two* things to prove.

1. If $x < y$, then $x^2 < y^2$.
2. If $x^2 < y^2$, then $x < y$.

Proof of 1. We have $x < y$ and $0 < x$. Hence, by **O4**,

$$x^2 = x.x < xy$$

Likewise as $x < y$ and $0 < y$ we have

$$xy < y.y = y^2.$$

Then, by **O2**,

$$x^2 < xy < y^2$$

as required.

Proof of 2. We argue by contradiction. Thus we assume that the conclusion is false, i.e. $y \leq x$. There are two possibilities. First $y = x$. Then we would have $x^2 = y^2$ contradicting the hypothesis.

The second possibility is $y < x$. Then by the first part of the theorem we would have $y^2 < x^2$ which again contradicts the hypothesis. □

At this point it is convenient to remind ourselves of some standard notation for an interval, which makes sense once we have an ordering.

def:twoint **Definition 2.4.** *When $a \leq b$ we can define various kinds of intervals.*

$$\begin{aligned}
(a,b) &= \{x : a < x < b\} && \textit{an open interval,}\\
[a,b] &= \{x : a \leq x \leq b\} && \textit{a closed interval,}\\
[a,b) &= \{x : a \leq x < b\} && \textit{half closed - half open interval,}\\
(a,b] &= \{x : a < x \leq b\} && \textit{half open - half closed interval,}\\
(a,\infty) &= \{x : a < x\},\\
[a,\infty) &= \{x : a \leq x\},\\
(-\infty,b) &= \{x : x < b\},\\
(-\infty,b] &= \{x : x \leq b\}.
\end{aligned}$$

2.1.1 Exercises

subsec:two4

1. (i) Prove that if $1 < x$, then $x < x^2$.
   (ii) Prove that if $0 < x < 1$, then $x^2 < x$.

2. (i) Prove that if $a < b$, then $a < \frac{1}{2}(a+b) < b$.

   (ii) Prove that if $0 < a < b$, then $1/b < 1/a$.

3. Find all x such that
$$\frac{x+1}{x^2+3} < \frac{2}{x}.$$

4. Find all real values of x such that
$$\frac{x+5}{x^2+3} < \frac{2}{x}.$$

5. Find all real values of x such that
$$\frac{x+1}{x-1} < \frac{1}{x}.$$

6. Prove that if $0 < x$ and $0 < y$, then $x < y$ if and only if $x^3 < y^3$.

7. Determine which x belong to the set
$$\mathcal{A} = \left\{ x \in \mathbb{R} : \frac{2x+1}{x+2} < 1 \right\}.$$

8. Let a, b, α, β be real numbers with $b > 0, \beta > 0$ and
$$\frac{a}{b} < \frac{\alpha}{\beta}.$$

   (i) Prove that $a\beta < \alpha b$.

   (ii) Prove that
$$\frac{a}{b} < \frac{a+\alpha}{b+\beta} < \frac{\alpha}{\beta}.$$

9. Suppose that x and y satisfy $x < y$. Prove that $2x < x + y$ and $x + y < 2y$. Here $2x$ means $x + x$, of course.

10. Suppose that a and b satisfy $0 < ab$. Prove that either (i) $0 < a$ and $0 < b$, or (ii) $a < 0$ and $b < 0$.

11. Prove that $0 < x < 1$ if and only if $0 < x^3 < 1$.

12. Find all real values of x such that
$$\frac{x}{x+1} < \frac{x-1}{x+2}.$$

13. Determine which real numbers x belong to the set

$$\mathcal{A} = \left\{ x \in \mathbb{R} : \frac{2x+1}{x+2} < 1 \right\}.$$

14. Find all real values of x such that

$$\frac{x+1}{x^2+3} < \frac{2}{x}.$$

15. Suppose that $x \geq -1$.
Prove that (i) $(1+x)^2 \geq 1+2x$,
(ii) $(1+x)^3 \geq 1+3x$,
(iii) and that $(1+x)^4 \geq 1+4x$.
Make a guess as to a possible generalization concerning $(1+x)^n$ when n is a whole number larger than 4. (iv) Prove that $(1+x) \leq (1+\frac{1}{2}x)^2$.

16. Suppose that x is a real number with $0 < x < 1$.
(i) Prove that $1 < x^{-1} < x^{-2}$.
(ii) Prove that $x^{-2} < x^{-4}$.

17. Determine the set

$$\mathcal{A} = \left\{ x : \frac{x+3}{x^2+1} < \frac{2}{x} \right\}$$

18. Determine the set

$$\mathcal{A} = \left\{ x : \frac{x+5}{x^2+2} < \frac{2}{x} \right\}.$$

2.2 Inequalities

sec:two5

Inequalities are fundamental to analysis and it is desirable to obtain some facility in their manipulation. They can mostly be treated like equations except for the important caveat that multiplication can flip an inequality if the multiplicand is negative.

The following is very famous and frequently made use of.

thm:two4 **Theorem 2.3** (Cauchy)**.** *Suppose that x and y are elements of an ordered field. Then*

$$2xy \leq x^2 + y^2$$

Proof. By the remark following Example 2.5 we have

$$0 \leq (x-y)^2 = x^2 - 2xy + y^2.$$

Hence

$$2xy = 2xy + 0 \leq 2xy + x^2 - 2xy + y^2 = x^2 + y^2.$$

$\square$

Note that strictly speaking we should have divided the proof into two cases, one with $<$ and one with $=$, but with increasing familiarity there is less need to be so pedantic. The following is closely related albeit more complicated.

thm:two5

Theorem 2.4 (Cauchy-Schwarz). *Suppose that $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are $2n$ elements of an ordered field. Then*

$$(a_1b_1 + \cdots + a_nb_n)^2 \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

One reason this is important is because it tells us that in n-dimensional Euclidean space the scalar product of two vectors is bounded by the product of their sizes

Proof. Let

$$\begin{aligned} A &= a_1^2 + \cdots + a_n^2, \\ B &= a_1b_1 + \cdots + a_nb_n, \\ C &= b_1^2 + \cdots + b_n^2. \end{aligned}$$

If $A = 0$, then we have $a_1 = \cdots = a_n = 0$, since otherwise at least one of the terms in A is positive and the others are non-negative and by repeated use of the order axioms A would have to be positive. Thus if $A = 0$ both sides of the displayed formula in the theorem are 0. A fortiori we cannot have $A < 0$. Hence we may suppose that $A > 0$.

Let x be an element of the field which we will give a special value to later, and consider

$$\begin{aligned} Ax^2 + 2Bx + C &= a_1^2x^2 + 2a_1xb_1 + b_1^2 + a_2^2x^2 + 2a_2xb_2 + b_2^2 + \cdots a_n^2x^2 + 2a_nxb_n + b_n^2 \\ &= (a_1x + b_1)^2 + (a_2x + b_2)^2 + \cdots (a_nx + b_n)^2 \\ &\geq 0. \end{aligned}$$

Now multiply both sides by A. This gives

$$0 \leq A^2x^2 + 2ABx + AC = (Ax + B)^2 + AC - B^2.$$

Now take $x = -B/A$. Thus

$$0 \leq AC - B^2, \quad B^2 \leq AC$$

as required. $\square$

There are many different proofs of this.

2.2.1 Exercises

subsec:two5

1. Prove that $x^2 - x + 1 \geq \frac{3}{4}$.
2. Prove that $x^4 - 4x^2y^2 + 6y^4 \geq 0$.
3. Prove that $4abcd \leq a^4 + b^4 + c^4 + d^4$.
4. Let a, b, c, d be real numbers. Prove that $(ab + cd)^2 \leq (a^2 + c^2)(b^2 + d^2)$.
5. Let x and y be real numbers. Prove that $x^4 - 4x^2y^2 + 6y^4 \geq 0$
6. Let x and y be real numbers. Prove that $x^2 - xy + y^2 \geq 0$.

sec:two6

2.3 Absolute values

Before we can discuss anything connected with convergence we need to know what we mean by "small", or to be more precise we need to have some measure of the size of a number. The standard way for real numbers is as follows.

def:two10 **Definition 2.5** (Absolute Value)**.** *Let x be an element of an ordered field. Then we define the absolute value, or modulus, of x by*

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

ex:two17 **Example 2.8.**

$$|-\pi| = \pi, \left|\frac{3}{2}\right| = \frac{3}{2}, |0| = 0.$$

Note 1. That $|x| = 0$ if and only if $x = 0$, but for any $c \neq 0$ there are two choices of x with $|x| = c$, namely $x = \pm c$.

2. For every x we have $|x| \geq 0$.

3. For every x we have $|-x| = |x|$. To see this, separate out the three cases $x > 0$, $x = 0$, $x < 0$. When $x = 0$ we have $|-x| = |0| = 0 = |0| = |x|$. When $x > 0$ we have $-x < 0$ and so $|-x| = -(-x) = x = |x|$ and when $x < 0$ we have $-x > 0$ so that $|-x| = -x = |x|$.

thm:two6 **Theorem 2.5.** *For every x we have $-|x| \leq x \leq |x|$.*

Proof. Two cases.

1. If $x \geq 0$, then
$$-|x| \leq 0 \leq x = |x|.$$

2. If $x < 0$, then
$$-|x| = (-1)|x| = (-1)(-x) = x < 0 \leq |x|.$$

□

The very useful feature of the absolute value is that it preserves multiplicative structure.

thm:two7 **Theorem 2.6.** *Let a and b be elements of an ordered field. Then $|ab| = |a|.|b|$.*

Proof. As usual for the absolute value, this is a division into cases. There are two choices of sign for a and likewise for b, so there should be four cases.

1. $a \geq 0$, $b \geq 0$. Then $ab \geq 0$ so
$$|ab| = ab = a.b = |a|.|b|.$$

2. $a \geq 0$, $b < 0$. Then
$$|ab| = |-(ab)| = |a(-b)| = |a|.|-b| = |a|.|b|$$
3. $a < 0$, $b \geq 0$. Imitate 2. with a and b switched.
4. $a < 0$, $b < 0$. Then $ab > 0$ and
$$|ab| = ab = (-a)(-b) = |a|.|b|.$$

□

cor:two1 **Corollary 2.7.** *Suppose that $b \neq 0$. Then*
$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$$

Proof. We have
$$\left|\frac{a}{b}\right| |b| = \left|\frac{a}{b}b\right| = |a|.$$
Since $b \neq 0$ we have $|b| \neq 0$ and so we can divide both sides by $|b|$. □

Now we come to something we will use all the time.

thm:two8 **Theorem 2.8** (The Triangle Inequality)**.** *Suppose that x, y are elements of an ordered field. Then*
$$|x + y| \leq |x| + |y|.$$

Proof. We argue by contradiction. Suppose there are x and y so that $|x| + |y| < |x + y|$. Then
$$(|x| + |y|)^2 < |x + y|^2.$$
But by the definition of absolute value we have
$$\begin{aligned} |x+y|^2 &= (x+y)^2 = x^2 + 2xy + y^2 \\ &\leq x^2 + |2xy| + y^2 = |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2. \end{aligned}$$

□

ex:two18 **Example 2.9.**
$$|1 - 2| = |-1| = 1 \leq 3 = |1| + |2|.$$

There are several important generalisations of the triangle inequality.

thm:two9 **Theorem 2.9** (Generalised Triangle Inequality)**.** *Suppose that t and u are elements of an ordered field. Then*
$$\big||t| - |u|\big| \leq |t - u|.$$

Proof. By the triangle inequality

$$|t| = |t - u + u| \leq |t - u| + |u|.$$

Hence

$$|t| - |u| \leq |t - u|.$$

Interchanging t and u gives

$$|u| - |t| \leq |u - t| = |t - u|.$$

But one of $|t| - |u|$ and $|u| - |t| = -(|t| - |u|)$ is non-negative, so is

$$||t| - |u||.$$

□

ex:two19 **Example 2.10.** *Determine the set $\mathcal{A}$ of x such that $|2x + 3| < 7$*

Proof. The simple way is to use the definition of absolute value. There are two cases.

1. $2x + 3 \geq 0$. Then we also have $2x + 3 = |2x + 3| < 7$. Combining the two we need $-3/2 \leq x < (7 - 3)/2 = 2$. Thus in this case the inequality only holds when

$$-\frac{3}{2} \leq x < 2.$$

2. $2x + 3 < 0$. Now we have $-2x - 3 = |2x + 3| < 7$ so that $(-7 - 3)/2 < x < -3/2$. Thus in the second case the inequality only holds when

$$-5 < x < -3/2.$$

Combining the two cases we see that the inequality holds if and only if $-5 < x < 2$, so

$$\mathcal{A} = (-5, 2).$$

□

ex:two20 **Example 2.11.** *Find all x such that $|x + 3| + |x - 1| = 6$.*

Proof. The simple way is to look at the four possible cases for the absolute values.

1. $x + 3 \geq 0$ and $x - 1 \geq 0$. Then we have $x \geq -3$ and $x \geq 1$ so we are forced to take $x \geq 1$. Then the equation becomes

$$2x + 2 = x + 3 + x - 1 = 6, \; x = 2.$$

2. $x + 3 \geq 0$ and $x - 1 < 0$. Then $x \geq -3$ and $x < 1$ so we are restricted to $-3 \leq x < 1$. Then the equation is

$$4 = x + 3 - (x - 1) = 6$$

which is impossible, so no solutions in this case.

3. $x+3<0$ and $x-1\geq 0$. Now we would have $1\leq x<-3$ which is impossible, so no solutions in this case.

4. $x+3<0$ and $x-1<0$. This requires $x<-3$ and $x<1$, so it forces $x<-3$. Then the equation becomes

$$-2x-2=-(x+3)-(x-1)=|x+3|+|x-1|=6,\ x=-4.$$

Hence the complete solution is

$$x=-4 \text{ or } 2.$$

$\square$

2.3.1 Exercises

subsec:two6

1. (i) Prove that, for any real number a, $|a|^2=a^2$.

(ii) Let a and b be real numbers. Show that $|a+b|=|a|+|b|$ if and only if $ab\geq 0$.

2. Suppose that a, b, x, y are real numbers satisfying $a<x<b$ and $a<y<b$. Show that $|x-y|<b-a$.

3. Sketch the graph of the equation $y=|x|-|x-1|$.

4. Find all x such that $|x+1|+|x-2|=7$.

5. Find all real numbers x that satisfy the inequality $4<|x+2|+|x-1|<5$.

6. Sketch the set of pairs (x,y) that satisfy $|x|+|y|=1$.

7. Let x, y, z be real numbers with $x\leq z$. Prove that $x\leq y\leq z$ if and only if $|x-y|+|y-z|=|x-z|$.

8. Let x, y, z satisfy $x\leq z$. Prove that $x\leq y\leq z$ if and only if $|x-y|+|y-z|=|x-z|$.

9. Sketch the graph of the equation $y=|x+1|-|x-2|$.

10. Find all x such that $|x+3|+|x-3|=8$.

11. Find all real numbers x that satisfy the inequality $4<|x+2|+|x|<6$.

12. Sketch the set of pairs (x,y) which satisfy $3|x|+2|y|=5$.

13. Find all real numbers x that satisfy $-1<2|x-1|-|3x+2|<1$

14. Find all real numbers x that satisfy the inequality $4<|x+2|+|x-1|<5$.

15. Let x, a, ε be real numbers with $\varepsilon>0$. Show that $|x-a|<\varepsilon$ if and only if $a-\varepsilon<x<a+\varepsilon$.

16. Sketch the graph of the equation $y=|x+2|-|x-1|$.

17. Find all real numbers x that satisfy the inequality $4<|x+1|+|x-1|<6$.

18. (i) Prove that, for any real number a, $|a|^2 = a^2$.
    (ii) Let a and b be real numbers. Show that $|a+b| = |a| + |b|$ if and only if $ab \geq 0$.

19. Sketch the graph of the equation $y = |x| - |x-1|$.

20. Find all x such that $|x+1| + |x-2| = 7$.

21. Sketch the set of pairs (x, y) that satisfy $|x| + |y| = 1$.

22. Let x, y, z be real numbers with $x \leq z$. Prove that $x \leq y \leq z$ if and only if $|x-y| + |y-z| = |x-z|$.

23. Suppose that a is a real number and $\varepsilon > 0$. Prove that

$$\{x : |x-a| \leq \varepsilon\} = [a-\varepsilon, a+\varepsilon].$$

24. Find all real x that satisfy

$$2 < |x-1| + |x+1| < 4$$

25. Find all real numbers x such that

$$\big||x-1| - |x|\big| < \frac{1}{2}.$$

2.4 The Continuum Property

sec:two7

We have already seen that it is possible to use ordered pairs to construct the integers from the natural numbers and then the rational numbers from the integers. Because we have to somehow build in limiting processes to obtain the real numbers we have to do something more sophisticated. There are several different ways of doing this. The approach we choose is essentially due to Dedekind. In place of ordered pairs we should, at least initially think of real numbers as being infinite sets of rational numbers. Thus we could think of $\sqrt{2}$ as being

$$\text{“}\sqrt{2}\text{”} = \{a : a \in \mathbb{Q}, a > 0, a^2 < 2 \text{ or } a \leq 0\}.$$

In other words we think of $\sqrt{2}$ as being the set of all rational numbers to the left of where we expect $\sqrt{2}$ to be. Then we need to show that these new objects, namely sets of rational numbers, can be made to satisfy all the previous axioms.

In order to do this systematically we need to set up some language. In what follows we should think of the various definitions as giving us language we can use once we are satisfied that constructions such as that above give us a set of real numbers $\mathbb{R}$ with the required properties.

def:two11 **Definition 2.6.** *A set $\mathcal{S}$ of real numbers is* **bounded above** *when there exists a real number H such that for every $x \in \mathcal{S}$ we have $x \leq H$.*

Any such number H is called an upper bound for $\mathcal{S}$.

ex:two22 **Example 2.12.** *Let* $\mathcal{S} = \{-3/2, \pi, 19\}$. *Then* 19, 19.1, 20, 100, 10^{60} *are all upper bounds for* $\mathcal{S}$.

There is a corresponding definition of *bounded below.*

def:two12 **Definition 2.7.** *A set* $\mathcal{S}$ *of real numbers is* **bounded below** *when there exists a real number* h *such that for every* $x \in \mathcal{S}$ *we have* $h \leq x$.

Any such number h is called a lower bound for $\mathcal{S}$.

def:two13 **Definition 2.8.** *A set* $\mathcal{S}$ *of real numbers which is both bounded above and bounded below is called* **bounded**. *If it is not bounded, then it is called unbounded.*

The set $\mathcal{S}$ of Example 2.12 is bounded below and bounded. The set $\mathbb{N}$ is unbounded (presumably - later we will prove this).

ex:two23 **Example 2.13.** *1.* $\{\sin x : x \in \mathbb{R}\}$ *is bounded because* $-1 \leq \sin x \leq 1$ *for every* x.

2. $\{x^2 : x \in \mathbb{R}\}$ *is bounded below but unbounded.*

3. $\mathcal{A} = \{x : x^2 - 3x + 2 < 0\}$ *is interesting. It is the set of* x *for which the polynomial* $x^2 - 3x + 2 = (x-1)(x-2)$ *is negative. The factorisation shows that it is only negative when* $1 < x < 2$. *Hence the set* $\mathcal{A}$ *is bounded with* 1 *as a lower bound and* 2 *as an upper bound..*

We have already suggested above that real numbers like $\sqrt{2}$ can be constructed through the use of a set which in some sense is the set of all rational numbers to the left of $\sqrt{2}$. Here is another example.

ex:two24 **Example 2.14.** *Can we assign a meaning to*

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots ?$$

Look at the sum S_n *after* n *terms, so that*

$$\begin{aligned}
S_1 &= 1, \\
S_2 &= 1 + \frac{1}{2^2}, \\
S_3 &= 1 + \frac{1}{2^2} + \frac{1}{3^2}, \\
S_4 &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2}, \\
&\vdots, \\
S_n &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots + \frac{1}{n^2}, \\
&\vdots
\end{aligned}$$

Obviously

$$S_1 < S_2 < S_3 < \ldots < S_n < \ldots .$$

Let

$$\mathcal{A} = \{S_1, S_2, S_3, \ldots, S_n, \ldots\}$$

Suppose that $\mathcal{A}$ is bounded above, so there are real numbers y such that $S_n \leq y$ for every n. Let x be the smallest such number. Then surely this means that the series is converging to x? Oh, but perhaps there is no smallest such number! Well surely there should be. The job of the axiom we are missing is to ensure that there is always a smallest such number.

By the way,

$$\begin{aligned} S_n &= 1 + \frac{2}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots + \frac{1}{n^2} \\ &\leq 1 + \frac{1}{2.1} + \frac{1}{2.3} + \frac{1}{3.4} + \cdots + \frac{1}{(n-1)n} \\ &= 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 2 - \frac{1}{n} \\ &< 2, \end{aligned}$$

so the set $\mathcal{A}$ is bounded above by 2. Moreover the series is well known and converges to

$$1 + \frac{2}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6}.$$

Thus we can now state the axiom which distinguishes the real numbers from the rational numbers.

def:two14 **Definition 2.9** (The Continuum Property)**.** *Every non-empty subset $\mathcal{S}$ of $\mathbb{R}$ which is bounded above has a* **least upper bound***, also called a* **supremum***, and we denote it by* $\sup \mathcal{S}$.

ex:two25 **Example 2.15.** *Here are some examples*

1. $\sup\{1, 2, 3\} = 3$.
2. $\sup(1, 2) = 2$.
3. $\sup(0, \infty)$ *does not exist.*
4. $\sup\left\{\frac{1}{2}, \frac{3}{4}, \ldots, 1 - \frac{1}{2^n}, \ldots\right\} = 1$.

ex:two26 **Example 2.16.** *Suppose that $\mathcal{A}$ is s non-empty set of real numbers which is bounded above. The* $\sup \mathcal{A}$ *is unique.*

Proof. *Suppose that $s_1 < s_2$ are two different suprema of $\mathcal{A}$. By the definition of supremum we have $a \leq s_1$ for every $a \in \mathcal{A}$ and so s_2 could not be a least upper bound.*

It is sometimes useful to deal with (non-empty) sets which are bounded below rather than bounded above. The corresponding term is *infimum*. Fortunately we do not need yet another axiom.

thm:two10 **Theorem 2.10.** *Suppose that $\mathcal{B}$ is a non-empty set of real numbers which is bounded below. Then $\mathcal{B}$ has a greatest lower bound, the infimum of $\mathcal{B}$.*

Proof. Let
$$\mathcal{A} = \{-b : b \in \mathcal{B}\}$$
and let h be a lower bound for $\mathcal{B}$, so that $h \le b$ for every $b \in \mathcal{B}$. Then by Theorem 2.1, $-b = (-1)b \le (-1)h = -h$ for every $b \in \mathcal{B}$. Hence $-h$ ia an upper bound for $\mathcal{A}$, i.e. $\mathcal{A}$ is non-empty and bounded above. Thus it has a supremum, s. Now we show that $-s$ acts as an infimum of $\mathcal{B}$. Clearly $s \ge -b$ for every $b \in \mathcal{B}$ and so $-s \le b$ for every b in $\mathcal{B}$. We show that there can be no larger lower bound. Suppose on the contrary that there is a $t > -s$ such that t is a lower bound for $\mathcal{B}$. i.e. for every $b \in \mathcal{B}$. Then $-b \le -t < s$. Thus $-t$ would be a lower upper bound for $\mathcal{A}$ than its supremum s which is absurd. □

Before moving on to study the properties of the real numbers we just give an inkling of how it is possible to pull over to $\mathbb{R}$ the various axioms which are satisfied by $\mathbb{Q}$

thm:two11 **Theorem 2.11.** *Suppose that $\mathcal{A}$ is a non-empty set of real numbers which is bounded above, $y > 0$ and*
$$\mathcal{B} = \{ya : a \in \mathcal{A}\}.$$
Then $\sup \mathcal{B}$ exists and
$$\sup \mathcal{B} = y \sup \mathcal{A}.$$

Proof. Since $\mathcal{A}$ is non-empty, so is $\mathcal{B}$. Moreover if H is an upper bound for $\mathcal{A}$, then yH is an upper bound for $\mathcal{B}$. Hence $s = \sup \mathcal{A}$ and $t = \sup \mathcal{B}$ both exist. Moreover sy will be an upper bound for $\mathcal{B}$ and t/y will be an upper bound for $\mathcal{A}$. Hence
$$t \le sy \text{ and } s \le t/y \le s,$$
whence
$$t = ys.$$
□

ex:two27 **Example 2.17.** *Let $y \in \mathbb{R}$, $\mathcal{S} \subset \mathbb{R}$, $\mathcal{S} \ne \emptyset$ and suppose that $\mathcal{S}$ is bounded above. Let $\mathcal{T} = \{x + y : x \in \mathcal{S}\}$. Then $\sup \mathcal{T}$ exists and*
$$\sup \mathcal{T} = y + \sup \mathcal{S}.$$

Proof. Let $s = \sup \mathcal{S}$. Since $\mathcal{S}$ is non-empty, then so is $\mathcal{T}$. We have to show two things.
1. $\mathcal{T}$ is bounded above by $y + s$, so that $t = \sup \mathcal{T}$ exists and $t \le y + s$, and
2. $\mathcal{S}$ is bounded above by $-y + t$, so that $s \le -y + t$.

Proof of 1. Let $v \in \mathcal{T}$. Then there is a $u \in \mathcal{S}$ so that $v = y + u$. Thus $v \leq y + s$. Since this holds for every $v \in \mathcal{T}$ and u is bounded above by s it follows that $\mathcal{T}$ is bounded above by $y + s$ and so $t \leq y + s$.

Proof of 2. We invert this argument. Let $u \in \mathcal{S}$. Then $y + u \in \mathcal{T}$. Hence $y + u \leq t$ and so $u \leq -y + t$. This holds for every $u \in \mathcal{S}$. Thus $-y + t$ is an upper bound for $\mathcal{S}$. Therefore $s \leq -y + t$. □

Now we have the machinery to prove the existence of square roots, cube roots it is necessary to make the following general definition.

def:two15 **Definition 2.10.** *Given any real number x we define x^n for $n \in \mathbb{N}$ inductively by $x_1 = x$ and $x^{n+1} = x^n.x$. When $x \neq 0$ we can extend this by defining $x^{-n} = (1/x)^n$. When $x \neq 0$ we take $x^0 = 1$. Except in peculiar circumstance which we will mention later we do not define 0^0.*

Given a non-negative real number x for $n \in \mathbb{N}$ we define $x^{1/n}$ to be that positive real number y such that $y^n = x$ and extend this to non-zero integers by taking $x^{1/(-n)} = (1/x)^{1/n}$. Then when m is a non-zero integer we can extend this further to rational exponents by taking

$$x^{\frac{m}{n}} = (x^m)^{1/n}.$$

It is then possible to check the standard rules for exponents such as

$$\left(x^{\frac{m}{n}}\right)^{\frac{q}{r}} = x^{\frac{mq}{nr}}, \quad x^{\frac{m}{n}} y^{\frac{m}{n}} = (xy)^{\frac{m}{n}},$$

$$x^{\frac{m}{n}} x^{\frac{q}{r}} = x^{\frac{mr+nq}{nr}}.$$

2.4.1 Exercises

subsec:two7

1. Let $\mathcal{A} = \{x : 2x - x^2 > 0\}$. Prove that this set is bounded above. Is it bounded below?

2. Decide in each of the following cases whether or not the given set is bounded above. For those which are bounded above give three different upper bounds including the smallest one.

   (i) $\{-2, 0, 2, 4, 7, 8, 20\}$, (ii) $[2, \infty)$, (iii) $(-\infty, 2)$, (iv) $[1, 2]$, (v) $(1, 2)$.

3. Give an example of a set which has least upper bound 1 but contains no element x satisfying $x < 1$.

4. Let a be any element of the open interval $(0, 1)$.

   (i) Show that there is another $b \in (0, 1)$ with $b > a$.

   (ii) Prove that $(0, 1)$ has no maximum.

5. Suppose that $\mathcal{A}$ is bounded above, $\mathcal{B} \subset \mathcal{A}$ and $\mathcal{B}$ is non-empty. Prove that $\sup \mathcal{B}$ exists and $\sup \mathcal{B} \leq \sup \mathcal{A}$.

6. Let $\mathcal{A} = \{x : x + x^2 < 0\}$. Prove that this set is non-empty and bounded above. What is the least upper bound? Is it bounded below?

7. Let $\mathcal{A}$, $\mathcal{B}$ be non–empty sets of real numbers which are bounded above, and let $\mathcal{A} + \mathcal{B}$ denote the set of numbers of the form $a + b$ with $a \in \mathcal{A}$ and $b \in \mathcal{B}$.

(i) Prove that $\sup(\mathcal{A} + \mathcal{B})$ exists.

(ii) Prove that $\sup(\mathcal{A} + \mathcal{B}) \leq \sup \mathcal{A} + \sup \mathcal{B}$.

(iii) Let $\delta > 0$. Prove that there are $a \in \mathcal{A}$ and $b \in \mathcal{B}$ such that $a > \sup \mathcal{A} - \delta$ and $b > \sup \mathcal{B} - \delta$.

(iv) Deduce that $\sup(\mathcal{A} + \mathcal{B}) = \sup \mathcal{A} + \sup B$.

8. Suppose that $\mathcal{A}$ is a non-empty set of real numbers which is bounded above, $y < 0$ and

$$\mathcal{B} = \{ya : a \in \mathcal{A}\}.$$

Then prove that $\inf \mathcal{B}$ exists and

$$\inf \mathcal{B} = y \sup \mathcal{A}.$$

9. Let $\mathcal{S} = \left\{2 + \frac{1}{\sqrt{n}} : n \in \mathbb{N}\right\}$.

(i) Prove that $\mathcal{S}$ is non-empty and bounded below by 2.

(ii) Prove that if a is a real number with $a > 2$, then there is an $n \in \mathbb{N}$ such that $2 + \frac{1}{\sqrt{n}} < a$.

(iii) Prove that $\inf \mathcal{S} = 2$.

10. Let $S = \{x \in \mathbb{Q}, x > 0, x^2 < 3\}$, $T = \{y \in \mathbb{Q}, y > 0, y^2 > 3\}$. Prove that

(i) $a = \sup S$ exists,

(ii) $a^2 \leq 3$ (hint: Suppose $a^2 > 3$ and choose $\delta = \frac{a^2-3}{2a}$ and an $x \in \mathbb{Q}$ with $a - \delta < x < a$),

(iii) $b = \inf T$ exists and $3 \leq b^2$ (suppose $b^2 < 3$ and choose $\delta = \min\{1, \frac{3-b^2}{2b+1}\}$) and hence $a \leq b$.

(iv) Prove that there is no rational number r with $a < r < b$.

(v) Deduce that $a = b$ and $a^2 = b^2 = 3$.

11. Simplify the following.

(i) $64^{2/3}$,

(ii) $3125^{1/5}$,

(iii) $27^{-4/3}$. 12. Simplify the following.

(i) $16^{-3/4}$, (ii) $243^{1/5}$, (iii) $125^{2/3}$.

13. Decide in each of the following cases whether or not the given set is bounded above. For those which are bounded above give three different upper bounds including the smallest one.

(i) $\{-4, -2, 1, 5, 6, 19\}$, (ii) $[-2, \infty)$, (iii) $(-\infty, -5)$, (iv) $[-17, 31]$, (v) $(12, 13)$.

14. Give an example of a set which has least upper bound 5 but contains no element x satisfying $x < 5$.

15. Let $\mathcal{A} = \{x : x^2 + 4x + 3 < 0\}$. Prove that this set is non-empty and bounded above. What is the least upper bound? Is it bounded below?

16. Let a be any element of the open interval $(0, 1)$.
  (i) Show that there is another $b \in (0, 1)$ with $b > a$.
  (ii) Prove that $(0, 1)$ has no maximum.

17. Suppose that $\mathcal{A}$ is bounded above, $\mathcal{B} \subset \mathcal{A}$ and $\mathcal{B}$ is non-empty. Prove that $\sup \mathcal{B}$ exists and $\sup \mathcal{B} \le \sup \mathcal{A}$.

18. Let $\mathcal{A} = \{x : x \in \mathbb{Q}, x^3 < 2\}$. Prove that $\sup \mathcal{A}$ exists. Guess the value of $\sup \mathcal{A}$.

19. Let $\mathcal{A} = \{x : 2x - x^2 > 0\}$. Prove that this set is bounded above. Is it bounded below?

20. Which of the following statements is true?
  (i) $4 \in (-5, 3)$, (ii) $2 \in (1, \infty]$, (iii) $3 \in (3, 4)$, (iv) $2 \in [2, 3]$, (v) $-1 \in [-1, 1)$.

21. Suppose that $a \in \mathbb{R}$ and $|a| < \varepsilon$ for every $\varepsilon > 0$. Prove that $a = 0$.

2.5 Notes

sec:two8

§2.4 For the work of Dedekind and others on the construction of real numbers, see
`https://en.wikipedia.org/wiki/Dedekind_cut`

Chapter 3

The Natural Numbers

ch:three

3.1 The Archmidean Property

sec:three1

We have seen that the natural numbers $\mathbb{N}$ are embedded in $\mathbb{Z}$ ("n" is the equivalence class $\mathcal{A}(n+1,1)$) and that is embedded in $\mathbb{Q}$ which in turn is embedded in $\mathbb{R}$. We now see what impact the Continuum property has on $\mathbb{N}$.

thm:three1

Theorem 3.1 (Archimedean Property)**.** *The set $\mathbb{N}$ is unbounded above.*

It is perhaps surprising that the continuum property makes a crucial contribution.

Proof. It is immediate from the fact that $0 < 1$ and the principle of induction that $\mathbb{N}$ is bounded below by 1. Thus it remains to show that $\mathbb{N}$ is unbounded above. We argue by contradiction.

Suppose $\mathbb{N}$ is bounded above. Since $1 \in \mathbb{N}$ we have $\mathbb{N} \neq \emptyset$. Thus $B = \sup \mathbb{N}$ exists. Then $B - 1$ is not an upper bound of $\mathbb{N}$. Hence there is an element n of $\mathbb{N}$ such that $B - 1 < n$. But $n + 1 \in \mathbb{N}$ and $B < n + 1$ gives a contradiction. $\square$

There are many ways in which we can use this.

ex:three1

Example 3.1. *Let*

$$\mathcal{A} = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\}.$$

Then $\mathcal{A}$ is bounded, and $\inf \mathcal{A} = 0$, $\sup \mathcal{A} = 1$.

Proof. We have $1/1 = 1 \in \mathcal{A}$, so $\mathcal{A} \neq \emptyset$.

Since $n \geq 1$ for $n \in \mathbb{N}$ we have $1/n \leq 1$. Hence $\mathcal{A}$ is bounded above by 1 and as $1 \in \mathcal{A}$ there cannot be any smaller upper bound. Thus $\sup \mathcal{A} = 1$.

We also have $n \geq 1 > 0$. Thus $1/n > 0$ also, so 0 is a lower bound for $\mathcal{A}$. Hence $\inf \mathcal{A}$ exists. We now show that there is no larger lower bound. We argue by contradiction. Let $b = \inf \mathcal{A}$ and suppose that $b > 0$. Then for every $n \in \mathbb{N}$ we have $b \leq 1/n$. Hence $n \leq 1/b$ which contradicts the Archimedean property.

This is the first instance in this text of a "limiting" process, and the connection with the Archimedean property, and through that the continuum property is crucial. We have not yet defined what we mean by a limit, but it suggests that such a concept is already built in to the definition of the continuum property. □

ex:three1a **Example 3.2.** *Let*

$$\mathcal{B} = \left\{ \frac{2n}{3n-1} : n \in \mathbb{N} \right\}.$$

Then $\mathcal{B}$ *is bounded, and* $\inf \mathcal{B} = \frac{2}{3}$, $\sup \mathcal{B} = 1$.

Proof. We first deal with the upper bound. Such proofs should be divided into three parts. (i) Prove that $\mathcal{B} \neq \emptyset$, (ii) prove that 1 is an upper bound, and (iii) prove that there is no smaller upper bound.

(i) We have $1 = \frac{2\times 1}{3\times 1-1} \in \mathcal{B}$, so $1 \in \mathcal{B}$ and $\mathcal{B} \neq \emptyset$.

(ii) Since $n \geq 1$ for $n \in \mathbb{N}$ we have $3n-1 = 2n+n-1 \geq 2n$. Hence $\frac{2n}{3n-1} \leq \frac{2n}{2n} = 1$ and so 1 is an upper bound.

(iii) As $1 \in \mathcal{B}$ there can be no smaller upper bound.

We can deal with the lower bound in the same kind of way. We have already established (i) above, so we do not have to do it again! It remains to show (ii) that $\frac{2}{3}$ is a lower bound and (iii) that there is no larger lower bound.

(ii) We have $3n-1 < 3n$ for each $n \in \mathbb{N}$, so that $\frac{2n}{3n-1} > \frac{2n}{3n} = \frac{2}{3}$. Hence $\frac{2}{3}$ is a lower bound for the set. Thus $b = \inf \mathcal{B}$ exists and $b \geq \frac{2}{3}$.

(iii) We now prove that $b = \frac{2}{3}$. This is the trickiest part of the question. We argue by contradiction, so suppose on the contrary that $b > \frac{2}{3}$, so that $b - \frac{2}{3} > 0$. By the Archimedean property we can choose an $n \in \mathbb{N}$ so that

$$n > \frac{b}{3b-2}.$$

Then, as $b > 2/3$,

$$\begin{aligned} 3bn - 2n = n(3b-2) &> b, \\ b(3n-1) = 3bn - b &> 2n, . \end{aligned}$$

so that

$$\frac{2n}{3n-1} < b.$$

Alternative proof. We could argue as follows. Suppose as before that $b > \frac{2}{3}$. Then for every $n \in \mathbb{N}$ we have

$$b \leq \frac{2n}{3n-1}.$$

Solve for n. We have $3bn - b \leq 2n$ and this can be rearranged to give $(3b-2)n \leq b$. Since $b > \frac{2}{3}$ this gives

$$n \leq \frac{b}{3b-2}$$

contradicting the Archimedean property. □

Here is something which one might think is self evident, but which nevertheless requires proof.

thm:three2 **Theorem 3.2.** *Every non-empty subset of* $\mathbb{N}$ *has a minimum.*

Before embarking on the proof we should be clear what we mean by the maximum or minimum of a set.

def:three1 **Definition 3.1.** *When a set* $\mathcal{A}$ *of real numbers has the property that it has a lower bound with* $m \in \mathcal{A}$*, then we say that* m *is the* **minimum** *of* $\mathcal{A}$*. When a set* $\mathcal{B}$ *of real numbers has the property that it has an upper bound* M *with* $M \in \mathcal{B}$*, then we say that* M *is the* **maximum** *of* $\mathcal{B}$*.*

ex:three2 **Example 3.3.** *1. The set* $\mathbb{N}$ *has* 1 *as its minimum.*

2. The open interval $(2, 3)$ *has neither a maximum nor a minimum.*

3. The closed interval $[1, 2]$ *has* 1 *as its minimum and* 2 *as its maximum.*

4. Note that 2. shows that, even when a set has an infimum or a supremum, that does not guarantee that it has a corresponding minimum or maximum. In other words, extrema may not be members of the set.

Now we return to

Proof of Theorem 3.2. Let $\mathcal{A}$ be a non-empty subset of $\mathbb{N}$. Every element of $\mathbb{N}$ is bounded below by 1 so $\mathcal{A}$ is bounded below. Let $b = \inf \mathcal{A}$. Then $b + 1$ is not a lower bound. Hence there is an element n of $\mathcal{A}$ such that $n < b + 1$. If $m \geq n$ for every element m of $\mathcal{A}$, then n is a lower bound of $\mathcal{A}$ *and* $n \in \mathcal{A}$, and we would be done. If there would be an element m of $\mathcal{A}$ with $m < n$, then we would have $m + 1 \leq n < b + 1$, so that m would satisfy $m < b$ which is impossible because b is a lower bound for $\mathcal{A}$. □

This principle can be extended to $\mathbb{Z}$.

ex:three3 **Example 3.4.** *Every non-empty subset of* $\mathbb{Z}$ *which is bounded below has a minimum.*

Proof. Let $\mathcal{S}$ be the set and let b be a lower bound for S. By the Archimedean property there is an $n \in \mathbb{N}$ such that $n > -b$. Let $\mathcal{T} = \{n + s : s \in \mathcal{S}\}$. For each $s \in \mathcal{S}$ we have $s + n > b + (-b) = 0$, so that $s + n \geq 1$. Hence $\mathcal{T}$ is a subset of $\mathbb{N}$ and so by Theorem 3.2 has a minimum m. Thus $m \in \mathcal{T}$, so that and $m \leq s + n$ for every $s \in \mathcal{S}$ and $m = s_0 + n$ for some $s_0 \in \mathcal{S}$. Thus $m - n \leq s$ for every $s \in \mathcal{S}$ and $m - n = s_0 \in \mathcal{S}$. Thus $m - n$ is the minimum for $\mathcal{S}$. □

Another consequence of the Archimedean property is that the rationals are everywhere dense amongst the real numbers. Here is a simple way of expressing that

thm:three3 **Theorem 3.3.** *Suppose that* a *and* b *are real numbers with* $a < b$*. Then there is a rational number* r *with*

$$a < r < b.$$

By a rational number we mean $r = \frac{m}{n}$ with $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. We also use the term *irrational* to mean a real number which is not rational.

Proof. First we find a suitable n. Since $a < b$ we have $0 < b - a$ and so

$$0 < \frac{1}{b-a}.$$

By the Archimedean property there is an $n \in \mathbb{N}$ such that $n > \frac{1}{b-a}$. Hence $n(b-a) > 1$, and so

$$1 + na < nb.$$

Let $\mathcal{A}$ be the subset of $\mathcal{Z}$,

$$\mathcal{A} = \{\ell \in \mathbb{Z} : \ell > na\}.$$

Then from above $\mathcal{A}$ has a minimal element. Call it m. If $m \geq nb$, then we would have

$$m - 1 \geq nb - 1 > na,$$

so we would have $m - 1 \in \mathcal{A}$ contradicting the minimality of m. Thus $na < m < nb$ and dividing by n gives the desired conclusion. □

3.1.1 Exercises

subsec:three1

1. Let

$$\mathcal{S} = \left\{ \frac{2n-1}{n+1} : n \in \mathbb{N} \right\}.$$

(i) Prove that $\mathcal{S} \neq \emptyset$ and $\mathcal{S}$ is bounded.
(ii) Prove that $\inf \mathcal{S} = \frac{1}{2}$.
(iii) Prove that $\sup \mathcal{S} = 2$.

2. Let $\mathcal{U} = \left\{ \frac{2n+1}{n+1} : n \in \mathbb{N} \right\}$.
(i) Prove that $\mathcal{U}$ is non-empty and bounded above by 2.
(ii) Prove that if a is a real number with $a < 2$, then there is an $n \in \mathbb{N}$ such that $a < \frac{2n+1}{n+1}$.
(iii) Prove that $\sup \mathcal{U} = 2$.

3. (i) Suppose that $\mathcal{A}$ is a non-empty subset of $\mathbb{N}$ which is bounded above. Then $\mathcal{A}$ has a maximum.
(ii) Suppose that $\mathcal{B}$ is a non-empty subset of $\mathbb{N}$ which is bounded above. Then $\mathcal{B}$ has a maximum.

4. Suppose that a, b are real numbers with $a < b$. Prove that there is an **irrational** number r with $a < r < b$.

5. Prove also that there is an irrational number b such that $2 < b^2 < 3$.

6. Let $\mathcal{U} = \left\{ \frac{2n+1}{n+1} : n \in \mathbb{N} \right\}$.

(i) Prove that $\mathcal{U}$ is non-empty and bounded above by 2.

(ii) Prove that if a is a real number with $a < 2$, then there is an $n \in \mathbb{N}$ such that $a < \frac{2n+1}{n+1}$.

(iii) Prove that $\sup \mathcal{U} = 2$.

7. Let $\mathcal{S} = \left\{2 + \frac{1}{\sqrt{n}} : n \in \mathbb{N}\right\}$.

(i) Prove that $\mathcal{S}$ is non-empty and bounded below by 2.

(ii) Prove that if a is a real number with $a > 2$, then there is an $n \in \mathbb{N}$ such that $2 + \frac{1}{\sqrt{n}} < a$.

(iii) Prove that $\inf \mathcal{S} = 2$.

8. Prove that there is an irrational number a such that $2 < a < 3$. Prove also that there is an irrational number b such that $2 < b^2 < 3$.

9. Prove that there is an irrational number a such that $2 < a < 3$. Prove also that there is an irrational number b such that $2 < b^2 < 3$.

10. Let $S = \{x \in \mathbb{Q}, x > 0, x^2 < 3\}$, $T = \{y \in \mathbb{Q}, y > 0, y^2 > 3\}$. Prove that

(i) $a = \sup S$ exists,

(ii) $a^2 \leq 3$ (hint: Suppose $a^2 > 3$ and choose $\delta = \frac{a^2-3}{2a}$ and an $x \in \mathbb{Q}$ with $a - \delta < x < a$),

(iii) $b = \inf T$ exists and $3 \leq b^2$ (suppose $b^2 < 3$ and choose $\delta = \min\{1, \frac{3-b^2}{2b+1}\}$) and hence $a \leq b$.

(iv) Prove that there is no rational number r with $a < r < b$.

(v) Deduce that $a = b$ and $a^2 = b^2 = 3$.

11. Let $\mathcal{S} = \left\{1 + \frac{1}{n} : n \in \mathbb{N}\right\}$. Prove that $\inf \mathcal{S}$ exists and $\inf \mathcal{S} = 1$.

12. Let $\mathcal{U} = \left\{1 - \frac{2}{\sqrt{n}} : n \in \mathbb{N}\right\}$.

(i) Prove that $\mathcal{U}$ is non-empty and bounded above by 1.

(ii) Prove that if a is a real number with $a < 1$, then there an $n \in \mathbb{N}$ such that $a < 1 - \frac{2}{\sqrt{n}}$.

(iii) Prove that $\sup \mathcal{U} = 1$.

13. Suppose that a and b are real numbers with $a < b$. Prove that there is an irrational number c such that $a < c < b$.

14. Let $\mathcal{A} = \left\{2 - \frac{1}{n} : n \in \mathbb{N}\right\}$. Prove that $\sup \mathcal{A}$ exists and $\sup \mathcal{A} = 2$.

15. Let $\mathcal{A} = \left\{\frac{3}{n} : n \in \mathbb{N}\right\}$.

(i) Prove that $\inf \mathcal{A}$ and $\sup \mathcal{A}$ exist.

(ii) Prove that $\inf \mathcal{A} = 0$.

(iii) Prove that $\sup \mathcal{A} = 3$.

(iv) Is $0 \in \mathcal{A}$?

3.2 The Principle of Induction

sec:three2

I want to say something more about the principle of induction and look at some applications which are a bit different from the kind that are normally met when the principle is introduced.

We recall from Definition 1.6 4. that the principle of induction says that if $\mathcal{S}$ is a set with the properties that (a) $1 \in \mathcal{S}$ and (b) whenever $n \in \mathcal{S}$ we have $n+1 \in \mathcal{S}$, then $\mathcal{S} = \mathbb{N}$.

It is often convenient to think of our set $\mathcal{S}$ as having some kind of defining statement for n to be an element. That is, there is some proposition or statement $P(n)$ which we would like to prove is true for every $n \in \mathbb{N}$. Then we take

$$\mathcal{S} = \{n : P(n) \text{ is true}\}.$$

Thus if we can show that

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

(ii) whenever $P(n)$ is true $P(n+1)$ is also true, then it follows that $\mathcal{S} = \mathbb{N}$ and $P(n)$ is true for every $n \in \mathbb{N}$.

ex:three4 **Example 3.5.** *1. A classic example is the formula*

$$1 + 2 + \cdots + n = \frac{1}{2}n(n+1).$$

Proof. Clearly it is true for $n = 1$. Suppose it holds for a particular n. Then

$$1 + 2 + \cdots + n + (n+1) = \frac{1}{2}n(n+1) + (n+1) = \frac{1}{2}(n+1)(n+2)$$

and so it also holds with n replaced by $n+1$. □

2. A more interesting example is to prove the proposition $P(n)$, that when $n \geq 4$ we have

$$n^2 \leq 2^n.$$

Proof. Let $\mathcal{S}$ be the set of n for which $P(n)$ is true. Since $P(n)$ is not making any claim for $n = 1$, $n = 2$, $n = 3$, it is trivial that $P(1)$, $P(2)$, $P(3)$ are true. We also have

$$4^2 = 16 = 4^2$$

so that $P(4)$ is true. Now suppose that $n \geq 4$ and $P(n)$ is true. Then

$$2^{n+1} = 2 \times 2^n \geq 2n^2 = n^2 + n^2 \geq n^2 + 4n \geq n^2 + 2n + 1 = (n+1)^2.$$

Hence $P(n+1)$ is true.

A different way to organise this would be to think of $P(4)$ as being the first case. Thus we prove that $P(n+3)$ holds for every $n \in \mathbb{N}$. □

3. Suppose that $0 < x < 1$*. Then for every* $n \in \mathbb{N}$ *we have* $0 < x^{n+1} < x^n < 1$

Proof. Suppose that $P(n)$ is the proposition "$0 < x^{n+1} < x^n < 1$". Then $x < 1$ is immediate from the hypothesis, $x^2 < x$ follows from order axiom O4, and we know $0 < x^2$, so $P(1)$ holds.

Now suppose that $P(n)$ is true. Then $0 < x$, so $0 < x^{n+1}.x = x^{n+2}$ and $x^{n+2} = x^{n+1}.x < x^n.x = x^{n+1}$. Moreover $x < 1$ so that $x^{n+1} = x^n.x < x^n < 1$. Hence $P(n+1)$ is true. □

3.2.1 Exercises

subsec:three2

1. Prove that if $n \geq 4$, then $2^n < n!$.

2. Prove that if $n \geq 10$, then $n^3 < 2^n$.

3. Prove that for all $n \in \mathbb{N}$ we have $n(n+2) \leq 2^{n+1}$.

4. Prove that for all $n \geq 7$ we have $3^n \leq n!$.

5. Prove that if x is a real number with $x \geq -1$, then for every $n \in \mathbb{N}$ we have

$$(1+x)^n \geq 1 + nx.$$

This is the **binomial inequality**, which is very useful and which we will use often.

6. Prove that for all $n \in \mathbb{N}$ we have $n(n+2) \leq 2^{n+1}$.

7. Prove that for all $n \geq 7$ we have $3^n \leq n!$.

8. Prove that for all $n \in \mathbb{N}$ we have

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

9. Prove that for all $n \in \mathbb{N}$ we have

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}.$$

10. Prove that for all $n \in \mathbb{N}$ we have $n^2 < 3^n$.

11. Prove that for all $n \in \mathbb{N}$ we have

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} \geq \sqrt{n}.$$

12. Prove that for all $n \in \mathbb{N}$ we have

$$n^2 + n + 1. \leq 3^n$$

3.3 Notes

sec:three3

Many of the most famous unsolved questions in mathematics are connected with the natural numbers. For a brief introduction see the Wikipedia article on number theory https://en.wikipedia.org/wiki/Number_theory

Chapter 4

Sequences

ch:four

4.1 Introduction

sec:four1

def:four1 **Definition 4.1.** *A sequence is a list of real numbers indexed by the members of* $\mathbb{N}$

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

and a_n *denotes the* $n-th$ *term.*

Hopefully in any particular case we might have a formula for a_n, but this is not always so easy to establish.

ex:four1 **Example 4.1.** *Examples of sequences are*

1. $-1, -4, -9, -16, \ldots, -n^2, \ldots,$
2. $1, 1, 1, \ldots, 1, \ldots,$
3. $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n+1}, \ldots,$
4. $2, 3, 5, \ldots, p_n, \ldots$ *where* p_n *denotes the* n*-th prime in order of magnitude.*

Note that repetitions are allowed so the list above is not simply a set of numbers.

The notation $\{a_n\}$ is often used to denote a sequence, but since it can be confused with the notation for a set here we will use the notation

$$\langle a_n \rangle.$$

The set $\mathcal{A} = \{a_n : n \in \mathbb{N}\}$ denotes the **range** of $\langle a_n \rangle$. In all but one of the examples above we do have $\mathcal{A} = \langle a_n \rangle$. The exception is the second one where we have $\mathcal{A} = \{1\}$.

We have some obvious terminology. A sequence $\langle a_n \rangle$ is *bounded above* (or *below*) when $\mathcal{A}$ is bounded above (or below). If $\mathcal{A}$ is both bounded above and below, then $\langle a_n \rangle$ is *bounded*. If it is not bounded, then we say that $\langle a_n \rangle$ is *unbounded*.

Recalling Definitions 2.6, 2.7, 2.8 we have at once the following theorem

thm:four1 **Theorem 4.1.** *A sequence* $\langle a_n \rangle$ *is bounded if and only if there is a real number* H *such that for every* $n \in \mathbb{N}$ *we have* $|a_n| \leq H$.

ex:four2 **Example 4.2.** *1. $\langle 1^n \rangle$ is bounded.*

2. $\langle n^2 \rangle$ is unbounded.

3. $\left\langle \frac{1}{n^2} \right\rangle$ is bounded, by 1 from above and by 0 from below.

4. Here is a more complicated sequence. We define x_n inductively by

$$x_1 = 2, x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

There is no really simple formula for x_n, although something could be worked out. However

$$\begin{aligned} x_2 &= \frac{1}{2}\left(2 + \frac{2}{2}\right) = \frac{3}{2} = 1.5, \\ x_3 &= \frac{1}{2}\left(\frac{3}{2} + \frac{2}{3/2}\right) = \frac{17}{12} = 1.416\ldots, \\ x_4 &= \frac{1}{2}\left(\frac{17}{12} + \frac{24}{17}\right) = \frac{577}{408} = 1.4142\ldots. \end{aligned}$$

Guess what is happening!

This leads on naturally to the next topic

4.1.1 Exercises

subsec:four1

1. Prove that

$$\left\langle \frac{2n^2+n+1}{n^2+3} = 2 \right\rangle$$

is bounded.

2. Prove that the sequence $\langle n^{1/3} \rangle$ is unbounded.

3. Suppose that the sequence $\langle a_n \rangle$ is bounded and for each $n \in \mathbb{N}$ define

$$b_n = \frac{a_1 - 2a_2 + 3a_3 + \cdots + (-1)^{n-1} n a_n}{n^2}.$$

Prove that $\langle b_n \rangle$ is bounded.

4.2 Convergent Sequences

sec:four2

A sequence $\langle a_n \rangle$ converges to the limit ℓ (where $\ell \in \mathbb{R}$) when the following holds.

def:four2 **Definition 4.2.** *Given any real number $\varepsilon > 0$ there is a real number N such that whenever $n \in \mathbb{N}$ and $n > N$ we have*

$$|a_n - \ell| < \varepsilon.$$

When this is satisfied we write

$$\lim_{n\to\infty} a_n = \ell$$

or

$$a_n \to \ell \ \text{as}n \to \infty,$$

and say that a_n tends to ℓ as n tends to infinity. Note that in general we would expect that N is a function of ε. Occasionally we can only prove its existence, but those proofs are usually pretty tricky.

This is the most important definition of the whole course. All other forms of convergence are modelled on this. There is one fundamental difficulty with this definition. What if one does not know the value of ℓ? Often in order to make progress one will need to have a good guess for ℓ. Later we will see ways which avoid this.

ex:four3 **Example 4.3.** *1. Let $a_n = 1/n$. We would guess that the limit exists and is* 0.

2. Suppose that b_n is a constant sequence, i.e there is a real number c such that for every $n \in \mathbb{N}$ we have $b_n = c$. Then $\lim_{n\to\infty} b_n = c$.

Proof. 1. Given any $\varepsilon > 0$ we need to find an N such that whenever $n > N$ we have $|a_n - 0| < \varepsilon$, i.e.

$$\frac{1}{n} = \left|\frac{1}{n}\right| = \left|\left(\frac{1}{n}\right) - 0\right| < \varepsilon.$$

Here we can choose $N = 1/\varepsilon$. Thus whenever $n > N$ we have

$$|a_n - \ell| = \frac{1}{n} < \frac{1}{N} = \varepsilon$$

2. is even easier. Let $\varepsilon > 0$ and choose $N = 1$, say. Then, whenever $n > N$ we have

$$|b_n - c| = |c - c| = 0 < \varepsilon.$$

and we are done. □

Note that to write down the formal proof we need to do some "rough work" to help us find a suitable N, but once we have a handle on N most of the rough work is redundant. It is necessary to get used to doing some "working out" before writing the formal proof, and that is part of the normal process of constructing formal proofs.

ex:four4 **Example 4.4.** *Let $b_n = 1/\sqrt{n}$. Prove that*

$$\lim_{n\to} b_n = 0.$$

Proof. Let $\ell = 0$ and $\varepsilon > 0$. Choose $N = \varepsilon^{-2}$. Thus whenever $n > N$ we have

$$|b_n - \ell| = \left|\frac{1}{\sqrt{n}}\right| = \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N}} = \varepsilon$$

and we are done. □

The following has its uses.

ex:four4a **Example 4.5.** *Suppose that $\langle a_n \rangle$ converges to ℓ. Let $b_n = a_{n+1}$. Then $\langle b_n \rangle$ converges to ℓ.*

Proof. This is immediate from the definition, since if $|a_n - \ell| < \varepsilon$ whenever $n > N$, then for such n we have $n + 1 > n > N$ and so $|b_n - \ell| = |a_{n+1} - \ell| < \varepsilon$. □

It might seem obvious that limits are unique, but it does need to be proved.

thm:four1a **Theorem 4.2.** *A sequence can have at most one limit.*

Proof. We argue by contradiction. Suppose that the sequence $\langle a_n \rangle$ has two different limits, k and ℓ. It is intuitive that when n is large a_n is close to the value of its limit, so it cannot be close to different limits. We can turn this into a proof. Let $\varepsilon = \frac{1}{2}|k - \ell|$. Choose N_1 so that $|a_n - k| < \varepsilon$ when $n > N_1$ and N_2 so that $|a_n - \ell| < \varepsilon$ when $n > N_2$. Suppose that $n > \max\{N_1, N_2\}$. Then, by the triangle inequality

$$|k - \ell| = |a_n - \ell - (a_n - k)| \leq |a_n - \ell| + |a_n - k| < 2\varepsilon = |k - \ell|$$

which is impossible. □

If a sequence is not convergent, then it is **divergent**. Proving that a sequence is divergent can be a little awkward. The following theorem tells us that unbounded sequences are divergent.

thm:four2 **Theorem 4.3.** *Every convergent sequence is bounded.*

Proof. Let $\langle a_n \rangle$ be the sequence in question and let ℓ be its limit. We can just use a special case of the definition of convergence. Let $\varepsilon = 1$ and choose N so that whenever $n > N$ we have

$$|a_n - \ell| < 1.$$

Then, by the triangle inequality, whenever $n > N$

$$|a_n| = |(a_n - \ell) + \ell| \leq |a_n - \ell| + |\ell| < 1 + |\ell|.$$

Now let

$$H = \max\left(\{1 + |\ell|\} \cup \{|a_n| : n \leq N\}\right).$$

Then, for every $n \in \mathbb{N}$, either $n > N$ or $n \leq N$ and so

$$|a_n| \leq H.$$

□

ex:four5 **Example 4.6.** *The sequence $\langle \sqrt{n} \rangle$ is divergent.*

The above is not the only way a sequence might be divergent. This is illustrated by the following example.

ex:four6 **Example 4.7.** *The sequence $\langle (-1)^n \rangle$ is divergent.*

Proof. The idea of the proof is very simple. If the sequence were to be convergent, then successive terms will have to get closer together as n grows. But here they are spaced a distance 2 apart. We argue by contradiction, of course, and make use of the triangle inequality once more.

So, suppose the sequence converges to ℓ, let $\varepsilon = 1$ (any number ≤ 1 would do) and choose N accordingly. Then whenever $n > N$ we have

$$\begin{aligned} 2 &= |(-1)^n + (-1)^n| = |(-1)^n - (-1)^{n+1}| \\ &= |(-1)^n - \ell - \big((-1)^{n+1} - \ell\big)| \\ &\leq |(-1)^n - \ell| + |(-1)^{n+1} - \ell| \\ &< 1 + 1 = 2 \end{aligned}$$

which is impossible. □

Note that it diverges even though it is bounded. In other words being bounded is not enough to confer convergence on a sequence.

How about more complicated sequences such as

$$\left\langle \left(1 + \frac{1}{n}\right)^n \right\rangle ?$$

Proving anything using the definition of convergence might be annoying.

There are a number of theorems which enable us to establish the convergence of more complicated sequences, at least if we understand simpler one.

thm:four3 **Theorem 4.4** (The Combination Theorem for sequences)**.** *Suppose that $\langle a_n \rangle$ converges to α and $\langle b_n \rangle$ converges to β as $n \to \infty$, and let λ and μ be real numbers. Then*

(i) $\langle \lambda a_n + \mu b_n \rangle$ converges to $\lambda\alpha + \mu\beta$ as $n \to \infty$,

(ii) $\langle a_n b_n \rangle$ converges to $\alpha\beta$ as $n \to \infty$.

(iii) If $\beta \neq 0$, then

$$\frac{a_n}{b_n} \to \frac{\alpha}{\beta} \text{ as } n \to \infty.$$

We will see many variants of this as the subject progresses. In part (iii) there is an underlying convention. Since $\beta \neq 0$ we are assured that there is some N_0 such that for $n > N_0$ we have $b_n \neq 0$. It is possible that for some of the $n \leq N_0$ we have $b_n = 0$. In that case the convention is that we suppose that $n > N_0$ and ignore the $n \leq N_0$.

Proof. (i) Let $\varepsilon > 0$. Choose N_1 so that

$$|a_n - \alpha| < \frac{\varepsilon}{2(1+|\lambda|)} \text{ whenever } n > N_1$$

and choose N_2 so that

$$|b_n - \beta| < \frac{\varepsilon}{2(1+|\mu|)} \text{ whenever } n > N_2.$$

Let

$$N = \max\{N_1, N_2\}$$

and suppose that $n > N$, so that both the above inequalities hold. Then, by the triangle inequality and properties of the absolute value,

$$\begin{aligned}
|\lambda a_n + \mu b_n - \lambda\alpha - \mu\beta| &= |\lambda(a_n - \alpha) + \mu(b_n - \beta)| \\
&\leq |\lambda||a_n - \alpha| + |\mu||b_n - \beta| \\
&\leq |\lambda|\frac{\varepsilon}{2(1+|\lambda|)} + |\mu|\frac{\varepsilon}{2(1+|\mu|)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

(ii) Here we need to relate $a_n b_n - \alpha\beta$ to $a_n - \alpha$ and $b_n - \beta$. We can do this *via*

$$a_n b_n - \alpha\beta = (a_n - \alpha)b_n + \alpha(b_n - \beta). \tag{4.1}$$

eq:four3

Since $\langle b_n \rangle$ is convergent we know from Theorem 4.3 that there is an H such that $|b_n| \leq H$. Now we can proceed somewhat as in (i). Let $\varepsilon > 0$, choose N_1 so that whenever $n > N_1$ we have

$$|a_n - \alpha| < \frac{\varepsilon}{2(1+H)},$$

choose N_2 so that whenever $n > N_2$ we have

$$|b_n - \beta| < \frac{\varepsilon}{2(1+|\alpha|)}$$

and suppose that $n > N = \max\{N_1, N_2\}$. Then, by (4.1), and the triangle inequality

$$\begin{aligned}
|a_n b_n - \alpha\beta| &\leq |a_n - \alpha||b_n| + |\alpha||b_n - \beta| \\
&\leq \frac{\varepsilon}{2(1+H)}H + |\alpha|\frac{\varepsilon}{2(1+|\alpha|} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

(iii) Here it suffices to prove that

$$\frac{1}{b_n} \to \frac{1}{\beta} \text{ as } n \to \infty$$

since then we could combine it with (ii). Somehow we need to make use of $\beta \neq 0$. From the special case $\varepsilon = \frac{1}{2}|\beta|$ we know that there is an N_1 such that whenever $n > N_1$ we have $|b_n - \beta| < \frac{1}{2}|\beta|$ so that by the triangle inequality we have

$$|b_n| > |\beta|/2.$$

Now choose an arbitrary $\varepsilon > 0$ and N_2 so that whenever $n > N_2$ we have

$$|b_n - \beta| < \frac{\varepsilon|\beta|^2}{2}.$$

Let $N = \max\{N_1, N_2\}$. Then whenever $n > N$ we have

$$\begin{aligned}\left|\frac{1}{b_n} - \frac{1}{\beta}\right| &= \left|\frac{\beta - b_n}{b_n\beta}\right| \\ &= \frac{|\beta - b_n|}{|b_n||\beta|} \\ &< \frac{2|\beta - b_n|}{|\beta|^2} \\ &< \varepsilon.\end{aligned}$$

□

ex:four7 **Example 4.8.** *Prove that*

$$\lim_{n\to\infty} \frac{n^4 - 3n^2 + 5}{4n^4 + 5n^3 - 3n} = \frac{1}{4}.$$

Proof. We have

$$\frac{n^4 - 3n^2 + 5}{4n^4 + 5n^3 - 3n} = \frac{1 - 3n^{-2} + 5n^{-4}}{4 + 5n^{-1} - 3n^{-3}}$$

and we know from Example 4.3 1. that $n^{-1} \to 0$ as $n \to \infty$ and that $\lim_{n\to\infty} c = c$. Hence we can apply Theorem 4.4 multiple times and obtain successively

$$\begin{aligned}&n^{-2} \to 0, \quad n^{-3} \to 0, \quad n^{-4} \to 0, \\ &1 - 3n^{-2} + 5n^{-4} \to 1, \\ &4 + 5n^{-1} - 3n^{-3} \to 4, \\ &\frac{1 - 3n^{-2} + 5n^{-4}}{4 + 5n^{-1} - 3n^{-3}} \to \frac{1}{4}.\end{aligned}$$

□

What if we do not have an exact formula for the general term of the sequence? Not to despair. The next theorem is very useful in such circumstances.

thm:four4 **Theorem 4.5** (The Sandwich Theorem)**.** *Suppose that $\langle a_n\rangle$, $\langle b_n\rangle$, $\langle c_n\rangle$ are three real sequences with $a_n \le b_n \le c_n$ for every $n \in \mathbb{N}$, and $a_n \to \ell$ as $n \to \infty$ and $c_n \to \ell$ as $n \to \infty$. Then $b_n \to \ell$ as $n \to \infty$*

Proof. Let $\varepsilon > 0$. Choose N_1 so that whenever $n > N_1$ we have $|a_n - \ell| < \varepsilon$ and choose N_2 so that whenever $n > N_2$ we have $|c_n - \ell| < \varepsilon$. Let $N = \max\{N_1, N_2\}$. Then, whenever $n > N$, we have

$$-\varepsilon < a_n - \ell \le b_n - \ell \le c_n - \ell < \varepsilon,$$
$$-\varepsilon < b_n - \ell < \varepsilon,$$
$$|b_n - \ell| < \varepsilon.$$

□

ex:four8 **Example 4.9.** *Suppose that $|x| < 1$. Then $x^n \to 0$ as $n \to \infty$.*

Proof. If $x = 0$, so that $x^n = 0$, then we already know the result. Thus we may suppose that $x \ne 0$, and thus $|x|^{-1} > 1$. Let $y = |x|^{-1} - 1$ so that $y > 0$ and $|x|^{-1} = 1 + y$. By the binomial inequality

$$|x|^{-n} = (1 + y)^n \ge 1 + ny > ny.$$

Hence

$$0 \le |x|^n < \frac{1}{ny}.$$

Now both sides have limit 0 so we can apply the sandwich theorem. □

Another nice example.

ex:four9 **Example 4.10.** *Suppose that $x > 0$. Then $x^{1/n} \to 1$ as $n \to \infty$.*

By the way, by $x^{1/n}$ we mean that positive number t such that $t^n = x$. We have not established that such an object exists, although it not too hard to do so using the completeness axiom. In a while we will look at the special case $2^{1/2}$. However after we have studied monotonic sequences in the next section proofs of such things become easier.

Proof. We first suppose that $x \ge 1$. Then $x^{1/n} \ge 1$, for if on the contrary we had $x^{1/n} < 1$ it would follow by repeated use of the order axioms that $x = (x^{1/n})^n < 1^n = 1$.

Let

$$y_n = x^{1/n} - 1.$$

Then $y_n \ge 0$. Also

$$(1 + y_n)^n = (x^{1/n})^n = x.$$

Hence, by the binomial inequality,

$$x = (1 + y_n)^n \ge 1 + ny_n = 1 + n(x^{1/n} - 1)$$

which can be rearranged to give

$$1 \leq x^{1/n} \leq 1 + \frac{x-1}{n}$$

and again the sandwich theorem comes to our aid.

If instead we have $0 < x < 1$, then

$$\frac{1}{x^{1/n}} = \left(\frac{1}{x}\right)^{1/n} \to 1.$$

Hence, by the combination theorem we have the desired conclusion. □

There is another theorem whose consequences we will use frequently, often without further comment.

thm:four5 **Theorem 4.6.** *Suppose that $c \in \mathbb{R}$ and $\langle a_n \rangle$ is a convergent sequence with $a_n \leq c$ for every $n \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} a_n \leq c.$$

cor:four6 **Corollary 4.7.** *Suppose that $\langle a_n \rangle$ and $\langle b_n \rangle$ are convergent sequences and $a_n \leq b_n$ for every $n \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

Proof of Theorem 4.6. . We argue by contradiction. Suppose that

$$\lim_{n\to\infty} a_n > c.$$

Let

$$\ell = \lim_{n\to\infty} a_n,$$

so that $l > c$. Let $\varepsilon = l - c$. Then there is an N and an $n > N$ such that

$$|a_n - l| < \varepsilon.$$

But then

$$a_n = \ell + a_n - \ell \geq \ell - |a_n - \ell| > \ell - \varepsilon = c$$

contradicting the hypothesis. □

Proof of Corollary 4.7. . For each $n \in \mathbb{N}$, let $d_n = a_n - b_n$. Then $d_n \leq 0$, by the combination theorem

$$\lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n = \lim_{n\to\infty} d_n$$

exists, and by the theorem

$$\lim_{n\to\infty} d_n \leq 0.$$

□

4.2.1 Exercises

subsec:four2

1. Prove that if $\lim_{n\to\infty} a_n = \ell$, then $\lim_{n\to\infty} |a_n| = |\ell|$.

2. Suppose that a_n converges to ℓ and let m be a given integer. When $n \geq \max\{1, 1-m\}$ define $b_n = a_{n+m}$ and when $1 \leq n < \max\{1, 1-m\}$ define $b_n = 0$ (when $m \geq 0$ there are no such n). Prove that b_n converges to ℓ.

3. Prove, using only the definition of a limit, that

$$\lim_{n\to\infty} \frac{n}{n^2+1} = 0.$$

4. Prove, using only the definition of a limit, that

$$\lim_{n\to\infty} \frac{2n^2+n+1}{n^2+3} = 2.$$

5. Let c be a fixed positive number and

$$a_n = \frac{1}{1+nc} \text{ where } c > 0.$$

Using only the definition of a limit to prove that $\langle a_n \rangle$ converges.

6. Prove, using the definition of a limit, that

$$\lim_{n\to\infty} \frac{n^2+1}{n^2+3} = 1.$$

7. Prove, using only the definition of a limit, that

$$\lim_{n\to\infty} \frac{\sqrt{n}}{2+3\sqrt{n}} = \frac{1}{3}.$$

8. Prove that if $x > 0$ and $\langle x_n \rangle$ is a sequence with $\lim_{n\to\infty} x_n = x$, then there is a real number N such that whenever $n > N$ we have $x_n > 0$.

9. Suppose that the sequence $\langle a_n \rangle$ converges to ℓ. Let

$$b_n = \frac{a_1 + a_2 + \cdots + a_n}{n},$$

the "average" of a_n. Using only the definition of limit, prove that $\langle b_n \rangle$ converges to ℓ.

10. Let $a_n = 1 + (-1)^n$.
(i) Prove that $\langle a_n \rangle$ diverges.
(ii) Let

$$b_n = \frac{a_1 + \cdots + a_n}{n}.$$

Prove that $\langle b_n \rangle$ converges.

11. Prove, using only definitions and results established in the course, that

$$\text{(i)} \quad \lim_{n\to\infty} \left\{ \frac{4n^5 + 5n^3 + 6n}{2n^5 + 1} \right\} = 2,$$
$$\text{(ii)} \quad \lim_{n\to\infty} \frac{(-1)^n n}{n^2 + 1} = 0,$$

12. Prove that

$$\lim_{n\to\infty} \frac{3n^5 - 4n^3 + 2n + 7}{4n^5 + 5n^4 + 6n^3 + n^2 + 1} = \frac{3}{4}.$$

13. Suppose that $0 < k < 1$ and $\langle x_n \rangle$ satisfies $|x_{n+1}| < k|x_n|$ for $n = 1, 2, 3, \ldots$. Prove that

(i) $|x_n| \le k^{n-1}|x_1|$,
(ii) $\lim_{n\to\infty} x_n = 0$.

14. Suppose that $c \in \mathbb{R}$ and $\langle a_n \rangle$ is a convergent sequence with $a_n \ge c$ for every $n \in \mathbb{N}$. Prove that $\lim\limits_{n\to\infty} a_n \ge c$.

15. Prove that if $x > 0$ and $\langle x_n \rangle$ is a sequence with $\lim_{n\to\infty} x_n = x$, then there is a real number N such that whenever $n > N$ we have $x_n > 0$.

16. (i) Prove that, if $n \in \mathbb{N}$ and $n \ge 4$, then $2^n < n!$, and deduce that $2^n \le 2((n-1)!)$.
(ii) Prove that

$$\lim_{n\to\infty} \frac{2^n}{n!} = 0.$$

17. The sequence $\langle x_n \rangle$ is defined by $x_1 = 1$ and $x_n = \frac{n}{2(n-1)} x_{n-1}$ $(n = 2, 3, 4, \ldots)$. Prove that

(i) for each $n \in \mathbb{N}$, $x_n > 0$.
(ii) for each $n \in \mathbb{N}$, $x_{n+1} \le x_n$.
(iii) $\lim_{n\to\infty} x_n$ exists, and find its value.

18. (i) Prove that, for each $n \in \mathbb{N}$, $n < 2^n$.
(ii) Prove, using only the definition of limit, that

$$\lim_{n\to\infty} \frac{\sqrt{n}}{2^n} = 0.$$

4.3 Divergence to Infinity

sec:four3

Definition 4.3. *A sequence $\langle a_n \rangle$* **diverges** *to $+\infty$ (and we write $x_n \to +\infty$) as $n \to \infty$ when for any $B > 0$ there exists a real number N such that whenever $n > N$ we have $a_n > B$.*

Likewise $\langle a_n \rangle$ **diverges** *to $-\infty$ (and we write $x_n \to -\infty$) as $n \to \infty$ when for any $b < 0$ there exists a real number N such that whenever $n > N$ we have $a_n < b$.*

ex:four10 **Example 4.11.** *1. Let $a_n = \sqrt{n}$ for $n \in \mathbb{N}$. Then $\langle a_n \rangle$ diverges to $+\infty$.*

2. Let $b_n = n + (-1)^n \sqrt{n}$ for $n \in \mathbb{N}$. Then $\langle b_n \rangle$ diverges to $+\infty$.

Proof. 1. Let $B > 0$ and choose $N = B^2$. Then, whenever $n > N$ we have $a_n = \sqrt{n} > \sqrt{N} = B$.

2. Let $B > 0$ and choose $N = (\sqrt{B} + 1)^2$. Then

$$\begin{aligned} b_n &= n + (-1)^n \sqrt{n} \geq n - \sqrt{n} \\ &= \left(\sqrt{n} - \frac{1}{2} \right)^2 - \frac{1}{4} \\ &> \left(\sqrt{N} - \frac{1}{2} \right)^2 - \frac{1}{4} \\ &= \left(\sqrt{B} + \frac{1}{2} \right)^2 - \frac{1}{4} \\ &= B + \sqrt{B} \\ &> B. \end{aligned}$$

□

4.3.1 Exercises

subsec:four3

1. Prove that the sequence $\langle n^{1/3} \rangle$ diverges to $+\infty$.
2. Let v_n be defined inductively by $v_1 = 1$ and $v_{n+1} = v_n + 1$. Prove that $\langle v_n \rangle$ diverges.

4.4 Notes

sec:four6

There is a brief history of the development of the concept of a convergent sequence at `https://en.wikipedia.org/wiki/Limit_of_a_sequence`

Chapter 5

Monotonic Sequences and Subsequences

ch:five

5.1 Monotonic Sequences

sec:five1

def:five1 **Definition 5.1.** *1. We say that* $\langle a_n \rangle$ *is* **increasing** *when* $a_n \le a_{n+1}$ *for every* $n \in \mathbb{N}$*, and it is* **decreasing** *when* $a_n \ge a_{n+1}$ *for every* $n \in \mathbb{N}$*.*

2. When $a_n < a_{n+1}$ *for every* $n \in \mathbb{N}$ *we call it* **strictly increasing***, and on the other hand when* $a_n > a_{n+1}$ *for every* $n \in \mathbb{N}$ *we call it* **strictly decreasing***.*

3. Such sequences are called **monotonic** *in case 1. and* **strictly monotonic** *in case 2.*

There is a modern tendency to use increasing to mean strictly increasing and, by a terrible misuse of language, to use non-decreasing to mean increasing, and a concomitant variant for the other two cases. A student of the English language would expect that the sequence $\langle (-1)^n \rangle$ is non-decreasing.

ex:five1 **Example 5.1.** *1. Note that the only sequences which are both increasing and decreasing are the constant, such as*

$$1, 1, 1, 1, 1, 1, \ldots$$

2. The sequence $\langle \frac{1}{n} \rangle$ *is strictly decreasing.*

3. The sequence $\langle n^2 \rangle$ *is strictly increasing.*

4. The sequence

$$\left\langle \frac{1}{n} + \frac{(-1)^n}{\sqrt{n}} \right\rangle$$

is neither increasing nor decreasing. 5. The sequence

$$1, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \ldots$$

is decresing but not strictly decreasing.

All the above, except 4. are monotonic, 2. and 3. are strictly monotonic

The next theorem explains the power of the concept. It says that you do not have to know much about a sequence to be sure of its convergence.

thm:five1 **Theorem 5.1.** *Every monotonic bounded sequence* $\langle a_n \rangle$ *converges. When it is increasing the limit is given by* $\sup\{a_n : n \in \mathbb{N}\}$ *and when it is decreasing it is given by* $\inf\{a_n : n \in \mathbb{N}\}$*.*

Proof. If $\langle a_n \rangle$ is bounded and decreasing, then $\langle -a_n \rangle$ is bounded and increasing and

$$\inf\{a_n : n \in \mathbb{N}\} = \sup\{-a_n : n \in \mathbb{N}\},$$

so it suffices to just treat the when $\langle a_n \rangle$ is increasing, which we henceforward assume.

Since the sequence is bounded the set

$$\mathcal{A} = \{a_n : n \in \mathbb{N}\}$$

is bounded above. Moreover as $a_1 \in \mathcal{A}$ it is also non-empty. Hence $\sup \mathcal{A}$ exists. Let $A = \sup \mathcal{A}$ and let $\varepsilon > 0$. Then, by the definition of supremum we cannot have $a_n \leq A - \varepsilon$ for every $n \in \mathbb{N}$. Hence there exists an $N \in \mathbb{N}$ so that

$$A - \varepsilon < a_N \leq A.$$

It then follows by the increasing property and induction on n that

$$A - \varepsilon < a_{N+n} \leq A$$

for every $n \in \mathbb{N}$. Hence

$$|a_n - A| < \varepsilon$$

for every $n > N$. $\square$

ex:five2 **Example 5.2.** *Recall Example 4.2 where we defined inductively*

$$x_1 = 2, x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

There are various observations we can make.

1. *It is a simple induction on* n *to show that* $x_n > 0$ *for every* $n \in \mathbb{N}$.
2. *Squaring both sides and multiplying out gives*

$$\begin{aligned} x_{n+1}^2 &= \frac{1}{4}(x_n^2 + 4 + 4x_n^{-2}), \\ x_{n+1}^2 - 2 &= \frac{1}{4}(x_n^2 - 4 + 4x_n^{-2}) \\ &= \frac{1}{4}(x_n - 2/x_n)^2 \geq 0. \end{aligned}$$

Hence $x_n^2 \geq 2$ *for every* $n \in \mathbb{N}$.

3. Again rearranging the original definition gives

$$x_n - x_{n+1} = \frac{x_n}{2} - \frac{1}{x_n} = \frac{x_n^2 - 2}{2x_n} \geq 0$$
$$x_{n+1} \leq x_n$$

for every $n \in \mathbb{N}$, *so* $\langle x_n \rangle$ *is decreasing and bounded below.*

4. By the monotonic convergence theorem

$$\ell = \lim_{n\to\infty} x_n$$

exists.

5. By 1. and 2. we have $x_n^2 \geq 2 > 1$ *and so* $x_n > 1$. *Thus, since* $\ell = \inf\{x_n\}$ *we have* $\ell \geq 1$.

6. Now reverting to the definition of x_n, *the combination theorem and Example 4.5 we have*

$$\begin{aligned}\ell &= \lim_{n\to\infty} x_{n+1}\\ &= \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)\\ &= \frac{1}{2}\left(\ell + \frac{2}{\ell}\right).\end{aligned}$$

Solving for ℓ *we have*

$$\frac{1}{2}\ell = \frac{1}{\ell}, \quad \ell^2 = 2$$

so we just proved there is a positive real number ℓ *whose square is 2, i.e.* $\sqrt{2}$ *exists.*

5.1.1 Exercises

subsec:five1

1. Suppose that $x_1 = 3$ and for each $n \in \mathbb{N}$ we have $x_{n+1} = \frac{2x_n+1}{x_n+1}$. Prove that
   (i) for every $n \in \mathbb{N}$, $x_n > 1$,
   (ii) the sequence $\langle x_n \rangle$ is decreasing and
   (iii) the sequence $\langle x_n \rangle$ converges.
   (iv) Find the limit.

2. Suppose that $x_1 = 7$ and for each $n \in \mathbb{N}$ we have $x_{n+1} = 2\sqrt{x_n}$. Prove that
   (i) for every $n \in \mathbb{N}$, $x_n > 1$,
   (ii) the sequence $\langle x_n \rangle$ is decreasing and
   (iii) the sequence $\langle x_n \rangle$ converges.
   (iv) Find the limit.

3. Let $c > 0$ and let $\langle y_n \rangle$ be defined iterative by

$$y_{n+1} = \frac{1}{2}\left(y_n + \frac{c}{y_n}\right).$$

Prove that $\langle y_n \rangle$ converges and the limit ℓ has the property that $\ell > 0$ and $\ell^2 = c$.

4. The binomial inequality, Exercise 3.2.1.5, $(1+x)^m \geq 1 + mx$, which holds whenever $x > -1$ and $m \in \mathbb{N}$, is very useful for several parts of this question.

Let $x_n = \left(1 + \frac{1}{n}\right)^n$ when $n \in \mathbb{N}$, and $y_n = \left(1 - \frac{1}{n}\right)^{-n}$ when $n \in \mathbb{N}$ and $n > 1$.
Prove that

(i)

$$\left(\frac{n^2 + 2n}{(n+1)^2}\right)^{n+1} \geq \frac{n}{n+1},$$

(ii) $\langle x_n \rangle$ is an increasing sequence,

(iii) if $n > 1$, then

$$\left(\frac{n^2}{n^2 - 1}\right)^{n+1} \geq \frac{n}{n-1},$$

(iv) $\langle y_n \rangle$ is a decreasing sequence,

(v) if $n > 1$, then $x_n / y_n < 1$,

(vi) if $n > 1$, then $2 \leq x_n < y_n \leq 4$,

(vii) $\langle x_n \rangle$ and $\langle y_n \rangle$ converge, and $2 \leq \lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n \leq 4$,

(viii) $y_{n+1} = x_n \left(1 + \frac{1}{n}\right)$,

(ix) $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$.

The common limit is $e = 2.71828...$, the base of the natural logarithms.

5.2 Subsequences

sec:five2

def:five2 **Definition 5.2.** *Suppose that $\langle a_n \rangle$ is a sequence and $\langle m_n \rangle$ is a strictly increasing sequence of natural numbers. That is, $m_n \in \mathbb{N}$ and $m_n < m_{n+1}$ for every $n \in \mathbb{N}$. Then we call the sequence $\langle a_{m_n} \rangle$ a subsequence of $\langle a_n \rangle$.*

ex:five3 **Example 5.3.** *Suppose that*

$$a_n = \frac{1}{\sqrt{n}}$$

and $m_n = n^2$, so that $\langle m_n \rangle = 1, 4, 9, 16, \ldots$. Then

$$a_{m_n} = \frac{1}{\sqrt{n^2}} = \frac{1}{n}.$$

Subsequences are very useful as a "way in" to the behaviour of a sequence, since a nasty subsequence may well have subsequences which are much easier to deal with and then give us a handle on the original sequence.

thm:five2 **Theorem 5.2.** *Suppose that the sequence $\langle a_n \rangle$ converges to ℓ. Then every subsequence of $\langle a_n \rangle$ converges to ℓ.*

Proof. Let $\langle m_n \rangle$ be a strictly increasing sequence of elements of $\mathbb{N}$. Then a simple induction shows that $m_n \geq n$.

Let $\varepsilon > 0$ and choose N so that whenever $n > N$ we have $|a_n - \ell| < \varepsilon$. Since $m_n \geq n$ we also have $m_n > N$ when $n > N$. Therefore for every $n > N$ we have $|a_{m_n} - \ell| < \varepsilon$ and so $\langle a_{m_n} \rangle$ converges to ℓ. □

ex:five4 **Example 5.4.** *We can now give a simple proof that $\langle (-1)^n \rangle$ diverges.*

Proof. Suppose on the contrary that the sequence converges. Then the subsequences $\langle (-1)^{2n} \rangle$ and $\langle (-1)^{2n-1} \rangle$ would both converge. But the first one converges to $+1$ and the second one to -1 and this would contradict Theorem 4.3. □

Now we come to a more complex example.

ex:five5 **Example 5.5.** *Let $a_n = n^{1/n}$, $b_n = (n+1)^{1/n}$, $c_n = \left(\frac{n}{n+1}\right)^{1/n}$.*

By the binomial inequality

$$\left(\frac{n+1}{n+2}\right)^{n+1} = \left(1 - \frac{1}{n+2}\right)^{n+1} \geq 1 - \frac{n+1}{n+2} = \frac{1}{n+2}.$$
$$(n+1)^{n+1} \geq (n+2)^n, \quad (n+1)^{1/n} \geq (n+2)^{1/(n+1)}$$

So $\langle b_n \rangle$ is decreasing, bounded below and convergent.

We also have

$$1 > \tfrac{n}{n+1} > \left(\tfrac{n}{n+1}\right)^n, \quad 1 > \left(\tfrac{n}{n+1}\right)^{1/n} = c_n > \tfrac{n}{n+1}$$

so, by the sandwich theorem,

$$\lim_{n\to\infty} c_n = 1.$$

By the definitions of the sequences we have

$$b_n c_n = (n+1)^{1/n} \left(\frac{n}{n+1}\right)^{1/n} = n^{1/n} = a_n.$$

Thus a_n converges,

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n$$

and as $a_n > 1$ we have

$$\lim_{n\to\infty} a_n \geq 1.$$

Now consider the subsequence $\langle a_{2n} \rangle$. Then

$$a_{2n}^2 = (2n)^{1/n} = 2^{1/n} a_n.$$

By Exercise **4.10** $2^{1/n} \to 1$. *Hence*

$$\ell^2 = \ell, \quad \ell = 1.$$

Thus

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = 1.$$

The next two theorems are extremely useful when a sequence is not necessarily monotonic.

thm:five3 **Theorem 5.3.** *Every sequence has a monotonic subsequence.*

Proof. Let $\langle a_n \rangle$ be the sequence in question. We call an index m extremal when it has the property that $a_k \leq a_m$ whenever $k \geq m$. If a sequence has infinitely many extrema, then the extrema form a sequence $m_1 < m_2 < \dots$ and since $m_{k+1} > m_k$ we have $a_{m_{k+1}} \leq a_{m_k}$. Thus $\langle a_{m_k} \rangle$ is a decreasing sequence.

Now suppose there are at most a finite number of extrema. Let n_0 denote the last extremum, or in the case that there are no extrema let $n_0 = 1$. Let $m_1 = n_0 + 1$. Since this is not an extremum there will be an $m_2 > m_1$ so that $a_{m_2} > a_{m_1}$. Then we can proceed iteratively. Given a_{m_k}, as $m_k \geq m_1$ and so is not an extremum there will be an $m_{k+1} > m_k$ so that $a_{m_{k+1}} > a_{m_k}$. Thus in this case we have constructed an increasing sequence. □

thm:five4 **Theorem 5.4** (The Bolzano-Weierstrass Theorem)**.** *Every bounded sequence has a convergent subsequence.*

Proof. At once by Theorem 5.3 and the monotonic convergence theorem, Theorem 5.1. □

ex:five6 **Example 5.6.** *1. Recall the Example 4.2, which we examined in detail in Example 5.2 where we defined inductively*

$$x_1 = 2, x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

If $1 \leq x_n \leq 2$, then it follows that

$$1 = \frac{1}{2}\left(1 + \frac{2}{2}\right) \leq x_{n+1} \leq \frac{1}{2}\left(2 + \frac{2}{1}\right) = 2.$$

Hence, by induction, x_n is bounded between 1 and 2. Thus the sequence has a convergent subsequence.

2. In the example $\langle (-1)^n \rangle$ we looked at in Examples 4.7 and 5.4, each of the subsequences $\langle (-1)^{2n} \rangle$ and $\langle (-1)^{2n-1} \rangle$ are monotonic and convergent.

5.2.1 Exercises

1. Here is an outline of an alternative proof of the Bolzano-Weierstrass theorem. Suppose that u and v are lower and upper bounds for the sequence $\langle a_n \rangle$. Prove that there two sequences $\langle u_n \rangle$ and $\langle v_n \rangle$ with the following properties.
(i) $\langle u_n \rangle$ is increasing and $\langle v_n \rangle$ is decreasing.
(ii) $v_n - u_n = \frac{v-u}{2^{n-1}}$.
(iii) The interval $[u_n, v_n]$ contains infinitely many members of the sequence $\langle a_n \rangle$.
(iv) There is a subsequence $\langle a_{m_n} \rangle$ with $u_n \le a_{m_n} \le v_n$ for every n.
(v) $\lim_{n\to\infty} u_n$, $\lim_{n\to\infty} v_n$, $\lim_{n\to\infty} a_{m_m}$ all exist and are equal.

5.3 Limit Inferior and Limit Superior

sec:three

We can also study the limiting behaviour of sequences through the following objects. Given a sequence $\langle a_n \rangle$, let

$$\mathcal{A}_n = \{a_m : m \ge n\} \tag{5.1}$$ eq:five11

and when the sequence is bounded above we write

$$t_n = \sup \mathcal{A}_n.$$

We could also adopt the convention that if the sequence is unbounded above we write $t_n = \infty$, but we should be aware that we cannot then treat t_n as a number, and here we will avoid this convention.

When the sequence is bounded below we likewise write

$$s_n = \inf \mathcal{A}_n.$$

Since $\mathcal{A}_{n+1} \subset \mathcal{A}_n$ it follows that when the sequence is bounded above, so that t_n and t_{n+1} exist, we have

$$t_{n+1} \le t_n.$$

In other words we have a decreasing sequence. If the sequence is also bounded below, then each of the sets $\mathcal{A}_n$ is bounded below by the same bound. Hence $\langle t_n \rangle$ is decreasing and bounded below and so convergent.

A similar argument shows that s_n is increasing and bounded above, and so convergent.

Definition 5.3. *When a sequence $\langle a_n \rangle$ is bounded we define*

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} t_n = \lim_{n\to\infty} \sup\{a_m : m \ge n\}$$

and

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \inf\{a_m : m \ge n\}.$$

The important thing is that when a sequence is bounded these limits always exist. Moreover if we were to adopt a general version of the convention mentioned above, then we could say that they exist even when the sequence is unbounded. This can be very useful and avoids having to deal with objects which might not exist.

ex:five7 **Example 5.7.** *Let $\langle a_n \rangle$ be bounded and let $\langle a_{m_n} \rangle$ be a convergent subsequence. Then*

$$\liminf_{n\to\infty} a_n \le \lim_{n\to\infty} a_{m_n} \le \limsup_{n\to\infty} a_n.$$

Proof. We have $m_n \ge n$. Hence $a_{m_n} \in \mathcal{A}_n$ and so

$$s_n \le a_{m_n} \le t_n$$

and the conclusion follows by Corollary 4.7. □

The power of the concept is illustrated by the next theorem.

thm:five5 **Theorem 5.5.** *Suppose that $\langle a_n \rangle$ is bounded. Then it converges if and only if*

$$\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$$

and then it converges to the common value.

Proof. Note that "if and only if" means we have two tasks.

1. Suppose that $\langle a_n \rangle$ converges. Let ℓ be its limit and let $\varepsilon > 0$. Choose N so that whenever $n > N$ we have

$$|a_n - \ell| < \tfrac{\varepsilon}{2}.$$

When $a_m \in \mathcal{A}_n$ we have $m \ge n > N$ so that

$$|a_m - \ell| < \tfrac{\varepsilon}{2}, \quad \ell - \tfrac{\varepsilon}{2} < a_m < \ell + \tfrac{\varepsilon}{2}.$$

Since these bounds hold for every element of $\mathcal{A}_n$, in the notation used in the preamble we have

$$\ell - \varepsilon < \ell - \frac{\varepsilon}{2} \le s_n \le t_n \le \ell + \frac{\varepsilon}{2} < \ell + \varepsilon.$$

Thus for every $n > N$ we have

$$|s_n - \ell| < \varepsilon, \quad |t_n - \ell| < \varepsilon$$

and so

$$\liminf_{n\to\infty} a_n = l = \limsup_{n\to\infty} a_n.$$

2. As in Example 5.7 we have

$$s_n \le a_n \le t_n.$$

Then the conclusion follows from the sandwich theorem, Theorem 4.5. □

ex:five8 **Example 5.8.** *Define $\langle a_n \rangle$ as follows . Let $k \in \mathbb{N}$ and define*

$$m = n - \frac{k(k-1)}{2},\ a_n = \frac{m}{k} \text{ when } \frac{k(k-1)}{2} < n \leq \frac{(k+1)k}{2}.$$

Since

$$\frac{(k+1)k}{2} - \frac{k(k-1)}{2} = k$$

for this range of n we have everything of the form

$$\frac{1}{k}, \frac{2}{k}, \dots, \frac{k}{k}.$$

Hence our sequence is just an ordering of all the rational numbers in $(0, 1]$, with repetitions of course

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

Thus the sequence is bounded between 0 and 1.

One subsequence is that given by

$$m_k = \frac{k(k-1)}{2} + 1,\ a_{m_k} = \frac{1}{k}$$

and this converges to 0. Another is given by

$$m_k = \frac{k(k-1)}{2} + k = \frac{(k+1)k}{2},\ a_{m_k} = \frac{k}{k} = 1$$

and this converges to 1. It follows that

$$\liminf_{n\to\infty} a_n = 0,\ \limsup_{n\to\infty} a_n = 1.$$

5.3.1 Exercises

subsec:five3

1. Let $D(n)$ denote the number of decimal digits in n and let $Z(n)$ denote the number of zero decimal digits in n. Let

$$a_n = \frac{Z(n)}{D(n)}.$$

Prove that

$$\liminf_{n\to\infty} a_n = 0,\ \limsup_{n\to\infty} a_n = 1.$$

5.4 Cauchy Sequences

sec:five4

When we introduced the idea of convergence we mentioned that one of the difficulties with the definition is the need to know the value of the limit. As we have seen this is a major issue and we have used various work rounds. Cauchy introduced an idea which avoids knowing *a priori* anything about the value of the limit.

def:five3 **Definition 5.4.** *A sequence $\langle a_n \rangle$ is a* **Cauchy sequence** *when for every $\varepsilon > 0$ there is an $N > 0$ such that whenever $n > N$ and $m > N$ we have*

$$|a_n - a_m| < \varepsilon.$$

We remark that in order to satisfy the criterion for being a Cauchy sequence it suffices to know that the above holds for $n > m > N$ because that gives the case $m < n$, the case $n < m$ holds by interchanging the values of m and n, and the case $m = n$ is clear.

There is an immediately useful theorem.

thm:five6 **Theorem 5.6.** *A sequence converges if and only if it is a Cauchy sequence.*

Suddenly we do not have to know anything about the limit!

Proof. We have two tasks.

1. Suppose that the sequence $\langle a_n \rangle$ converges. Let ℓ be the limit and let $\varepsilon > 0$. Choose N so that whenever $n > N$ we have

$$|a_n - \ell| < \frac{\varepsilon}{2}.$$

Then for any m, n with $n > N$ and $m > N$ we have, by the triangle inequality,

$$|a_n - a_m| = |a_n - \ell - (a_m - \ell)| \leq |a_m - \ell| + |a_m - \ell| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

2. Suppose that the sequence $\langle a_n \rangle$ is a Cauchy sequence. Choose N_0 so that whenever $n > m > N_0$ we have $|a_n - a_m| < 1$, and then choose $M \in \mathbb{N}$ so that $N_0 < M \leq N_0 + 1$ and M is fixed by N_0. Then for every $n > M$ we have, again by the triangle inequality,

$$|a_n| = |a_n - a_M + a_M| \leq |a_n - a_M| + |a_M| < 1 + |a_M|.$$

Thus $\langle a_n \rangle$ is bounded by

$$\max\{|a_1|, |a_2|, \ldots, |a_M|, 1 + |a_M|\}.$$

Hence, by the Bolzano-Weierstrass theorem, Theorem 5.4, $\langle a_n \rangle$ has a convergent subsequence, $\langle a_{m_n} \rangle$. Let

$$\ell = \lim_{n \to \infty} a_{m_n}.$$

Let $\varepsilon > 0$. Choose N_1 so that whenever $n > N_1$ we have

$$|a_{m_n} - \ell| < \frac{\varepsilon}{2}. \tag{5.2}$$ eq:five12

We are assuming that the sequence is a Cauchy sequence. Hence we can choose N_2 so that whenever $n > N_2$ and $m > N_2$ we have

$$|a_m - a_n| < \frac{\varepsilon}{2}. \tag{5.3}$$ eq:five13

Now choose $N = \max\{N_1, N_2\}$, so that whenever $n > N$ we have $n > N_2$ and $m_n > N_1$. Then $m_n \geq n > N_2$ also. Hence, by the triangle inequality, (5.2) and (5.3), when $n > N$ we have

$$|a_n - \ell| = |a_n - a_{m_n} + a_{m_n} - \ell| \leq |a_n - a_{m_n}| + |a_{m_n} - \ell| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

ex:five9 **Example 5.9.** *Suppose that $0 < \lambda < 1$ and $\langle a_n \rangle$ is a sequence which satisfies for each $n \geq 1$*

$$|a_{n+1} - a_n| < \lambda^n.$$

We have, for $n > m \geq 1$, by generalizing the triangle inequality to $n - m$ terms (an easy induction),

$$\begin{aligned} |a_n - a_m| &= |(a_n - a_{n-1}) + (a_{n-1} - a_{n-2}) + \cdots + (a_{m+1} - a_m)| \\ &\leq |a_n - a_{n-1}| + |a_{n-1} - a_{n-2}| + \cdots + |a_{m+1} - a_m| \\ &\leq \lambda^{n-1} + \lambda^{n-2} + \cdots + \lambda^m \end{aligned}$$

This last expression is the sum of the first $n - m$ terms of a geometric progression and summing this gives

$$0 < \frac{\lambda^m - \lambda^n}{1 - \lambda} < \frac{\lambda^m}{1 - \lambda}.$$

By Example 4.9 the expression on the right has limit 0 as $m \to \infty$. Hence, for every $\varepsilon > 0$ there is an N so that whenever $n > m > N$ we have

$$|a_n - a_m| < \varepsilon.$$

Thus the sequence is a Cauchy sequence and so it converges. Note that we do not know what the limit is, only that it exists! Indeed given any real number ℓ it is possible to construct a sequence which satisfies the hypothesis and converges to ℓ! For example, take $a_n = \ell + \lambda^n$.

5.4.1 Exercises

subsec:five4

1. Suppose that $\langle a_n \rangle$ is a real sequence and for each $n \in \mathbb{N}$

$$b_n = \frac{a_{n+1} + \cdots + a_{2n}}{n}$$

(i) Prove that if $\langle a_n \rangle$ is a Cauchy sequence, then so is $\langle b_n \rangle$.
(ii) Prove that if $\langle a_n \rangle$ converges, then so does $\langle b_n \rangle$.

5.5 Notes

sec:five

$$5.2 and 5.4. For background to the Bolzano-Weierstrass theorem and Cauchy sequences see `https://en.wikipedia.org/wiki/Bolzano-Weierstrass_theorem` and `https://en.wikipedia.org/wiki/Cauchy_sequence` respectively.

Chapter 6

Series

ch:six

6.1 Series

sec:six1

A series is a sum of the kind

$$a_1 + a_2 + \cdots a_n$$

which is often abbreviated to

$$\sum_{m=1}^{n} a_m.$$

Thus given a sequence $\langle a_n \rangle$ we can form a new sequence $\langle s_n \rangle$ defined by

$$s_n = \sum_{m=1}^{n} a_m. \tag{6.1}$$

eq:six1

def:six0 **Definition 6.1.** *If the sequence $\langle s_n \rangle$ converges, then we say that the* **infinite series**

$$\sum_{m=1}^{\infty} a_m = a_1 + a_2 + \cdots + a_n + \cdots \tag{6.2}$$

eq:six2

converges *and the* **sum** *of the series is the limit*

$$\lim_{n \to \infty} s_n.$$

The s_n are called the **partial sums** *of the infinite series (6.2). When a series converges the sum*

$$t_n = \sum_{m=n+1}^{\infty} a_m \tag{6.3}$$

eq:six2a

is called the **tail** *of the series.*

rem:six1 **Remark 6.1.** *One comment that needs to be said straight away. There is no reason that a series has to start with $n = 1$, so we could equally work with*

$$\sum_{n=M}^{\infty} a_n$$

where M is any integer. Moreover if we can establish the convergence for some M , then it follows for any M by adding or subtracting a finite number of terms.

ex:six1 **Example 6.1.** *Let $x \in \mathbb{R}$ and $a_n = x^n$, so that*

$$s_n = x + x^2 + \ldots + x^n = \frac{x - x^{n+1}}{1 - x}\,(x \neq 1).$$

By Example 4.9, when $|x| < 1$ we have $\lim_{n\to\infty} x^n = 0$. Thus, in that case the series converges and we have

$$\lim_{n\to\infty} s_n = \frac{x}{1-x}\,(|x| < 1).$$

If $x = 1$, then $s_n = n$ and so is unbounded and thus divergent. If $|x| > 1$. Let $y = |x| - 1$. Then by the binomial inequality we have

$$|x|^n = (1 + y)^n \geq 1 + ny$$

and, as $y > 0$, $\langle s_n \rangle$ is unbounded once more and so divergent.

If $x = -1$, then

$$s_n = -1 + 1 - 1 + 1 - \cdots + (-1)^n = \begin{cases} -1 & \text{when } n \text{ is odd,} \\ 0 & \text{when } n \text{ is even.} \end{cases}$$

Since a sequence cannot have two limits the series again diverges, even though it is bounded.

Thus we conclude that

$$\sum_{n=1}^{\infty} x^n$$

converges if and only if $|x| < 1$, and in that case it sums to

$$\frac{x}{1-x}.$$

ex:six2 **Example 6.2.** *Let*

$$a_n = \frac{1}{n(n+1)}.$$

Then

$$s_n = \sum_{m=1}^{n} a_m = \frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + \cdots + \frac{1}{n(n+1)}$$

The nice thing about this series is there is an exact formula for the sum of the first n terms. In fact

$$s_n = 1 - \frac{1}{n+1}.$$

One way to see this is to apply induction. The base case $n = 1$ gives

$$s_1 = \frac{1}{2} = 1 - \frac{1}{1+1}.$$

Now suppose the above formula has been verified for n. Then

$$\begin{aligned} s_{n+1} = s_n + \frac{1}{(n+1)(n+2)} &= 1 - \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} \\ = 1 - \frac{(n+2)-1}{(n+1)(n+2)} &= 1 - \frac{1}{(n+1)+1}. \end{aligned}$$

Now we let $n \to \infty$. Thus $s_n \to 1$. Hence

$$\sum_{m=1}^{\infty} \frac{1}{m(m+1)} = 1.$$

Here is another trick up our sleeve for series.

ex:six3 **Example 6.3.** *Let $b_n = \frac{1}{n^2}$ and*

$$u_n = \sum_{m=1}^{n} b_m.$$

Since each $b_m > 0$, $\langle u_n \rangle$ is an increasing sequence. Moreover, when $m \geq 2$ we have

$$\frac{1}{m^2} \leq \frac{1}{m(m-1)}$$

so

$$u_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} \leq 1 + \frac{1}{1.2} + \frac{1}{2.3} + \cdots \frac{1}{(n-1)n} = 1 + s_{n-1}$$

in the notation of the previous example. Therefore for $n \geq 2$

$$u_n \leq 2 - \frac{1}{n} < 2.$$

Hence we have an increasing sequence which is bounded above. Thus by the monotonic convergence theorem u_n converges.

This is yet another example where we have established convergence but do not yet have the tools to give the value of the limit.

An immediate consequence of the definition is the following.

thm:six0 **Theorem 6.1.** *Suppose that the series (6.1) converges. Then the tail of the series (6.3) satisfies*

$$\lim_{n\to\infty} t_n = 0.$$

Proof. Let ℓ denote the value of the infinite series (6.2). Then

$$t_n = \ell - s_n \to 0 \text{ as } n \to \infty.$$

□

We can now port over the theory of sequences to the theory of series. For example, the following theorem is very useful.

thm:six1 **Theorem 6.2** (The Combination Theorem for Series)**.** *Suppose that*

$$\sum_{n=1}^{\infty} a_n \textit{ and } \sum_{n=1}^{\infty} b_n$$

converge to α and β respectively and λ and μ are real numbers. Let

$$c_n = \lambda a_n + \mu b_n \, (n \in \mathbb{N}).$$

Then

$$\sum_{n=1}^{\infty} c_n$$

converges to $\lambda\alpha + \mu\beta$.

6.1.1 Exercises

subsec:six1

1. Prove that

$$\sum_{m=1}^{n} \frac{1}{m(m+2)} = \frac{3}{4} - \frac{1}{2(n+1)} - \frac{1}{2(n+2)}$$

and that

$$\sum_{m=1}^{\infty} \frac{1}{m(m+2}) = \frac{3}{4}.$$

2. Suppose that $|x| < 1$. Prove that

$$\sum_{m=1}^{n} m x^m = \frac{x - (n+1)x^{n+1} + n x^{n+2}}{(1-x)^2}$$

and that

$$\sum_{m=1}^{\infty} m x^m$$

converges.

3. Suppose that $\langle a_n \rangle$ is a real sequence and

$$\sum_{n=1}^{\infty} a_n$$

converges and for every $n \in \mathbb{N}$ we have $a_n \geq a_{n+1}$. Prove that $\lim_{n\to\infty} n a_n = 0$.

6.2 Tests for Convergence of Series

sec:six2

Because series are so important there are various tests and criteria for their convergence, and these can be presented in the form of an algorithm. Be warned that most of the really interesting series fall outside the scope of this algorithm!

Suppose that $\langle a_n \rangle$ is a real sequence and s_n is defined by (6.1). Then we are concerned with the existence of (6.2).

Step 1. *If* $\lim_{n\to} a_n$ *does not exist, or it does but it is not* 0*, then (6.2) diverges.*

Step 2. *The Comparison Test.* Comparison with a known series. There are two cases.

2.1. Suppose that $|a_n| \leq b_n$ for every $n \in \mathbb{N}$ and

$$\sum_{n=1}^{\infty} b_n$$

converges. Then so does (6.2).

2.2. Suppose that $0 \leq c_n \leq a_n$ for every $n \in \mathbb{N}$ and

$$\sum_{n=1}^{\infty} c_n$$

diverges. Then so does (6.2).

Step 3. *The ratio test.* Suppose that $a_n \neq 0$ for every large n and

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

exists. Let its value be ℓ.

If $\ell < 1$, then (6.2) converges.

If $\ell > 1$, then (6.2) diverges.

If $\ell = 1$, then no conclusion can be made.

Step 4. *The Leibnitz (or alternating series) test.* Suppose there is a sequence $\langle d_n \rangle$ which is (i) non-negative, (ii) decreasing and (iii) satisfies

$$\lim_{n\to\infty} d_n = 0$$

and (iv) $a_n = (-1)^{n-1} d_n$. Then (6.2) converges.

There are some more sophisticated versions of **3.**, such as the n-th root test, but if Step **3.** fails to decide convergence or divergence these more sophisticated versions are unlikely to do any better. Typically if the algorithm fails to determine convergence or divergence, then an *ad hoc* method is usually the way to go.

ex:six4 **Example 6.4.**

$$\sum_{n=1}^{\infty} (-1)^n$$

diverges because $(-1)^n \not\to$ *limit as* $n \to$.

ex:six5 **Example 6.5.**

$$\sum_{n=1}^{\infty} (1 - 1/n)^2$$

diverges because

$$\lim_{n\to\infty} (1 - 1/n)^2 = 1 \neq 0.$$

Example 6.2 gives an example in which **2.1.** holds.

Crucial for the utility of the comparison test is a range of useful examples. We will show later that

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges. Then it follows from **Step 2.2** that if $c < 1$, then

$$\sum_{n=1}^{\infty} \frac{1}{n^c}$$

diverges.

ex:six6 **Example 6.6.** *Let*

$$a_n = \frac{(n!)^2}{(2n)!}.$$

Then

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(2n)!((n+1)!)^2}{(2n+2)!(n!)^2} = \frac{(n+1)^2}{(2n+1)(2n+2)} \to \frac{1}{4}.$$

Hence

$$\sum_{n=1}^{\infty} a_n$$

converges by the ratio test.

Here is a more elaborate version of this.

ex:six7 **Example 6.7.** *Let $x \in \mathbb{R}$ and*

$$b_n = \frac{(n!)^2}{(2n)!} x^n.$$

Then

$$\left|\frac{b_{n+1}}{b_n}\right| = \frac{(2n)!((n+1)!)^2}{(2n+2)!(n!)^2}|x| = \frac{(n+1)^2}{(2n=1)(2n+2)}|x| \to \frac{|x|}{4}.$$

Hence

$$\sum_{n=1}^{\infty} b_n$$

converges when $|x| < 4$ and diverges when $|x| > 4$, by the ratio test. Note that nothing can be concluded when $|x| = \frac{1}{4}$. By more sophisticated arguments the series can be shown to converge when $x = -\frac{1}{4}$ and diverge when $x = \frac{1}{4}$.

ex:six8 **Example 6.8.** *Let $x \in \mathbb{R}$ and*

$$c_n = \frac{x^n}{n!}.$$

Then

$$\left|\frac{c_{n+1}}{c_n}\right| = \frac{n!}{(n+1)!}|x| = \frac{|x|}{n+1} \to 0$$

regardless of the value of x. Hence

$$\sum_{n=1}^{\infty} \frac{x^n}{n!}$$

converges for every real x.

We remark that the function

$$\exp(x) = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

is very important. Note that here we have deployed the conventions $0! = 1$ and that in such series $x^0 = 1$ even when $x = 0$.

ex:six9 **Example 6.9.** *If $a_n = 1$ for every n we have $s_n = n$ and so*

$$\sum_{n=1}^{\infty} a_n$$

diverges. If instead $a_n = \frac{1}{n^2}$, then the series converges. But in either case we have

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1.$$

ex:six10 **Example 6.10.** *Let*

$$a_n = \frac{(-1)^{n-1}}{\sqrt{n}}.$$

We apply the alternating series test with

$$d_n = \frac{1}{\sqrt{n}}.$$

For every $n \in \mathbb{N}$ *we have* $d_n > 0$ *and*

$$d_{n+1} = \frac{1}{\sqrt{n+1}} < \frac{1}{\sqrt{n}} = d_n$$

so d_n *is decreasing and*

$$\lim_{n\to\infty} d_n = 0.$$

Thus

$$\sum_{n=1}^{\infty} a_n$$

converges by the Leibnitz test.

6.2.1 Exercises

subsec:six2

1. Decide the convergence of the each of the following series, in each case proving your assertion by using the appropriate tests for convergence and divergence.

$$\text{(i)} \sum_{n=1}^{\infty} \frac{2}{n^2+n-1} \quad \text{(ii)} \sum_{n=1}^{\infty} \frac{1}{2n+(-1)^n} \quad \text{(iii)} \sum_{n=1}^{\infty} \frac{(n!)^4}{(4n)!}(255)^n$$

$$\text{(iv)} \sum_{n=1}^{\infty} \frac{(n!)^4}{(4n)!}(257)^n \quad \text{(v)} \sum_{n=1}^{\infty}(-1)^{n-1}n^{-2/3} \quad \text{(vi)} \sum_{n=1}^{\infty}(-1)^{n-1}\left(1+\frac{1}{\sqrt{n}}\right).$$

2. Decide the convergence of the each of the following series, in each case proving your assertion.

$$\text{(i)} \sum_{n=1}^{\infty} \frac{3}{n^3+2} \quad \text{(ii)} \sum_{n=1}^{\infty} \frac{4}{3n+2} \quad \text{(iii)} \sum_{n=1}^{\infty} \frac{(n!)^3}{(3n)!}(26)^n$$

$$\text{(iv)} \sum_{n=1}^{\infty} \frac{(n!)^3}{(3n)!}(28)^n \quad \text{(v)} \sum_{n=1}^{\infty}(-1)^{n-1}n^{-1/4} \quad \text{(vi)} \sum_{n=1}^{\infty}\left(1+\frac{1}{n}\right)(-1)^n.$$

3. Prove that

$$\sum_{n=1}^{\infty} x^n \frac{(n!)^2}{(2n)!}$$

converges when $|x| < 4$ and diverges when $|x| > 4$.

4. Prove that

$$\sum_{n=1}^{\infty} \frac{(-x)^n}{(2n-1)!}$$

converges for all real x.

5. State in each case whether the series below converges, and justify your assertions.

$$\text{(i)} \quad \sum_{n=1}^{\infty} \frac{1}{n(n+1)}, \quad \text{(ii)} \quad \sum_{n=1}^{\infty} \frac{2}{n+1}, \quad \text{(iii)} \quad \sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!} 3^n,$$

$$\text{(iv)} \quad \sum_{n=1}^{\infty} (-1)^{n-1} n^{-1/2}, \quad \text{(v)} \quad \sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{n+1}.$$

6. Decide the convergence of the each of the following series, in each case proving your assertion

$$\text{(i)} \ \sum_{n=1}^{\infty} \frac{2}{n^3+1}, \ \text{(ii)} \ \sum_{n=1}^{\infty} \frac{3}{2n+1}, \ \text{(iii)} \ \sum_{n=1}^{\infty} \frac{(n!)^2}{(2n-1)!} 3^n,$$

$$\text{(iv)} \ \sum_{n=1}^{\infty} \frac{(n!)^2}{(2n-1)!} 5^n, \ \text{(v)} \ \sum_{n=1}^{\infty} (-1)^{n-1} n^{-1/3}, \ \text{(vi)} \ \sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right) (-1)^n.$$

6.3 Proofs of the Tests

sec:six3

The first test is easily dealt with.

thm:six2

Theorem 6.3. *If* $\lim_{n\to\infty} a_n$ *does not exist, or it does but is not* 0*, then*

$$\sum_{n=1}^{\infty} a_n$$

diverges.

Proof. Suppose on the contrary that

$$\lim_{n\to\infty} s_n$$

exist and its value is ℓ. Then $a_n = s_n - s_{n-1}$ and so by the combination theorem

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} s_n - s_{n-1} = \ell - \ell = 0$$

contradicting the hypothesis. □

The remaining tests are more demanding.

thm:six3 **Theorem 6.4.** *1. Suppose that* $|a_n| \le b_n$ *for every* $n \in \mathbb{N}$ *and*

$$\sum_{n=1}^{\infty} b_n$$

converges. Then so does

$$\sum_{n=1}^{\infty} a_n.$$

2. Suppose that $0 \le c_n \le a_n$ *for every* $n \in \mathbb{N}$ *and*

$$\sum_{n=1}^{\infty} c_n$$

diverges. Then so does

$$\sum_{n=1}^{\infty} a_n.$$

Proof. 1. We first treat a special case. Suppose $0 \le A_n \le b_n$. Let

$$u_n = \sum_{m=1}^{n} A_m$$

and

$$B = \sum_{m=1}^{\infty} b_m.$$

Then

$$u_n \le \sum_{m=1}^{n} b_m \le \sum_{m=1}^{\infty} b_n = B,$$

so $\langle u_n \rangle$ is bounded above and, since the terms A_n are non-negative, the sequence is increasing. Hence $\langle u_n \rangle$ converges.

Now we turn to the general case $|a_n| \le b_n$ for every $n \in \mathbb{N}$. Let

$$D_n = \begin{cases} a_n & \text{when } (a_n \ge 0), \\ 0 & \text{when } (a_n < 0), \end{cases}$$

$$E_n = \begin{cases} 0 & \text{when } (a_n \ge 0), \\ -a_n & \text{when } (a_n < 0). \end{cases}$$

Then $0 \le D_n \le b_n$ and $0 \le E_n \le b_n$. Hence

$$\sum_{n=1}^{\infty} D_n$$

and

$$\sum_{n=1}^{\infty} E_n$$

both converge. Thus by the combination theorem, Theorem 6.2,

$$\sum_{n=1}^{\infty} (D_n - E_n)$$

converges. But $D_n - E_n = a_n$ for every $n \in \mathbb{N}$.

2. We have $0 \leq c_n \leq a_n$ and

$$\sum_{n=1}^{\infty} c_n \tag{6.4}$$ eq:six3

diverges. Let

$$t_n = \sum_{m=1}^{n} c_m.$$

Since each $c_m \geq 0$, $\langle t_n \rangle$ is an increasing sequence. If the sequence $\langle t_n \rangle$ were bounded then the series (6.4) would have to converge by the monotone convergence theorem. Hence it is unbounded. But $s_n \geq t_n$, so $\langle s_n \rangle$ is unbounded and hence (6.2) diverges. □

thm:six4 **Theorem 6.5** (The Ratio Test)**.** *Suppose that $a_n \neq 0$ for every large n and*

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

exists. Let its value be ℓ.

If $\ell < 1$, then (6.2) converges.

If $\ell > 1$, then (6.2) diverges.

Proof. Of course $\ell \geq 0$. Suppose first of all that $\ell < 1$. Then we want to compare (6.2) with a series of the form

$$\sum_{n=1}^{\infty} x^n.$$

Let

$$\varepsilon = \frac{1-\ell}{2}.$$

Choose $N \in \mathbb{N}$ so that whenever $n > N$ we have

$$\left| \left| \frac{a_{n+1}}{a_n} \right| - \ell \right| < \varepsilon.$$

Thus

$$\left| \frac{a_{n+1}}{a_n} \right| - \ell < \varepsilon.$$

Put $x = \ell + \varepsilon$ so that

$$0 < x = \ell + \frac{1-\ell}{2} = \frac{1+\ell}{2} < 1$$

and

$$\left|\frac{a_{n+1}}{a_n}\right| < x$$

whenever $n > N$.

Now by induction on $n \geq N$ we have

$$|a_n| \leq x^n |a_N| x^{-N}$$

To see this take the base case as $n = N$ and then given $n \leq N$ we have

$$|a_{n+1}| < x|a_n| \leq x^{n+1}|a_N|x^{-N}.$$

Now by Example 6.1

$$\sum_{n=1}^{\infty} x^n$$

converges so, by the combination theorem,

$$\sum_{n=1}^{\infty} x^n |a_N| x^{-N}.$$

converges. Hence, by the comparison test

$$\sum_{n=N}^{\infty} a_n$$

converges. Thus, by Remark 6.1 the first part of the theorem follows.

Now suppose that $\ell > 1$. Then, by taking $\varepsilon = l - 1$ in the definition of convergence it follows that there is an $N \in \mathbb{N}$ so that whenever $n \geq N$ we have

$$\left|\frac{a_{n+1}}{a_n}\right| > 1.$$

Hence

$$|a_{n+1}| > |a_n| > \ldots |a_N| > 0.$$

Thus either $\lim_{n\to\infty} a_n$ does not exist or $|\lim_{n\to\infty} a_n| \geq |a_N| > 0$, so the second part of the theorem follows from Theorem 6.3. □

For completeness at this stage, we include the following test. For most applications of this test it is easier to use the ratio test. However it does have the merit of not requiring that $a_n \neq 0$ and there is an important application later in §6.5 to power series. Given any non-negative number c we mean by $c^{1/n}$ the positive real number x such that $x^n = c$. We can establish the existence of such a number by taking

$$x = \sup\{r : r \in \mathbb{Q}, r \geq 0, r^n \leq c\}$$

thm:six4a **Theorem 6.6** (The Root Test). *If the sequence*

$$b_n = |a_n|^{1/n}$$

is bounded and

$$\limsup_{n\to\infty} b_n < 1,$$

then the series

$$\sum_{n=1}^{\infty} a_n$$

converges absolutely. On the other hand if $\langle b_n \rangle$ is unbounded, or it is bounded but

$$\limsup_{n\to\infty} b_n > 1,$$

then the series diverges.

Proof. The second case is easy. In that case there will be infinitely many n so that $|a_n|^{1/n} \geq 1$ and so $|a_n| \geq 1$. Hence we cannot have $\lim_{n\to} a_n = 0$ and can appeal to Theorem 6.3.

The proof in the first case has a similar structure to the proof of the Ratio test when $\ell < 1$. let

$$\ell = \limsup_{n\to\infty} b_n$$

and choose $\varepsilon = \frac{1-\ell}{2}$. Then there is an $N \in \mathbb{N}$ such that whenever $n > N$ we have

$$|\sup\{b_m : m \geq n\} - \ell| < \varepsilon$$

and so

$$\sup\{b_m : m \geq N+1\} < \ell + \varepsilon = x$$

where $x = \frac{1+\ell}{2} < 1$. Now for every $m \geq N+1$ we have

$$|a_m|^{1/m} = b_m \leq x, \quad |a_m| \leq x^m$$

and we can proceed much as in the ratio test. □

We now come to the final part of our algorithm.

thm:six5 **Theorem 6.7** (The Leibnitz Test). *Suppose there is a sequence $\langle d_n \rangle$ which is (i) non-negative, (ii) decreasing and (iii) satisfies*

$$\lim_{n\to\infty} d_n = 0$$

and (iv) $a_n = (-1)^{n-1} d_n$. Then (6.2) converges.

Proof. As usual, let
$$s_n = \sum_{m=1}^{n} a_n.$$
Then
$$s_{2n+2} = s_{2n} + a_{2n+2} + a_{2n+1} = s_{2n} - d_{2n+2} + d_{2n+1} \geq s_{2n}$$
since $d_{2n+2} \leq d_{2n+1}$.

Likewise
$$s_{2n+1} = s_{2n-1} + a_{2n+1} + a_{2n} = s_{2n-1} + d_{2n+1} - d_{2n} \leq s_{2n-1}.$$
Hence the subsequences $\langle s_{2n} \rangle$ and $\langle s_{2n-1} \rangle$ are increasing and decreasing respectively. We also have
$$s_{2n} = s_{2n-1} + a_{2n} = s_{2n-1} - d_{2n} \leq s_{2n-1}$$
so that
$$s_2 \leq s_4 \leq s_6 \leq \ldots \leq s_{2n} \leq s_{2n-1} \leq \cdots \leq s_5 \leq s_3 \leq s_1.$$
Thus $\langle s_{2n} \rangle$ is increasing and bounded above by s_1 and $\langle s_{2n-1} \rangle$ is decreasing and bounded below by s_2. Hence both subsequences converge. Let
$$\ell_1 = \lim_{n\to\infty} s_{2n-1}, \quad \ell_2 = \lim_{n\to\infty} s_{2n}.$$
Then
$$\ell_1 - \ell_2 = \lim_{n\to\infty} (s_{2n-1} - s_{2n}) = \lim_{n\to\infty} d_{2n} = 0.$$
Let $\ell = \ell_1 = \ell_2$. It follows that $\lim_{n\to\infty} s_n = \ell$. □

6.3.1 Exercises

subsec:six3

1. Suppose that $\langle a_n \rangle$ is a real sequence and that for every $\varepsilon > 0$ there is an N such that whenever $n > m > N$ we have
$$\left| \sum_{k=m+1}^{n} a_k \right| < \varepsilon.$$
Prove that
$$\sum_{n=1}^{\infty} a_n$$
converges.

2. (Dirichlet's Test.) Suppose that the real sequence $\langle a_n \rangle$ is monotonice, that
$$\lim_{n\to\infty} a_n = 0$$

and the sequence $\langle b_n \rangle$ has the property that there is real number B such that for every $n \in \mathbb{N}$ we have

$$\left| \sum_{m=1}^{n} b_m \right| \leq B.$$

Prove that

$$\sum_{n=1}^{\infty} a_n b_n$$

converges.

3. (Abel's Test.) Suppose that the series

$$\sum_{n=1}^{\infty} a_n$$

converges and that the sequence $\langle b_n \rangle$ is monotonic and bounded. Prove that

$$\sum_{n=1}^{\infty} a_n b_n$$

converges.

6.4 Further Theorems and Examples

sec:six4

There is a terminology which can now be introduced, following Theorem 6.4.

def:six1 **Definition 6.2.** *A series*

$$\sum_{n=1}^{\infty} a_n \tag{6.5}$$

eq:six5

is **absolutely convergent** *when*

$$\sum_{n=1}^{\infty} |a_n| \tag{6.6}$$

eq:six6

converges. When (6.5) converges but (6.6) diverges we call the series (6.5) **conditionally convergent.**

Note that a convergent series is not necessarily absolutely convergent. For example

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{\sqrt{n}}$$

converges by the Leibnitz test, Theorem 6.7, but

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$$

diverges since the n-th partial sum is bounded below by $\sqrt{n}$ and so is unbounded.

The following is an immediate corollary of Theorem 6.4. Indeed any series which passes part 1. of that theorem is automatically absolutely convergent.

thm:six8 **Theorem 6.8.** *Every absolutely convergent series is convergent.*

Proof. Take $b_n = |a_n|$ in part 1. of the comparison test. □

The last theorem confers a very important and useful further property.

thm:six8a **Theorem 6.9.** *Let $f : \mathbb{N} \to \mathbb{N}$ be a permutation of $\mathbb{N}$. That is, f is a bijection - for every $n \in \mathbb{N}$ there is a unique $m \in \mathbb{N}$ such that $f(m) = n$. Suppose, moreover, that*

$$\sum_{n=1}^{\infty} a_n$$

converges absolutely. Then so does

$$\sum_{n=1}^{\infty} a_{f(n)}$$

and

$$\sum_{n=1}^{\infty} a_{f(n)} = \sum_{n=1}^{\infty} a_n.$$

In other words, however one rearranges an absolutely convergent series the sum remains the same. Riemann showed that this is false for conditionally convergent series.

Proof. The Cauchy condition for convergence tells us that given any $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that whenever $n > m > N$ we have

$$\sum_{k=m+1}^{n} |a_k| < \varepsilon.$$

Let

$$M = \max\{m : f(m) \leq N\}$$

Thus, when $m > M$ we have $f(m) > N$. Hence, whenever $M < m < n$ we have

$$\sum_{k=m+1}^{n} |a_{f(k)}| \leq \sum_{k=M+1}^{L(n)} |a_k| < \varepsilon$$

where

$$L(n) = \max\{f(k) : M + 1 \leq k \leq n\}.$$

Thus, by the Cauchy condition for convergence,

$$\sum_{n=1}^{\infty} a_{f(n)}$$

converges absolutely.

Let $n_0(n)$ denote the smallest k such that $k \notin \{f(1), f(2), \dots f(n)\}$ and let $K(m,n) = \min\{n_0(n), m+1\}$. Then K is the smallest member of $\mathbb{N}$ which is in neither of the first two sums below, and so all the terms with smaller index k cancel. Hence

$$\left|\sum_{k=1}^{m} a_k - \sum_{k=1}^{n} a_{f(k)}\right| \le \sum_{k=K(m,n)}^{\infty} |a_k|.$$

Now let $n \to \infty$, and then $m \to \infty$. Then $K(m,n) \to \infty$ and

$$\left|\sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_{f(k)}\right| = 0.$$

□

For most series that one comes across, the ratio test and its allies are useless, because in principle the ratio test is making a comparison with a geometric series, and geometric series converge or diverges exponentially fast. Most series converge or diverge much more slowly. In such cases the normal process would be to compare with one of the series considered here.

thm:six6 **Theorem 6.10.** *Suppose that $\sigma \in \mathbb{R}$ and $\sigma \le 1$. Then the series*

$$\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$$

diverges.

Proof. We argue by contradiction. Suppose that the series converges, let

$$\ell = \sum_{n=1}^{\infty} \frac{1}{n^\sigma}$$

and let

$$s_n = \sum_{m=1}^{n} \frac{1}{m^\sigma}.$$

Then $\langle s_n \rangle$ converges to ℓ and hence so does the subsequence

$$\langle s_{2n} \rangle.$$

Therefore

$$\lim_{n\to\infty} (s_{2n} - s_n) = \ell - \ell = 0.$$

But

$$s_{2n} - s_n = \sum_{m=n+1}^{2n} \frac{1}{m^\sigma} \ge \sum_{m=n+1}^{2n} \frac{1}{(2n)^\sigma} = 2^{-\sigma} n^{1-\sigma} \ge \frac{1}{2}.$$

Taking limits we just showed that $0 \ge \frac{1}{2}$. □

One can contrast the previous theorem with the next one.

thm:six7 **Theorem 6.11.** *Suppose that* $\sigma \in \mathbb{R}$ *and* $\sigma > 1$. *Then*

$$\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$$

converges.

One can see that there is some delicacy in these conclusions, and that there is a switch over at $\sigma = 1$.

Proof. We have $n^\sigma > 0$ for every $n \in \mathbb{N}$. Thus the partial sums

$$s_n = \sum_{m=1}^{n} \frac{1}{m^\sigma}$$

form an increasing sequence. Hence it suffices to show that the subsequence $\langle s_{2^k} \rangle$ is bounded above, i.e. $s_{2_k} \le B$ for every $k \in \mathbb{N}$, because given n the Archimedean property ensures that there is a k with $n \le 2^k$ and then it follows that $s_n \le s_{2^k} \le B$.

Let

$$t_k = s_{2^k} - s_{2^{k-1}} = \sum_{n=2^{k-1}+1}^{2^k} \frac{1}{n^\sigma}.$$

Then

$$1 + t_1 + t_2 + \cdots t_k = 1 + (s_2 - s_1) + (s_4 - s_2) + \cdots (s_{2^k} - s_{2^{k-1}}) = s_{2^k} + 1 - s_1 = s_{2^k}. \quad (6.7)$$ eq:six4

Moreover

$$t_j = \sum_{n=2^{j-1}+1}^{2^j} \frac{1}{n^\sigma} \le \frac{2^{j-1}}{2^{(j-1)(\sigma)}} = x^{j-1}$$

where

$$x = 2^{1-\sigma}$$

so that $0 < x < 1$. Hence, by Example 6.1 and the comparison test,

$$\sum_{j=1}^{k} t_j$$

converges and so by (6.7) $\langle s_{2^k} \rangle$ converges and so is bounded, as required. □

6.4.1 Exercises

subsec:six4

1. Prove that if $\sigma > 1$ then

$$\sum_{n=1}^{\infty} \frac{\log n}{n^\sigma}$$

converges.

6.5 Power Series

sec:six5

We now examine a special class of series which give rise to many of the most important functions in mathematics and have myriad applications.

def:six2 **Definition 6.3.** *For a given sequence $\langle a_n \rangle$ of real numbers consider the series*

$$A(x) = \sum_{n=0}^{\infty} a_n x^n. \tag{6.8}$$

eq:six7

We call such a series a **power series**. *Note that we include a term with $n = 0$ and by convention $x^0 = 1$ regardless of the value of x.*

The following is the fundamental theorem of power series.

thm:six9 **Theorem 6.12.** *Given a sequence $\langle a_n \rangle$ of real numbers and the corresponding power series $A(x)$,*

(i) the series converges absolutely for every x and

$$\limsup_{n\to\infty} |a_n|^{1/n} = 0$$

or

(ii) there is a positive real number R such that the series converges absolutely for all x with $|x| < R$ and diverges for all x with $|x| > R$ and

$$\limsup_{n\to\infty} |a_n|^{1/n} = R^{-1}$$

or

(iii) the series converges for $x = 0$ only and

$$\langle |a_n|^{1/n} \rangle$$

is unbounded.

def:six3 **Definition 6.4.** *It is conventional to define R in case (ii) to be the* **radius of convergence** *of $A(x)$, and to extend this to be $R = \infty$ in case (i) and $R = 0$ in case (iii). Moreover, by an abuse of notation we could write in each case*

$$R = 1/\limsup_{n\to\infty} |a_n|^{1/n}.$$

Proof of Theorem 6.12. We can certainly suppose throughout that $x \neq 0$. Let

$$c_n = a_n x^n.$$

Then

$$|c_n|^{1/n} = |x||a_n|^{1/n}.$$

If $\langle |c_n|^{1/n} \rangle$ is unbounded, then so is $\langle |a_n|^{1/n} \rangle$. Hence by the root test, Theorem 6.6, the series diverges for all non-zero x, which gives case (iii).

If
$$\limsup_{n\to\infty} |c_n|^{1/n},$$
exists and is non-zero, then likewise for
$$\limsup_{n\to\infty} |a_n|^{1/n}$$
and we can define
$$R = \left(\limsup_{n\to\infty} |a_n|^{1/n}.\right)^{-1}$$
Then
$$\limsup_{n\to\infty} |c_n|^{1/n} = |x| R^{-1}$$
and by the root test the series converges absolutely when $|x| < R$ and diverges when $|x| > R$. which gives (ii).

Finally, when
$$\limsup_{n\to\infty} |c_n|^{1/n} = 0,$$
we have
$$|x| \limsup_{n\to\infty} |a_n|^{1/n} = 0$$
and so
$$\limsup_{n\to\infty} |a_n|^{1/n} = 0.$$
Thus by the root test the series converges absolutely for every value of x. This gives case (i) and completes the proof of the theorem. □

We can now introduce some important functions.

def:six4 **Definition 6.5.** *Whenever the corresponding series converges we define*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{6.9}$$ eq:six8

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \tag{6.10}$$ eq:six9

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \tag{6.11}$$ eq:six10

The first part of the following theorem is an easy consequence of the ratio test.

thm:six13 **Theorem 6.13.** *(i) Each of the series (6.9), (6.10), (6.11) has radius of convergence* ∞.

(ii) We have $\exp(0) = 1$, $\sin(0) = 0$, $\cos(0) = 1$.

(iii) For every pair of real numbers x and y we have

$$\exp(x + y) = \exp(x)\exp(y)$$

and

$$\exp(-x) = \frac{1}{\exp(x)}.$$

(iv) For every $x \in \mathbb{R}$ we have $\exp(x) > 0$ *and for every $x > 0$ we have* $\exp(x) > 1$.

(v) The function $\exp(x)$ *is unbounded above, and for every $\varepsilon > 0$ there are x such that* $\exp(x) < \varepsilon$.

thm:six14

Proofs of (ii), (iii), (iv), (v). The formulæ in (ii) are immediate from the definition.

To prove (iii) observe first that by the ratio test

$$\sum_{m=0}^{\infty}\sum_{k=0}^{\infty}\frac{|x|^m|y|^k}{m!k!}$$

converges absolutely and so by the rearrangement theorem

$$\exp(x)\exp(y) = \sum_{m=0}^{\infty}\sum_{k=0}^{\infty}\frac{x^m y^k}{m!k!}$$

can be rearranged in any way we like. Thus it is

$$\begin{aligned}
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{\substack{k=0\\ m+k=n}}^{\infty}\frac{x^m y^k}{m!k!}\\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{x^m y^{n-m}}{m!(n-m)!}\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}x^m y^{n-m}\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{m=0}^{n}\binom{n}{m}x^m y^{n-m}\\
&= \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!}\\
&= \exp(x+y).
\end{aligned}$$

To prove (iv) note that it follows at once from the definition (6.9) when $x \geq 0$ and then from the last part of (iii) when $x < 0$. To prove the first part of (v) note that for any $n \in \mathbb{N}$ we have

$$\exp(n) = \sum_{m=0}^{\infty} \frac{n^m}{m!} > n$$

and we can apply the Archimedean property. If further we choose $n > 1/\varepsilon$ it follows that $\exp(-n) = 1/\exp(n) < 1/n < \varepsilon$ which establishes the second part. □

ex:six11 **Example 6.11.** *The series*

$$\sum_{n=1}^{\infty} \frac{x^{n^2}}{n^3} \tag{6.12}$$

eq:six11

has radius of convergence 1.

Proof. The series converges when $x = 0$ because all the terms except the first one are 0.

Suppose $x \neq 0$. Then consider

$$\left| \frac{x^{(n+1)^2}}{(n+1)^3} \frac{n^2}{x^{n^3}} \right| = |x|^{2n+1}(1 + 1/n)^{-3}.$$

This converges to 0 when $|x| < 1$, so by the ratio test the series also converges when $0 < |x| < 1$. On the other hand, when $|x| > 1$ the ratio is unbounded and so the series diverges. □

We can now combine a number of the concepts we have developed to show a connection between the sequences $\langle x_n \rangle$ and $\langle y_n \rangle$ of Exercise 5.1.1.5. and $\exp(x)$.

thm:six15 **Theorem 6.14.** *Suppose that* $x \in \mathbb{R}$. *Then*

$$\lim_{n\to\infty} (1 + x/n)^n = \lim_{n\to\infty} (1 - x/n)^{-n} = \exp(x).$$

Proof. We can certainly suppose that $x \neq 0$. Suppose in the first instance that $x > 0$ and apply the binomial theorem, Exercise 3.2.1.1 to obtain

$$(1 + x/n)^n = \sum_{m=0}^{n} \binom{n}{m} \frac{x^m}{n^m} = \sum_{m=0}^{n} \frac{x^m}{m!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m}{n}\right).$$

Let $M \in \mathbb{N}$ with $M \leq n$. Since all the terms above are non-negative, and the factors on the far right are all < 1, we have

$$\sum_{m=0}^{M} \frac{x^m}{m!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m}{n}\right) \leq (1 + x/n)^n \leq (1 + x/n)^n \leq \sum_{m=0}^{n} \frac{x^m}{m!}.$$

Now let $n \to \infty$. The sum on the far right converges to $\exp(x)$ and the one on the left has a fixed number of terms, so it also converges. We do not yet know that $(1 + x/n)^n$ converges, so we have

$$\sum_{m=0}^{M} \frac{x^m}{m!} \leq \liminf_{n\to\infty}(1 + x/n)^n \leq \limsup_{n\to\infty}(1 + x/n)^n \leq \exp(x).$$

Now let $M \to \infty$. Then the sum on the left also converges to $\exp(x)$. Hence we have

$$\exp(x) \leq \liminf_{n\to\infty}(1 + x/n)^n \leq \limsup_{n\to\infty}(1 + x/n)^n \leq \exp(x).$$

Since we now have equality throughout it follows that

$$\lim_{n\to\infty}(1 + x/n)^n$$

exists and equals the common value.

Now consider

$$(1 - x/n)^n(1 + x/n)^n = (1 - x^2/n^2)^n.$$

When $n > x$ we have $0 < x^2/n^2 < 1$ and so the expression on the right is < 1. Hence by the binomial inequality, Exercise 3.2.1.5,

$$1 - x^2/n \leq (1 - x/n)^n(1 + x/n)^n < 1.$$

Thus

$$(1 + x/n)^n < (1 - x/n)^{-n} < (1 + x/n)^n(1 - x^2/n)^{-1}$$

and by the sandwich theorem, Theorem 4.5,

$$\lim_{n\to\infty}(1 - x/n)^{-n} = \lim_{n\to\infty}(1 + x/n)^n = \exp(x).$$

Now suppose that $x < 0$. Then

$$(1 + x/n)^n = \frac{1}{(1 - (-x)/n)^{-n}}$$

and since $-x > 0$ it follows from the above and the combination theorem for sequences, Theorem 4.4, that the expression on the right has limit

$$\frac{1}{\exp(-x)} = \exp(x).$$

A similar argument pertains for $(1 - x/n)^{-n}$. $\square$

subsec:six5

6.5.1 Exercises

1. Prove that for every pair of real numbers x and y we have

$$\begin{aligned}\cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y),\\ \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y).\\ \cos^2(x) + \sin^2(x) = 1,&\ \text{Pythagorus' Theorem},\end{aligned}$$

2. Let R be the radius of convergence of

$$\sum_{n=0}^{\infty} a_n x^n.$$

Prove that

$$\sum_{n=1}^{\infty} n a_n x^{n-1}$$

and

$$\sum_{n=2}^{\infty} n(n-1) a_n x^{n-2}$$

both have radius of convergence R.

3. For each real number x define

$$\cosh(x) = \frac{\exp(x) + \exp(-x)}{2},\ \sinh(x) = \frac{\exp(x) - \exp(-x)}{2}.$$

Prove that

$$\cosh(0) = 1,\ \sinh(0) = 0$$

and that for every pair of real numbers x and y

$$\begin{aligned}\cosh(-x) &= \cosh(x),\ \sinh(-x) = -\sinh(x),\\ \cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(y),\\ \sinh(x+y) &= \sinh(x)\cosh(y) + \cosh(x)\sinh(y).\\ \cosh^2(x) - \sinh^2(x) = 1.&\end{aligned}$$

4. We already introduced the number e in Exercise 3.2.1.5 as

$$e = \lim_{n\to\infty} (1 + 1/n)^n.$$

Prove that if $m \in \mathbb{Z}$ we have

$$\exp(m) = e^m$$

and if further $n \in \mathbb{N}$, then

$$\exp(m/n) = e^{m/n}.$$

6.6 Notes

sec:six

§6.5.1, Exercise 1. There is an extensive description of Pythagorus' Theorem and its history at https://en.wikipedia.org/wiki/Pythagorean_theorem

Chapter 7

Limits of Functions

ch:seven

7.1 Functions

sec:seven1

def:seven1 **Definition 7.1.** *A function f from a set $\mathcal{A}$ to a set $\mathcal{B}$*

$$f : \mathcal{A} \mapsto \mathcal{B} : f(x) = y$$

is a rule which assigns to each $x \in \mathcal{A}$ a unique element $y \in \mathcal{B}$. The element $y \in \mathcal{B}$ is called the image of the element $x \in \mathcal{A}$ and we write

$$y = f(x).$$

If we know a formula for $f(x)$ we may alternatively write

$$x \mapsto f(x).$$

The set $\mathcal{A}$ is called the **domain** *of f. For $\mathcal{S} \subset \mathcal{A}$ we use the notation*

$$f(\mathcal{S}) = \{f(x); x \in \mathcal{S}\}$$

and we call $f(\mathcal{S})$ the **image of $\mathcal{S}$ under** *f. In the special case $\mathcal{S} = \mathcal{A}$ we call $f(\mathcal{A})$ simply the* **image or range of** *f. The set $\mathcal{B}$, which may have elements which are not in $f(\mathcal{A})$ is called the* **codomain of** *f. We can also think of the function f as being the set of ordered pairs (x, y) in which x and y are connected by the rule $y = f(x)$.*

When no element y of the codomain appears in more than one ordered pair, then the function is called **bijective***, which means that to each point in the image there is a unique member of the domain, i.e. there is an* **inverse function** *$f^{-1}(y) = x$ with the property that $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$.*

ex:seven1 **Example 7.1.** *Let $\mathbb{R}$ be the domain and codomain of the following function defined as the set of ordered pairs (x, x^2) with $x \in \mathbb{R}$. Then each positive member y of the codomain occurs in both $(-\sqrt{y}, y)$ and $(\sqrt{y}, y)$, but no negative number appears in the image. Of course this is the function $f(x) = x^2$.*

ex:seven2 **Example 7.2.** *The equation $y^2 = x$ with $x \in \mathbb{R}$ and $y \in \mathbb{R}$* **does not** *define a function from $\mathbb{R}$ to $\mathbb{R}$ because given $x > 0$ there are two values of y for which this holds.*

However if we take $\mathcal{A} = \{x : x \geq 0\}$, $\mathcal{B} = \{y : y \geq 0\}$, then the equation $y^2 = x$ **does** *define a function because given $x \in \mathcal{A}$ there is only one corresponding $y \in \mathcal{B}$. Of course this is the function $f(x) = \sqrt{x}$, where as usual this denotes the non-negative square root.*

def:seven2 **Definition 7.2.** *Suppose that the function f is defined on a subset $\mathcal{S}$ of $\mathbb{R}$ and its codomain is $\mathbb{R}$. Then we say that f is* **bounded above by** *H when the image $f(\mathcal{S})$ is bounded above by H. Likewise we define* **bounded below by** *h when the image is bounded below by h, and* **bounded** *when it is both bounded above and below.*

If $f(\mathcal{S})$ is non-empty and bounded above, then by the continuum property $\sup f(\mathcal{S})$ exists.

def:seven3 **Definition 7.3.** *When $\sup f(\mathcal{S})$ is non-empty and bounded above, and there is a $\xi \in \mathcal{S}$ so that $f(\xi) = \sup f(\mathcal{S})$, then we say that f has a maximum and the maximum is attained at $x = \xi$. Likewise when $f(\mathcal{S})$ is bounded below we use the corresponding term* **minimum** *for infima which are attained.*

ex:seven3 **Example 7.3.** *The function $f : (0, 1] \mapsto \mathbb{R} : f(x) = \frac{1}{x}$ is unbounded, but it is bounded below and $\inf f((0, 1]) = 1$, so it has minimum 1 which is attained with $x = 1$.*

An important class of functions are monotonic, which we define analogously to that for monotonic sequences.

def:seven3a **Definition 7.4.** *1. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are subsets of $\mathbb{R}$ and that $f : \mathcal{A} \mapsto \mathcal{B}$. We say that f is* **increasing** *when $f(x_1) \leq f(x_2)$ for every $x_1, x_2 \in \mathbb{R}$ with $x_1 \leq x_2$, and it is* **decreasing** *when $f(x_1) \geq f(x_2)$ for every such x_1, x_2.*

2. When $f(x_1) < f(x_2)$ for every pair x_1, x_2 with $x_1 < x_2$ we call it **strictly increasing***, and on the other hand when $f(x_1) > f(x_2)$ for every pair x_1, x_2 witn $x_1 < x_2$ we call it* **strictly decreasing***.*

3. Such functions are called **monotonic** *in case 1. and* **strictly monotonic** *in case 2.*

4. With reference to the last paragraph of Definition 7.1 it follows that every strictly monotonic function has an inverse from its image.

def:seven3- **Example 7.4.** *The function $\exp(x)$ defined by (6.9) is strictly increasing. To see this note that when $x_1 < x_2$ we have*

$$\exp(x_2) = \exp(x_1)\exp(x_2 - x_1)$$

and

$$\exp(x_2 - x_1) = \sum_{n=0}^{\infty} \frac{(x_2 - x_1)^n}{n!} > 1,$$

and moreover by Theorem 6.13 (iv) we have $\exp(x_1) > 0$.

In view of 4. above it follows that exp has an inverse function.

def:seven3+ **Definition 7.5.** *We define the function* $\log(x)$*, sometimes written* $\ln(x)$*, to be the inverse function of* $\exp(x)$*. The domain of* exp *is* $\mathbb{R}$ *and we will show in Corollary 8.8 that its image is*

$$\mathbb{R}^+ = \{x : x \in \mathbb{R} \text{ and } x > 0\}, \tag{7.1}$$

eq:seven1

the set of positive real numbers. Hence $\log(x)$ *has domain* $\mathbb{R}^+$ *and image* $\mathbb{R}$*. It also satisfies*

$$\log\big(\exp(x)\big) = x \text{ and } \exp\big(\log(y)\big) = y \tag{7.2}$$

eq:seven2

for $x \in \mathbb{R}$ *and* $y \in \mathbb{R}^+$*.*

Given u, v *in the domain of* log *there will be* $x, y \in \mathbb{R}$ *so that* $x = \log u$*,* $y = \log v$ *and so* $u = \exp(x)$*,* $v = \exp(y)$*. Thus* $uv = \exp(x)\exp(y) = \exp(x+y)$ *and*

$$\log(uv) = x + y = \log(u) + \log(v). \tag{7.3}$$

eq:seven3

We can now use this to define, whenever $a > 0$*,*

$$a^x : \mathbb{R} \mapsto \mathbb{R}^+ : x \mapsto \exp\big(x \log(a)\big). \tag{7.4}$$

eq:seven3x

7.1.1 Exercises

subsec:seven1

1. Sketch the set of points (x, y) in $\mathbb{R}^2$ for which $x^4 + y^2 = 1$. Explain why
(i) they do not define a function $f : \mathbb{R} \to \mathbb{R}$,
(ii) they do not define a function $f : [-1, 1] \to [-1, 1]$,
(iii) they do define a function $f : [-1, 1] \to [0, 1]$.

2. Let $\exp(x)$ be the function defined for $x \in \mathbb{R}$ by (6.9) and let $\mathcal{B} = f(\mathbb{R})$. Prove that $\inf \mathcal{B} = 0$, but $0 \notin \mathcal{B}$.

7.2 Limits

sec:seven2

For functions of a real variable, when we consider limits we are fundamentally looking at a real variable getting closer and closer to some real number ξ, rather than in the case of sequences where the variable n is getting larger and larger. Moreover when we consider x getting closer and closer to ξ we need to be impartial as to the sign of $x - \xi$, that is we want to look at both $x < \xi$ and $x > \xi$. We also want to avoid making any assumptions about the behaviour of f at ξ Thus in the first instance given a ξ we will restrict our attention to functions whose domain contains the two open intervals (a, ξ) and (ξ, b) where $a < \xi < b$.

def:seven4 **Definition 7.6** (Limit of a function)**.** *Suppose that* $a < \xi < b$*,* $\mathcal{A} \subset \mathbb{R}$ *and* $\mathcal{B} \subset \mathbb{R}$*,* $f : \mathcal{A} \mapsto \mathcal{B}$ *and* $(a, \xi) \cup (\xi, b) \in \mathcal{A}$*. Then*

$$\lim_{x \to \xi} f(x) = \ell,$$

or equivalently

$$f(x) \to \ell \text{ as } x \to \xi,$$

means that there is an $\ell \in \mathbb{R}$ *such that for every* $\varepsilon > 0$ *there is a* $\delta > 0$ *so that whenever* $x \in \mathcal{A}$ *and*

$$0 < |x - \xi| < \delta$$

we have

$$|f(x) - \ell| < \varepsilon.$$

See how the definition has a similar structure to the definition of limits for sequences. There is an ε in both which plays the rôle of measuring how close we are to the limit, and instead of N we have a δ which plays a similar rôle to N. We should expect that, just as for N, when we come to find a suitable δ it depends on ε.

We should also note the condition $0 < |x - \xi|$. We want to include the possibility that the limit ℓ differs from $f(\xi)$ if the latter should exist.

ex:seven4 **Example 7.5.** *Suppose that* $f : (0, 1) \mapsto \mathbb{R}$ *is defined by*

$$f(x) = \begin{cases} 0 & x \neq \frac{1}{2}, \\ 1 & x = \frac{1}{2}. \end{cases}$$

Then we have

$$\lim_{x \to \frac{1}{2}} f(x) = 0 \neq f(1/2).$$

To see this take $\delta = \frac{1}{2}$ *in the definition. Then for* $0 < |x - \frac{1}{2}| < \delta$*, so that* $0 < x < \frac{1}{2}$ *or* $\frac{1}{2} < x < 1$ *we have*

$$|f(x) - 0| = |0 - 0| = 0 < \varepsilon.$$

Here is a more typical example.

ex:seven5 **Example 7.6.** *Let* $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = x^2$ *and* $\xi \in \mathbb{R}$*. Then*

$$\lim_{x \to \xi} f(x) = \xi^2.$$

Proof. We guess that $\ell = \xi^2$. Let $\varepsilon > 0$. Choose

$$\delta = \min \left\{ 1, \frac{\varepsilon}{1 + 2|\xi|} \right\}.$$

Then whenever $0 < |x - \xi| < \delta$ we have, by the triangle inequality,

$$\begin{aligned}
|f(x) - \xi^2| &= |x^2 - \xi^2| \\
&= |x - \xi||x + \xi| \\
&= |x - \xi||(x - \xi) + 2\xi| \\
&\leq |x - \xi|(|x - \xi| + 2|\xi|) \\
&< \delta(\delta + 2|\xi|) \\
&\leq \frac{\varepsilon}{1 + 2|\xi|}(1 + 2|\xi|) \\
&= \varepsilon.
\end{aligned}$$

□

Here is an example where the limit does not exist.

ex:seven6 **Example 7.7.** *Let $f : (0, 2) \mapsto \mathbb{R}$ be defined by*

$$f(x) = \begin{cases} 0 & (0 \leq x \leq 1), \\ 1 & (1 < x < 2). \end{cases}$$

Then $\lim_{x \to 1} f(x)$ does not exist.

Proof. We argue by contradiction. Suppose the limit exists and equals ℓ. Choose $\varepsilon = \frac{1}{3}$ and choose $\delta > 0$ so that whenever $|x - 1| < \delta$ we have $|f(x) - \ell| < \varepsilon = \frac{1}{3}$. When $1 - \delta < x_1 < 1$ we have $f(x_1) = 0$ and when $1 < x_2 < 1 + \delta$ we have $f(x_2) = 1$. Hence, by the triangle inequality

$$\begin{aligned}
1 &= |f(x_2) - f(x_1)| \\
&= |(f(x_2) - \ell) - (f(x_1) - \ell)| \\
&\leq |f(x_2) - \ell| + |f(x_2) - \ell| \\
&< \frac{1}{3} + \frac{1}{3} \\
&= \frac{2}{3}.
\end{aligned}$$

□

ex:seven6a **Example 7.8.** *Let $f : \mathbb{R} \mapsto \mathbb{R} : x \mapsto x^3 + x$. Prove that $\lim_{x \to 2} f(x) = 10$.*

Proof. Let $\varepsilon > 0$. Choose $\delta = \min\left\{1, \frac{\varepsilon}{20}\right\}$. Then whenever $|x-2| < \delta$ we have

$$\begin{aligned}
|f(x) - 10| &= |x^3 + x - 10| \\
&= |(x-2)(x^2 + 2x + 5)| \\
&= |x-2||(x-2)^2 + 6(x-2) + 13| \\
&\leq |x-2|(|x-2|^2 + 6|x-2| + 13) \\
&< \delta(\delta^2 + 6\delta + 13) \\
&\leq \frac{\varepsilon}{20}(1^2 + 6 + 13) \\
&= \varepsilon.
\end{aligned}$$

□

Of course, as with sequences we will need to combine limits in order to deal with more complicated functions. The proofs of the next two theorems follow precisely the same pattern as for those for sequences and are left exercises.

thm:seven-

Theorem 7.1 (Combination Theorem for Functions)**.** *Suppose that $a < \xi < b$, $f, g : (a, \xi) \cup (\xi, b) \mapsto \mathbb{R}$, and $f(x) \to \ell$ and $g(x) \to m$ as $x \to \xi$. Suppose further that $\lambda, \mu \in \mathbb{R}$. Then*

(i) $\lambda f(x) + \mu g(x) \to \lambda\ell + \mu m$ as $x \to \xi$,

(ii) $f(x)g(x) \to \ell m$ as $x \to \xi$,

(iii) and when $m \neq 0$ we have

$$\frac{f(x)}{g(x)} \to \frac{\ell}{m}$$

as $x \to \xi$.

thm:seven0

Theorem 7.2 (Sandwich Theorem for Functions)**.** *Suppose that $a < \xi < b$, $f, g, h : (a, \xi) \cup (\xi, b) \mapsto \mathbb{R}$,*

$$g(x) \leq f(x) \leq h(x) \text{ when } x \in (a, \xi) \cup (\xi, b),$$

and $g(x) \to \ell$ and $h(x) \to \ell$ as $x \to \xi$. Then

$$f(x) \to \ell \text{ as } x \to \xi.$$

7.2.1 Limits at Infinity

def:seven6

Definition 7.7. *Suppose that $a \in \mathbb{R}$ and $f : (a, \infty) \mapsto \mathbb{R}$. Then by*

$$\lim_{x \to \infty} f(x) = \ell$$

we mean that for every $\varepsilon > 0$ there is an X such that whenever $x > X$ we have

$$|f(x) - \ell| < \varepsilon.$$

Similarly when $f : (-\infty, a) \mapsto \mathbb{R}$

$$\lim_{x \to -\infty} f(x) = \ell$$

we mean that for every $\varepsilon > 0$ *there is an* X *such that whenever* $x < X$ *we have*

$$|f(x) - \ell| < \varepsilon.$$

7.2.2 Exercises

subsec:seven2

1. Using only the definition of limit, prove that $\lim_{x \to 2} x^2 = 4$.

2. Using only the definition of limit, prove that
   (i) $\lim_{x \to 1} \frac{x}{1+x} = \frac{1}{2}$,
   (ii) $\lim_{x \to 2} x^2 + 3x = 10$.

3. Using only the definition of limit, prove that $\lim_{x \to 0} (1 + x)^{1/2} = 1$ (The identity $b^{1/2} - a^{1/2} = (b - a)/(b^{1/2} + a^{1/2})$ could be helpful in this question).

4. Prove, using only the definition of limit, that $\lim_{x \to 1} (5x - 3) = 2$.

5. Prove, using only the definition of limit, that $\lim_{x \to 1} (2x - 2) = 0$.

6. Let $n \in \mathbb{N}$ and $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = x^n$. Let $c \in \mathbb{R}$ and $g : \mathbb{R} \mapsto \mathbb{R} : g(x) = c$
   (i) Prove that for every $\xi \in \mathbb{R}$ we have $\lim_{x \to \xi} f(x) = \xi^n$.
   (ii) Prove that for every $\xi \in \mathbb{R}$ we have $\lim_{x \to \xi} g(x) = c$.
   (iii) Let $c_0, c_1, \ldots, c_m \in \mathbb{R}$ and $P : \mathbb{R} \mapsto \mathbb{R} : P(x) = c_0 + c_1 x + \cdots c_m x^m$. Prove that $\lim_{x \to \xi} P(x) = P(\xi)$.
   (iv) Let $d_0, d_1 \ldots, d_n \in \mathbb{R}$ and $Q : \mathbb{R} \mapsto \mathbb{R} : Q(x) = d_0 + d_1 x + \cdots d_n x^n$. Prove that if $Q(\xi) \neq 0$, then
$$\lim_{x \to \xi} \frac{P(x)}{Q(x)} = \frac{P(\xi)}{Q(\xi)}.$$

7. Evaluate the following limits.
$$\text{(i) } \lim_{x \to 3} \frac{x^3 + 5x + 7}{x^4 + 6x^2 + 8} \quad \text{(ii) } \lim_{x \to 3} \frac{x^2 - 4x + 3}{x^2 - 2x - 3}.$$

8. Prove that if $a < \xi < b$, $f : (a, b) \mapsto \mathbb{R}$ and $\lim_{x \to \xi} f(x) = \ell$, then $\lim_{x \to \xi} |f(x)| = |\ell|$ Disprove the converse of this statement.

7.3 One Sided Limits

sec:seven3

It can happen that sometimes we want to restrict our attention to one of the cases $x < \xi$ or $x > \xi$. Typically this happens when a function is only defined on a closed interval

$[a, b]$ and we want to understand the limiting behaviour at a and b. It can also happen with examples like $f : [0, 2] \mapsto \mathbb{R}$

$$f(x) = \begin{cases} 0 & (0 \le x < 1), \\ 1 & (x = 1), \\ 2 & (1 < x \le 2) \end{cases}$$

when $\xi = 1$.

Thus we introduce a variant of our definition of limit.

def:seven5 **Definition 7.8** (Limit from above and below)**.** *Suppose that $\mathcal{A} \subset \mathbb{R}$ and $\mathcal{B} \subset \mathbb{R}$, $f : \mathcal{A} \mapsto \mathcal{B}$, $a < \xi$ and $(a, \xi) \in \mathcal{A}$. Then*

$$\lim_{x \to \xi -} f(x) = \ell$$

means that there is an $\ell \in \mathbb{R}$ such that for every $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $x \in \mathcal{A}$ and

$$\xi - \delta < x < \xi$$

we have

$$|f(x) - \ell| < \varepsilon$$

and we call ℓ the limit from below.

There is a corresponding definition for limit from above. Suppose that $\mathcal{A} \subset \mathbb{R}$ and $\mathcal{B} \subset \mathbb{R}$, $f : \mathcal{A} \mapsto \mathcal{B}$, $\xi < b$ and $(\xi, b) \in \mathcal{A}$. Then

$$\lim_{x \to \xi +} f(x) = \ell$$

means that there is an $\ell \in \mathbb{R}$ such that for every $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $x \in \mathcal{A}$ and

$$\xi < x < \xi + \delta$$

we have

$$|f(x) - \ell| < \varepsilon$$

and we call ℓ the limit from above.

ex:seven7 **Example 7.9.** *Suppose that $f : [0, \infty) \mapsto \mathbb{R} : f(x) = \sqrt{x}$. Then* $\lim_{x \to 0+} f(x) = 0$.

Proof. Let $\varepsilon > 0$. Choose $\delta = \varepsilon^2$. Then, whenever $0 < x < \delta$ we have

$$|f(x) - 0| = \sqrt{x} < \sqrt{\delta} = \varepsilon.$$

Note that $\lim_{x \to 0} f(x)$ and $\lim_{x \to -0} f(x)$ do not exist. $\square$

As you might expect, if the limits from below and above exist and agree, then the limit does exist.

thm:seven1 **Theorem 7.3.** *Suppose that $a < \xi < b$ and $f : (a, b) \mapsto \mathbb{R}$. Then the limit*

$$\lim_{x \to \xi} f(x)$$

exists and converges to ℓ if and only if both the limits

$$\lim_{x \to \xi -} f(x), \quad \lim_{x \to \xi +} f(x)$$

exist and converge to ℓ.

The proof is immediate on comparing the definitions.

7.3.1 Exercises

subsec:seven3

1. When $x \in \mathbb{R}$ let $\lfloor x \rfloor$ denote the largest integer m not exceeding x, for example $\lfloor -\sqrt{2} \rfloor = -2$, $\lfloor \frac{7}{2} \rfloor = 3$ and define

$$f(x) = x - \lfloor x \rfloor - \tfrac{1}{2}$$

and

$$g(x) = \tfrac{1}{2}(x - \lfloor x \rfloor)^2 - \tfrac{1}{2}(x - \lfloor x \rfloor) + \tfrac{1}{12}.$$

(i) Prove that $\lim_{x \to 0-} f(x) = \frac{1}{2}$ and $\lim_{x \to 0+} f(x) = -\frac{1}{2}$.
(ii) Prove that $\lim_{x \to 0} g(x) = \frac{1}{12}$.

2. Suppose that $a < b$, $f : (a, b) \mapsto \mathbb{R}$, and f is bounded and monotonic. Prove that $\lim_{x \to a+} f(x)$ and $\lim_{x \to b-} f(x)$ both exist.

3. State and prove the analogues for one-sided limits of Theorems 7.1 and 7.2.

4. Suppose that $f : (1, \infty) \mapsto \mathbb{R}$. Prove that $\lim_{x \to \infty} f(x) = \ell$ if and only if

$$\lim_{x \to 0+} f(1/x) = \ell.$$

7.4 Notes

sec:seven8

§7.3. The functions f and g of exercise 7.3.1.1 are the first two periodic Bernoulli polynomials, with g normalised so that when $x \notin \mathbb{Z}$ we have $g'(x) = f(x)$. See
https://en.wikipedia.org/wiki/Bernoulli_polynomials

Chapter 8

Continuity

ch:eight

8.1 Continuity at a Point

sec:eight1

The concept of continuity is fundamental to much of mathematics. We start with continuity at a point.

def:eight1 **Definition 8.1** (Continuity at a Point)**.** *Suppose that $a < \xi < b$ and $f : (a, b) \mapsto \mathbb{R}$. Then we say that f is* **continuous at** *ξ when $f(x) \to f(\xi)$ as $x \to \xi$. Otherwise it is* **discontinuous at** *ξ.*

ex:eight1 **Example 8.1.** *Suppose that $c_0, c_1, \ldots, c_m \in \mathbb{R}$, $\xi \in \mathbb{R}$ and*

$$P(x) = c_0 + c_1 x + \cdots c_m x^m \, (x \in \mathbb{R}).$$

Then by Exercise 7.2.2.4(iii) it follows that P is continuous at ξ. In other words every polynomial is continuous at every real number x. More generally, it follows from Exercise 7.2.2.4(iv) that every rational function $\frac{P(x)}{Q(x)}$ is continuous at every real number ξ for which $Q(\xi) \neq 0$.

ex:eight2 **Example 8.2.** *The function f of Exercise 7.3.1.1 is discontinuous at 0, but the function g of that exercise is continuous at 0.*

It is also important, especially when dealing with intervals, to deal with one sided continuity.

def:eight2 **Definition 8.2** (Continuity from the Left or Right)**.** *Suppose that $a < \xi$ and $f : (a, \xi] \mapsto \mathbb{R}$. Then we say that f is* **continuous from below, or from the left, at** *ξ when $f(x) \to f(\xi)$ as $x \to \xi-$. Otherwise we say that f is* **discontinuous from the left at** *ξ.*

Suppose that $\xi < b$ and $f : [\xi, b) \mapsto \mathbb{R}$. Then we say that f is **continuous from above, or from the right, at** *ξ when $f(x) \to f(\xi)$ as $x \to \xi+$. Otherwise we say that f is* **discontinuous from the right at** *ξ.*

ex:eight3 **Example 8.3.** *In Exercise 7.3.1.1 the function f is discontinuous from the left at 0, but continuous from the right at 0.*

Since continuity at a point is simply about a limit at one point the following is immediate from the combination theorem for functions, Theorem 7.1.

thm:eight0 **Theorem 8.1** (Combination Theorem for Pointwise Continuity). *Suppose that $a < \xi < b$, $f, g : (a, b) \mapsto \mathbb{R}$, and $f(x)$ and $g(x)$ are continuous at ξ. Suppose further that $\lambda, \mu \in \mathbb{R}$. Then*
(i) $\lambda f(x) + \mu g(x)$ is continuous at ξ,
(ii) $f(x)g(x)$ is continuous at ξ,
(iii) if $g(\xi) \neq 0$, then

$$\frac{f(x)}{g(x)}$$

is continuous at ξ.

thm:eight0+ **Theorem 8.2** (Combination Theorem for One Sided Continuity). *Suppose that $a < \xi$, $f, g : (a, \xi] \mapsto \mathbb{R}$, and $f(x)$ and $g(x)$ are continuous from below at ξ. Suppose further that $\lambda, \mu \in \mathbb{R}$. Then*
(i) $\lambda f(x) + \mu g(x)$ is continuous from below at ξ,
(ii) $f(x)g(x)$ is continuous from below at ξ,
(iii) if $g(\xi) \neq 0$, then

$$\frac{f(x)}{g(x)}$$

is continuous from below at ξ.

There are corresponding statements for continuity from above when f and g are defined on $[\xi, b)$.

The following is also useful.

thm:eight0- **Theorem 8.3.** *Suppose $a < b$, $\xi \in (a, b)$. Then $f : (a, b) \mapsto \mathbb{R}$ is continuous at ξ if and only if for* **every** *sequence $\langle x_n \rangle$ in (a, b) satisfying $\lim_{n\to\infty} x_n = \xi$ we have*

$$\lim_{n\to\infty} f(x_n) = f(\xi).$$

We leave the proof to the reader.

8.1.1 Exercises

subsec:eight1

1. Suppose that $a < \xi < b$, $f : (a, b) \mapsto \mathbb{R}$, $c < f(\xi) < d$, $g : (c, d) \mapsto \mathbb{R}$, f is continuous at ξ and g is continuous at $f(\xi)$. Let $h : (a, b) \mapsto \mathbb{R} : h(x) = f\big(g(x)\big)$. Prove that h is continuous at ξ.

2. Prove that

(i) $f(x) = 2x^3 - 4x^2 + 5$ is continuous at each $\xi \in \mathbb{R}$,
(ii) $g(x) = 1/(x^2 - 4)$ is continuous at each $\xi \in \mathbb{R}\backslash\{-2, 2\}$.

3. Let $\exp(x) : x \mapsto \mathbb{R}$ be the function defined in (6.9) and log the function defined in Definition 7.5.
(i) Prove that $\exp(x)$ is continuous at 0.
(ii) Prove that $\exp(x)$ is continuous at every $\xi \in \mathbb{R}$.
(iii) Prove that $\log(x)$ is continuous at every $\eta \in \mathbb{R}^+$.

2. Let $\cos(x) : x \mapsto \mathbb{R}$ and $\sin(x) : x \mapsto \mathbb{R}$ be the functions defined in (6.10) and (6.11).
(i) Prove that $\cos(x)$ and $\sin(x)$ are continuous at 0.
(ii) Prove that $\sin(x)$ and $\cos(x)$ are continuous at every $\xi \in \mathbb{R}$.
(iii) When $\cos(x) \neq 0$, define

$$\tan(x) = \frac{\sin(x)}{\cos(x)}.$$

Prove that if $\cos(\xi) \neq 0$, then $\tan(x)$ is continuous at ξ.

3. Let $\cosh(x) : x \mapsto \mathbb{R}$ and $\sinh(x) : x \mapsto \mathbb{R}$ be the functions defined in Exercise 6.5.1.3.
(i) Prove that $\cosh(x)$ and $\sinh(x)$ are continuous at 0.
(ii) Prove that $\sinh(x)$ and $\cosh(x)$ are continuous at every real $\xi \in \mathbb{R}$.
(iii) Define

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}.$$

Prove that $\tanh(x)$ is continuous at every $\xi \in \mathbb{R}$.

4. Define $f : \mathbb{R} \mapsto \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{when } x \text{ is irrational,} \\ \frac{1}{q} & \text{when } x = \frac{r}{q} \text{ with } r \in \mathbb{Z},\ q \in \mathbb{N}, (r, q) = 1. \end{cases}$$

Prove that f is continuous at each $x \in \mathbb{R} \setminus \mathbb{Q}$ and discontinuous at each $x \in \mathbb{Q}$.

5. Prove that $f : \mathbb{R} \mapsto R : f(x) = |x|$ is continuous at all $\xi \in \mathbb{R}$.

8.2 Continuity on an Interval

sec:eight2

Continuity at an individual point is not particularly useful. Most functions we meet are continuous on an interval.

def:eight3

Definition 8.3. *Suppose that $a < b$, I is the* **open interval** (a, b) *and $f : I \mapsto \mathbb{R}$. Then f is continuous on I when it is continuous at every point $\xi \in I$.*

Suppose instead that I is the **closed interval** $[a, b]$ *and $f : I \mapsto \mathbb{R}$. Then f is continuous on I when it is continuous at every point $\xi \in (a, b)$, continuous from the right at a and continuous from the left at b. When f is continuous at every $\xi \in \mathbb{R}$ then we say that f is continuous on $\mathbb{R}$.*

Of course when f is continuous on every interval $I \subset \mathbb{R}$, then it is continuous on $\mathbb{R}$.

ex:eight3a **Example 8.4.** *The function* $f : (0,1) \mapsto \mathbb{R} : f(x) = \frac{1}{x}$ *is continuous on* $(0,1)$*, even though* f *is unbounded.*

ex:eight4 **Example 8.5.** *Let* $c_0, c_1, \ldots, c_n \in \mathbb{R}$ *and* $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = c_0 + c_1 x + \cdots + c_n x^n$*. Then* f *is continuous on* $\mathbb{R}$*.*

ex:eight5 **Example 8.6.** *Let* f *be the example of Exercise 7.3.1.1,*

$$f(x) = x - \lfloor x \rfloor - \frac{1}{2}.$$

Then, for any $k \in \mathbb{Z}$*,* f *is continuous on* $(k, k+1)$ *but discontinuous on* $[k, k+1]$*.*

thm:eight1- **Theorem 8.4** (Combination Theorem for Continuity on an Interval)**.** *Suppose that* $a < b$ *and* $I = (a,b)$ *or* $[a,b]$*,* $f, g : I \mapsto \mathbb{R}$*, and* $f(x)$ *and* $g(x)$ *are continuous on* I*. Suppose further that* $\lambda, \mu \in \mathbb{R}$*. Then*
(i) $\lambda f(x) + \mu g(x)$ *is continuous on* I*,*
(ii) $f(x)g(x)$ *is continuous on* I*,*
(iii) if $g(x) \neq 0$ *for* $x \in I$*, then*

$$\frac{f(x)}{g(x)}$$

is continuous on I*.*

Continuity on a closed interval is much more constraining than continuity on an open interval, but comes with many benefits. Example 8.3 can be contrasted with the following important theorem.

thm:eight1 **Theorem 8.5.** *Suppose that* $a < b$ *and* $f : [a,b] \mapsto \mathbb{R}$ *is continuous on* $[a,b]$*. Then* f *is bounded. That is, there is a* $B \in \mathbb{R}$ *such that* $|f(x)| \leq B$ *for any* $x \in [a,b]$*.*

Proof. Suppose that $f([a,b])$ is unbounded above. If instead it is unbounded below we can replace f by $-f$. Then given any $n \in \mathbb{N}$ there is an $x_n \in [a,b]$ such that

$$f(x_n) > n. \tag{8.1}$$ eq:eight1

But $\langle x_n \rangle$ is a bounded sequence because $a \leq x_n \leq b$. Hence, by the Bolzano-Weierstrass theorem, Theorem 5.4, it has a convergent subsequence $\langle x_{m_n} \rangle$. Let the limit be ℓ. Then $a \leq \ell \leq b$ and

$$\ell = \lim_{n \to \infty} x_{m_n}.$$

The function f is continuous at ℓ. Hence there is a $\delta > 0$ so that whenever $|x_{m_n} - \ell| < \delta$ we have $|f(x_{m_n}) - f(\ell)| < 1$. Therefore, by the triangle inequality, for every $n \in \mathbb{N}$,

$$n \leq m_n < f(x_{m_n}) = (f(x_{m_n}) - \ell) + \ell \leq |f(x_{m_n} - \ell| + |\ell| < 1 + |\ell|$$

contradicting the Archimedean property of $\mathbb{N}$. $\square$

This leads to the following remarkable and very useful result.

thm:eight2 **Theorem 8.6.** *Suppose that $a < b$ and f is continuous on $[a,b]$. Then f attains its bounds. In other words there are $\eta, \xi \in [a,b]$ such that*

$$f(\eta) = \inf f([a,b]) \tag{8.2}$$ eq:eight4

and

$$f(\xi) = \sup f([a,b]). \tag{8.3}$$ eq:eight5

Proof. It suffices to prove (8.3) since (8.2) then follows by replacing f by $-f$.

We argue by contradiction. Let

$$\Lambda = \sup f([a,b]) \tag{8.4}$$ eq:eight6

and suppose that $f(x) < \Lambda$ for every $x \in [a,b]$. Let

$$g(x) = \frac{1}{\Lambda - f(x)}.$$

Then by the combination theorem for continuity on an interval, Theorem 8.4 it follows that g is continuous on $[a,b]$ and by Theorem 8.5 is bounded on I. Hence there is a $B > 0$ such that for every $x \in I$ we have

$$\frac{1}{\Lambda - f(x)} = g(x) < B.$$

Hence

$$0 < \frac{1}{B} < \Lambda - f(x),$$

$$f(x) < \Lambda - \frac{1}{B}$$

which contradicts the definition of Λ, (8.4). □

8.2.1 Exercises

subsec:eight2

1. Suppose that $a < b$. Prove that the function $f : (a,b) \mapsto \mathbb{R} : f(x) = \frac{1}{(x-a)(b-x)}$ is continuous on (a,b).

2. Alternative proof of Theorem 8.5. Let f be as in that theorem, but suppose it is unbounded. Construct two sequences $\langle a_n \rangle$ and $\langle b_n \rangle$ in $[a,b]$ with the following properties

(1) $\langle a_n \rangle$ is increasing,

(2) $\langle b_n \rangle$ is decreasing,

(3) $b_n - a_n = \dfrac{b-a}{2^{n-1}}$, (4) f is unbounded on $[a_n, b_n]$.

(i) Prove that $\lim_{n\to\infty} a_n$ and $\lim_{n\to\infty} b_n$ each exist and are equal.

(ii) Let ξ be the common limit. Prove that there is $\delta > 0$ such that f is bounded on $I = (\xi - \delta, \xi + \delta) \cap [a, b]$.

(iii) Prove that for some n, $[a_n, b_n] \subset I$.

3. Give an example of a function $f : (0, 1) \mapsto \mathbb{R}$ such that f is continuous at exactly one point of $(0, 1)$

8.3 The Intermediate Value Theorem

sec:eight3

We now come to a theorem which is used all the time in applications. It is especially important in that it underpins all zero finding techniques for continuous functions.

thm:eight3

Theorem 8.7. *Suppose that $a < b$, $f : [a, b] \mapsto \mathbb{R}$ is continuous on $[a, b]$ and*

$$\inf f([a, b]) \le \lambda \le \sup f([a, b]).$$

Then there is a $\xi \in [a, b]$ such that

$$f(\xi) = \lambda.$$

rem:eight1

Remark 8.1. *This theorem says in effect that, when f is continuous on a closed interval $[a, b]$, the set $f([a, b])$ is also an interval.*

Proof. We construct two sequences $\langle a_n \rangle$ and $\langle b_n \rangle$ such that

(1) $\langle a_n \rangle$ is increasing and $a \le a_n \le b$,

(2) $\langle b_n \rangle$ is decreasing and $a \le b_n \le b$,

(3) $0 < b_n - a_n = \frac{b_1 - a_1}{2^{n-1}}$,

(4) $(f(a_n) - \lambda)(f(b_n) - \lambda) \le 0$.

By Theorem 6 there are $u, v \in [a, b]$ so that $f(u) = \inf f([a, b])$, $f(v) = \sup f([a, b])$. Hence $(f(u) - \lambda))(f(v) - \lambda) \le 0$. Let $a_1 = \min\{u, v\}$, $b_1 = \max\{u, v\}$. Then (4) holds with $n = 1$,

Given a_n and b_n satisfying (3), (4), choose

$$c_n = \frac{a_n + b_n}{2}.$$

The inequality (4) says that at least one of the two factors $f(a_n) - \lambda$ and $f(b_n) - \lambda$ is 0, or they are both non-zero and have opposite signs. If $f(a_n) - \lambda = 0$ let $a_{n+1} = a_n$, $b_{n+1} = c_n$. If $f(a_n) - \lambda \ne 0$ but $f(c_n) - \lambda = 0$ let $a_{n+1} = c_n$, $b_{n+1} = b_n$. If $f(a_n) - \lambda \ne 0$ and $f(c_n) - \lambda \ne 0$ and they have opposite signs, then we take $a_{n+1} = a_n$, $b_{n+1} = c_n$. If $f(a_n) - \lambda \ne 0$ and $f(c_n) - \lambda \ne 0$ and they have the same sign, then we take $a_{n+1} = c_n$, $b_{n+1} = b_n$. In any case we have (1), (2), (3), (4) with n replaced by $n + 1$, and so the construction proceeds inductively.

By (1), (2), the monotonic convergence theorem and (3) the sequences $\langle a_n \rangle$ and $\langle b_n \rangle$ converge to a common value, say ξ. Thus, by Theorem 8.3,

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = f(\xi)$$

Hence, by Theorem 4.6 and (4)

$$(f(\xi) - \lambda)^2 \leq 0.$$

Therefore

$$f(\xi) = \lambda$$

as required.

□

thm:eight4 **Corollary 8.8.** *The image of* exp *is* $\mathbb{R}^+$

Proof. This follows at once from Theorems 8.7 and 6.13 (v). □

ex:eight6 **Example 8.7.** *Let* $f(x) = x^2 - x$ *be defined on* $I = \left[-\frac{1}{2}, \frac{3}{4}\right]$. *Then*

$$\inf f(I) = -\tfrac{1}{4},\ f\left(\tfrac{1}{2}\right) = -\tfrac{1}{4},$$

$$\sup f(I) = \tfrac{3}{4},\ f\left(-\tfrac{1}{2}\right) = \tfrac{3}{4},$$

$$-\tfrac{1}{4} < \tfrac{5}{16} < \tfrac{3}{4}, f\left(-\tfrac{1}{4}\right) = \tfrac{5}{16}.$$

ex:eight7 **Example 8.8.** *Prove that the cubic equation* $x^3 - 3x^2 + 1 = 0$ *has* 3 *real roots.*

Proof. For brevity write $f(x) = x^3 - 3x^2 + 1$. Then

$$f(-1) = -3,\ f(0) = 1,\ f(1) = -1,\ f(3) = 1$$

and f is continuous on each of the intervals $[-1, 0]$, $[0, 1]$, $[1, 3]$. Hence there are ξ_1, ξ_2, ξ_3 so that

$$-1 < \xi_1 < 0 < \xi_2 < 1 < \xi_3 < 3$$

and

$$f(\xi_1) = f(\xi_2) = f(\xi_3) = 0.$$

□

ex:eight8 **Example 8.9.** *Prove that the curve* $y = x^2$ *intersects the curve* $y = x^3 - 2x^2 + 1$ *in three places.*

Proof. At a point of intersection $x^2 = x^3 - 2x^2 + 1$, so that $x^3 - 3x^2 + 1 = 0$. Hence see previous example. □

ex:eight9 **Example 8.10.** *Suppose that* f *is continuous on* $[0, 1]$ *and* $f(0) = f(1)$. *Prove that there is a* $\xi \in [0, 1]$ *so that*

$$f(\xi) = f(\xi + 1/2).$$

This says that there are always two diametrically opposite points on the equator which have the same temperature.

Proof. Let $g(x) = f(x) - f(x + 1/2)$. Then g is continuous on $[0, 1/2]$. If $f(0) = f(1/2)$, then we are done. Suppose $f(0) \neq f(1/2)$. Then $g(0) = f(0) - f(1/2)$ and $g(1/2) = f(1/2) - f(1) = f(1/2) - f(0) = -(f(0) - f(1/2))$. Hence g changes sign on $[0, 1/2]$. Thus, by the Intermediate Value Theorem there is a $\xi \in (0, 1/2)$ such that $g(\xi) = 0$ and we are done once more. $\square$

We can also now say something more about the trigonometric functions sin and cos.

thm:eight5 **Theorem 8.9.** *The function* cos *changes sign on the interval* $[0, 2]$. *We define* $\frac{\pi}{2}$ *to be the smallest positive zero of* cos. *Then* cos *and* sin *are periodic with period* 2π *and*

$$\sin(0) = \sin(\pi) = 0,\ \sin\frac{\pi}{2} = 1,\ \sin\frac{3\pi}{2} = -1,\ \cos(x) = \sin\left(\frac{\pi}{2} - x\right).$$

Proof. By the definition of cos, (6.11), we have $\cos(0) = 1$ and

$$\begin{aligned}\cos(2) &= 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \sum_{k=2}^{\infty} \frac{2^{4k-2}}{(4k-2)!}\left(1 - \frac{2^2}{(4k-1)4k}\right) \\ &< 1 - 2 + \frac{2}{3} = -\frac{1}{3}.\end{aligned}$$

Hence, by the Intermediate Value Theorem, Theorem 8.7, there is an x with $0 < x < 2$ and $\cos(x) = 0$. Let $\varpi = \inf\{x : x > 0, \cos(x) = 0\}$. Then, by continuity, $\cos(\varpi) = 0$ and since $\cos(0) = 1$ we have $\varpi > 0$. Define

$$\pi = 2\varpi.$$

For any non-negative integer k, when $0 < x \leq 2$ we have

$$\frac{x^{4k+1}}{(4k+1)!} - \frac{x^{4k+3}}{(4k+3)!} = \frac{x^{4k+1}}{(4k+1)!}\left(1 - \frac{x^2}{(4k+2)(4k+3)}\right) > 0.$$

Hence, by the definition of sin, (6.10), we have $\sin(\varpi) > 0$.

By the addition formulæ Exercise 6.5.1.1, we have

$$\begin{aligned}\sin(\pi) &= 2\sin(\varpi)\cos(\varpi) = 0 \\ \cos(\pi) &= 2(\cos(\varpi))^2 - 1 = -1, \\ \cos(2\pi) &= 1 - 2(\sin(\pi))^2 = 1, \\ \sin(2\pi) &= 2\sin(\pi)\cos(\pi) = 0, \\ \sin(x + 2\pi) &= \sin(x)\cos(2\pi) + \cos(x)\sin(2\pi) = \sin(x), \\ \cos(x + 2\pi) &= \cos(x)\cos(2\pi) - \sin(x)\sin(2\pi) = \cos(x), \\ -1 = \cos(\pi) &= 1 - 2\sin^2(\varpi), \\ \sin^2(\varpi) &= 1, \\ \sin(\varpi) &= 1, \\ \cos(-x) &= \cos(x), \\ \sin(-x) &= -\sin(x).\end{aligned}$$

Thus

$$\begin{aligned}\sin(\varpi - x) &= \sin(\varpi)\cos(-x) + \cos(\varpi)\sin(-x)\\ &= \cos(x),\\ \sin(3\varpi) &= \sin(\varpi + \pi)\\ &= \cos(-\pi)\\ &= \cos(\pi)\\ &= -1.\end{aligned}$$

□

Observe that we nowhere used derivatives.

8.3.1 Exercises

subsec:eight3

1. Prove that the equation $x^{3456} + x^{1234} - 1 = 0$ has a solution with $0 < x < 1$.

2. Prove that the line $y = 2x$ intersects the cubic curve $y = x^3 - x + 1$ in at least three distinct points.

3. Prove that the curve $y = x^2$ intersects the cubic curve $y = x^4 - 2x^2 - x + 1$ in at least four distinct points.

4. Prove that the quintic equation $x^5 - 4x^2 + 2 = 0$ has at least three real roots.

8.4 Uniform Continuity

sec:eight4

Consider a real valued function defined on some domain $\mathcal{D} \in \mathbb{R}$, $f : \mathcal{D} \mapsto \mathbb{R}$. Then the definition of continuity, Definition 8.1 is a pointwise definition, even in the special case of an interval, Definition 8.3. This runs into the problem in applications that, given $\xi \in \mathcal{D}$ and $\varepsilon > 0$, the choice of δ can depend on ε *and* ξ.

ex:eight10

Example 8.11. *Let $f : (0,1) \mapsto \mathbb{R} : f(x) \mapsto \frac{1}{x}$. Suppose $0 < \varepsilon < \xi$. Then given $\xi \in (0,1)$ we need to find $\delta > 0$ so that when $0 < |x - \xi| < \delta$ we have $|f(x) - f(\xi)| < \varepsilon$, that is*

$$\left|\frac{1}{x} - \frac{1}{\xi}\right| < \varepsilon$$

or equivalently

$$|\xi - x| < \varepsilon x\xi < \varepsilon\xi(x - \xi) + \varepsilon\xi^2.$$

This has to hold for every x with $\xi - \delta < x < \xi + \delta$ and so taking x arbitrarily close to $x - \delta$ we must have $\delta \le -\varepsilon\xi\delta + \varepsilon\xi^2$ and so

$$\delta < \frac{\varepsilon\xi^2}{1 + \varepsilon\xi}.$$

Now δ cannot be taken to be independent of ξ, for taking ξ arbitrarily close to 0 would contradict $\delta > 0$.

When we have a situation in which it is possible to find a universal δ it is usual to associate the word *uniform* with it.

def:eight4 **Definition 8.4.** *Suppose that $\mathcal{S} \subset \mathbb{R}$ and $f : \mathcal{S} \mapsto \mathbb{R}$ has the property that for every $\varepsilon > 0$ there is a $\delta > 0$ such that whenever $x, y \in \mathcal{S}$ and $|x - y| < \delta$ we have*

$$|f(x) - f(y)| < \varepsilon \tag{8.5}$$ eq:eight7

then we say that f is uniformly continuous on $\mathcal{S}$. An equivalent statement is that for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sup\{|f(x) - f(y)| : x, y \in \mathcal{S} \text{ and } |x - y| < \delta\} < \varepsilon.$$

The following theorem highlights an important difference between open and closed intervals.

thm:eight6 **Theorem 8.10.** *Suppose that $a < b$, $f : [a, b] \mapsto \mathbb{R}$ and f is continuous on $[a, b]$. Then f is uniformly continuous on $[a, b]$.*

Proof. The important ingredients of the proof are (i) that $[a, b]$ is bounded, so that by the Bolzano-Weierstrass theorem any sequence restricted to the interval will have a convergent subsequence, and (ii) we can combine this with the definition of pointwise continuity.

As usual we argue by contradiction. Suppose that f is not uniformly continuous on $[a, b]$. Then there will be an $\varepsilon > 0$ such that for every $n \in \mathbb{N}$ there will be $x_n, y_n \in [a, b]$ such that $0 < |x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| \geq \varepsilon$. The members of the sequence $\langle x_n \rangle$ lie in $[a, b]$. Hence the sequence is bounded and so by Theorem 5.4 it has a convergent subsequence $\langle x_{m_n} \rangle$. Likewise the sequence $\langle y_{m_n} \rangle$ has a convergent subsequence $\langle y_{m_{k_n}} \rangle$. Moreover

$$|x_{m_{k_n}} - y_{m_{k_n}}| < \frac{1}{m_{k_n}} \to 0 \text{ as } n \to \infty.$$

Hence they have a common limit, ℓ, and $\ell \in [a, b]$. Thus, by the continuity of f at ℓ and Theorem 8.3

$$0 = \lim_{n \to} |f(x_{m_{k_n}}) - f(y_{m_{k_n}})| \geq \varepsilon$$

which gives the required contradiction. □

8.4.1 Exercises

subsec:eight4

1. Suppose that $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = |x|$. Show that f is uniformly continuous on $\mathbb{R}$.
2. Prove that
   (i) if $x \in \mathbb{R}$, then $|\sin(x)| \leq |x|$ and $|\cos(x) - 1| \leq \frac{|x|^2}{2}$,
   (ii) that $\cos(x)$ is uniformly continuous on $\mathbb{R}$,
   (iii) and that $\sin(x)$ is uniformly continuous on $\mathbb{R}$.

8.5 Notes

sec:eight8

The first definition of uniform continuity is in work of Bolzano and he gave the first proof that a continuous function on an open interval need not be uniformly continuous. He also states that a continuous function on a closed interval is uniformly continuous, but he does not give a complete proof. See `https://en.wikipedia.org/wiki/Uniform_continuity` The subject become more interesting when Weierstrass, *Über continuirliche Functionen eines reellen Arguments, die für keinen Werth des letzeren einen bestimmten Differentialquotienten besitzen*, Königlich Preussichen Akademie der Wissenschaften, 1872, gave an example of a function which is uniformly continuous on $\mathbb{R}$ but nowhere differentiable! See `https://en.wikipedia.org/wiki/Weierstrass_function`

Chapter 9

Differentiation

ch:nine

9.1 The Derivative

sec:nine1

The core idea driving differentiation is the need to find the slope of a curve $y = f(x)$ at a particular point $(\xi, f(\xi)$. In the definition below the motivation is the hope that the ratio

$$\frac{f(x) - f(\xi)}{x - \xi}$$

will approach that slope as $x \to \xi$.

def:nine1 **Definition 9.1** (Derivative)**.** *Suppose that $a < \xi < b$ and $f : (a, b) \mapsto \mathbb{R}$. Then we say that f is* **differentiable at** ξ *when*

$$\lim_{x \to \xi} \frac{f(x) - f(\xi)}{x - \xi}$$

exists, and then we write $f'(\xi)$ for the value of the derivative. Alternatively we might write

$$\lim_{h \to 0} \frac{f(\xi + h) - f(\xi)}{h}.$$

Note that it is crucial for this to make sense that in the definition of limit, Definition 7.6, we have $0 < |x - \xi|$ so as to avoid division by 0.

ex:nine0 **Example 9.1.** *Let $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = x^n$ where $n \in \mathbb{N}$. Then, by the binomial theorem*

$$(\xi + h)^n = \sum_{m=0}^{n} \binom{n}{m} h^m \xi^{n-m}$$

so that

$$\begin{aligned} \frac{f(\xi + h) - f(\xi)}{h} - nx^{n-1} &= \sum_{m=2}^{n} \binom{n}{m} h^{m-1} \xi^{n-m} \\ &\to 0 \text{ as } h \to 0. \end{aligned}$$

ex:nine1 **Example 9.2.** *For each $x \in \mathbb{R}$ let $f(x) = 2x^5 - 4x^3 + 2x - 1$. Then*

$$\begin{aligned}
(\xi + h)^5 &= \xi^5 + 5\xi^4 h + 10\xi^3 h^2 + 10\xi^2 h^3 + 5\xi h^4 + h^5 \\
(\xi + h)^3 &= \xi^3 + 3\xi^2 h + 3\xi h^2 + h^3 \\
f(\xi + h) - f(\xi) &= 10\xi^4 h + 20\xi^3 h^2 + 20\xi^2 h^3 + 10\xi h^4 + 2h^5 - 12\xi^2 h - 12\xi h^2 - 4h^3 + 2h \\
\frac{f(\xi_h) - f(\xi)}{h} &= 10\xi^4 + 20\xi^3 h + 20\xi^2 h^2 + 10\xi h^3 + 2h^4 - 12\xi - 12\xi h - 4h^2 + 2.
\end{aligned}$$

so

$$\lim_{h \to 0} \frac{f(\xi + h) - f(\xi)}{h} = 10\xi^4 - 12\xi^2 + 2.$$

A more important example is the following.

ex:nine3 **Example 9.3.** *Let $\exp(x)$ be the real valued function of the real variable x defined by (6.9). Then $\exp(x)$ is differentiable at every point ξ and the derivative at ξ is $\exp(\xi)$.*

Proof. We are concerned with the behaviour of

$$\frac{\exp(x + h) - \exp(x)}{h} - \exp(x)$$

as $h \to 0$. By the addition formula for exp this is

$$\left(\frac{\exp(h) - 1}{h} - 1\right) \exp(x)$$

so is suffices to show that

$$\frac{\exp(h) - 1}{h} - 1 \to 0 \text{ as } h \to 0.$$

By the definition of exp, (6.9) the right hand side is

$$\sum_{n=2}^{\infty} \frac{h^{n-1}}{n!}$$

and so when $|h| \leq 1$

$$\left|\frac{\exp(h) - 1}{h} - 1\right| \leq |h| \sum_{n=2}^{\infty} \frac{1}{n!} \leq |h| e.$$

Letting $h \to 0$ gives the desired conclusion. □

An immediate consequence of the definition is the following theorem.

thm:nine0 **Theorem 9.1.** *Suppose that $a < \xi < b$, $f : (a, b) \mapsto \mathbb{R}$ and f is differentiable at ξ. Then f is continuous at ξ.*

Proof. Let $\varepsilon > 0$. Choose δ_1 so that when $0 < |x - \xi| < \delta_1$ we have

$$\left| \frac{f(x) - f(\xi)}{x - \xi} - f'(\xi) \right| < \frac{\varepsilon}{2}.$$

Let

$$\delta = \min \left\{ \delta_1, 1, \frac{\varepsilon}{2 + 2|f'(\xi)|} \right\}$$

Then, whenever $0 < |x - \xi| < \delta$, we have, by the triangle inequality,

$$\begin{aligned}
|f(x) - f(\xi)| &\le |(x - \xi) f'(\xi)| + |x - \xi| \frac{\varepsilon}{2} \\
&< \frac{\varepsilon |f'(\xi)}{2 + 2|f'(\xi)|} + \frac{\varepsilon}{2} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon.
\end{aligned}$$

□

The next example shows that continuity at a point does not necessarily confer differentiability.

ex:nine4 **Example 9.4.** *Let $f : \mathbb{R} \mapsto \mathbb{R} : f(x) = |x|$. We have already seen in Exercise 8.1.5 that f is continuous at every point ξ. However when $h < 0$ we have $f(h)/h = -1$ and when $h > 0$ we have $f(h)/h = +1$. Hence*

$$\lim_{h \to 0-} \frac{f(0+h) - f(0)}{h} = -1 \neq 1 = \lim_{h \to 0+} \frac{f(0+h) - f(0)}{h}$$

and so

$$\lim_{h \to 0} \frac{f(0+h) - f(0)}{h}$$

does not exist.

Here are the well known rules for combining derivatives

thm:nine1 **Theorem 9.2** (Combination Theorem for Derivatives). *Suppose that $a < b$, $f : (a, b) \mapsto \mathbb{R}$, $g : (a, b) \mapsto \mathbb{R}$ and $\xi \in (a, b)$. Suppose further that f and g are differentiable at ξ and that λ and μ are real numbers. Then*

(i) When $x \in (a, b)$ let $h(x) = \lambda f(x) + \mu g(x)$. Then h is differentiable at ξ and

$$h'(\xi) = \lambda f'(\xi) + \mu g'(\xi).$$

(ii) When $x \in (a, b)$ let $j(x) = f(x) g(x)$. Then j is differentiable at ξ and

$$j'(\xi) = f'(\xi) g(\xi) + f(\xi) g'(\xi).$$

(iii) When $x \in (a,b)$ and $g(x) \neq 0$ let $k(x) = \frac{f(x)}{g(x)}$. Then k is differentiable at ξ and

$$k'(\xi) = \frac{f'(\xi)g(\xi) - f(\xi)g'(\xi)}{g^2(\xi)}.$$

Proof. These are largely easy consequences of the combination theorem for limits of functions. (i) is immediate from the observation that

$$\frac{h(x)-h(\xi)}{x-\xi} = \lambda\frac{f(x)-f(\xi)}{x-\xi} + \mu\frac{g(x)-g(\xi)}{x-\xi}.$$

For (ii) observe that

$$\frac{j(x)-j(\xi)}{x-\xi} = \frac{f(x)-f(\xi)}{x-\xi}(g(x)-g(\xi)) + \frac{f(x)-f(\xi)}{x-\xi}g(\xi) + f(\xi)\frac{g(x)-g(\xi)}{x-\xi}. \quad (9.1)$$ eq:nine1

By the previous theorem, Theorem 9.1, we have

$$\lim_{x\to\xi}(g(x)-g(\xi)) = 0.$$

Hence the right hand side of (9.1) has the required limit.

For (iii) we use

$$\frac{\frac{f(x)}{g(x)} - \frac{f(\xi)}{g(\xi)}}{x-\xi} = \left(\frac{f(x)-f(\xi)}{x-\xi}g(\xi) - f(\xi)\frac{g(x)-g(\xi)}{x-\xi}\right)\frac{1}{g(x)g(\xi)}.$$

□

The chain rule is a little trickier.

thm:nine2 **Theorem 9.3.** *Suppose that $a < b$, $g : (a,b) \mapsto \mathbb{R}$, $A < B$, $f : (A,B)) \mapsto \mathbb{R}$, $\xi \in (a,b)$ and $g((a,b)) \subset (A,B)$. Suppose further that g is differentiable at ξ and f is differentiable at $g(\xi)$. Then $h : (a,b) \mapsto \mathbb{R} : h(x) = f(g(x))$ is differentiable at ξ and*

$$h'(\xi) = f'(g(\xi))g'(\xi).$$

Proof. We want to use the identity

$$\frac{h(x)-h(\xi)}{x-\xi} = \frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{g(x)-g(\xi)}\frac{g(x)-g(\xi)}{x-\xi}$$

but we run in to the problem that we might have values of x arbitrarily close to ξ for which $g(x) = g(\xi)$ and we would be dividing by 0.

Thus we consider two cases.

Case 1. There is a $\delta_0 > 0$ such that whenever $0 < |x-\xi| < \delta_0$ we have $g(x) \neq g(\xi)$

Case 2. Suppose that for every $\delta_0 > 0$ there are x with $0 < |x - \xi| < \delta_0$ and $g(x) = g(\xi)$.

Proof in Case 1. Let $\varepsilon > 0$ and choose $\delta_1 > 0$ so that if $0 < |y - g(\xi)| < \delta_1$, then we have

$$\left|\frac{f(y) - f\big(g(\xi)\big)}{y - g(\xi)} - f'(g(\xi))\right| < \min\left(1, \frac{\varepsilon}{3 + 3|g'(\xi)|}\right). \tag{9.2}$$

eq:nine2

Then choose $\delta_2 > 0$ with $\delta_2 \le \delta_0$ so that when $0 < |x - \xi| < \delta_2$ we have $|g(x) - g(\xi)| < \delta_1$ (which is assured by the continuity of g at ξ) and

$$\left|\frac{g(x) - g(\xi)}{x - \xi} - g'(\xi)\right| < \min\left(1, \frac{\varepsilon}{3 + 3\big|f'\big(g(\xi)\big)\big|}\right). \tag{9.3}$$

eq:nine3

Then

$$\begin{aligned}
&\frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{x - \xi} - f'\big(g(\xi)\big)g'(\xi) = \\
&\qquad\left(\frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{g(x) - g(\xi)} - f'\big(g(\xi)\big)\right)\left(\frac{g(x) - g(\xi)}{x - \xi} - g'(\xi)\right) \\
&\qquad\qquad + \left(\frac{g(x) - g(\xi)}{x - \xi} - g'(\xi)\right) f'\big(g(\xi)\big) \\
&\qquad\qquad\qquad + \left(\frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{g(x) - g(\xi)} - f'\big(g(\xi)\big)\right) g'(\xi).
\end{aligned}$$

Inserting the bounds from (9.2) and (9.3) gives

$$\left|\frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{x - \xi} - f'\big(g(\xi)\big)g'(\xi)\right| \le \frac{\varepsilon}{3} + \frac{\big|f'\big(g(\xi)\big)\big|}{3 + 3\big|f'\big(g(\xi)\big)\big|} + \frac{\varepsilon}{3} < \varepsilon.$$

Proof in Case 2. In this case, given $\varepsilon > 0$ choose δ_1 so that whenever $0 < |x - \xi| < \delta_1$ we have

$$\left|\frac{g(x) - g(\xi)}{x - \xi} - g'(\xi)\right| < \varepsilon.$$

Take $\delta_0 = \delta_1$. Then there are x with $0 < |x - \xi| < \delta_1$ so that $g(x) = g(\xi)$ and hence

$$|g'(\xi)| < \varepsilon.$$

Since this holds for every $\varepsilon > 0$ it follows that $g'(\xi) = 0$.

For those x for which $g(x) \ne g(\xi)$ we can proceed as in Case 1. It then remains to consider what happens when x is such that $g(x) = g(\xi)$. Since $g(x) = g(\xi)$ and $g'(\xi) = 0$ we have

$$\frac{f\big(g(x)\big) - f\big(g(\xi)\big)}{x - \xi} - f'\big(g(\xi)\big)g'(\xi) = 0$$

so the bound
$$\left|\frac{f\big(g(x)\big)-f\big(g(\xi)\big)}{x-\xi}-f'\big(g(\xi)\big)g'(\xi)\right|<\varepsilon$$
holds anyway. □

It is often convenient, even for quite arbitrary functions, to restrict ones attention to subintervals in their domain on which the function is strictly monotonic. Then relative to that interval the function will have an inverse function which, hopefully, is well behaved. Thus we can hope to appeal to the next theorem.

thm:nine3 **Theorem 9.4** (Derivatives of Inverse Functions)**.** *Suppose that $a<b$ and $f:[a,b]\to\mathbb{R}$, that f is continuous and strictly monotonic on $[a,b]$ and differentiable on (a,b). Then $f([a,b])$ is an interval $[c,d]$, f^{-1} exists and is continuous on $[c,d]$, and is differentiable on (c,d). Moreover, for $\eta\in(c,d)$ we have*
$$(f^{-1})'(\eta)=\frac{1}{f'\big(f^{-1}(\eta)\big)}.$$

Proof. Since f is strictly monotonic on $[a,b]$ we have $f'(x)\neq 0$ for every $x\in(a,b)$. We establish the theorem when f is strictly increasing. In the strictly decreasing case we could then replace f by $-f$.

Thus $c=f(a)$ and $d=f(b)$. Now suppose
$$c\le\lambda\le d.$$
Then by the continuity of f we can invoke the intermediate value theorem, Theorem 8.7 which tells us that there is an $x\in[a,b]$ so that $f(x)=\lambda$. Thus $f([a,b])=[c,d]$ and so f^{-1} does indeed exist on $[c,d]$. Let $\eta\in(c,d)$ and put $\xi=f^{-1}(\eta)$. Since $c<\eta<d$ we have $a<\xi<b$ so $f'(\xi)$ exists. Moreover as f is strictly increasing we have $f'(\xi)\neq 0$. Hence, by the definition of a derivative and the combination theorem for limits we have
$$\lim_{x\to\xi}\frac{x-\xi}{f(x)-f(\xi)}=\frac{1}{f'(\xi)}.$$
Let $\varepsilon>0$. Then we may choose $\delta_0>0$ so that whenever $0<|x-\xi|<\delta$ we have
$$\left|\frac{x-\xi}{f(x)-f(\xi)}-\frac{1}{f'(\xi)}\right|<\varepsilon.$$
Let $\delta_-=\eta-f(\xi-\delta_0)$, $\delta_+=f(\xi+\delta_0)-\eta$ and take $\delta=\min\{\delta_-,\delta_+\}$. Then, whenever $0<|y-\eta|<\delta$, we have
$$f(\xi-\delta_0)\le\eta-\delta<y<\eta+\delta<f(\xi+\delta_0)$$

and so there is an x with $\xi - \delta_0 < x < \xi + \delta_0$ so that $y = f(x)$. Then

$$\left| \frac{f^{-1}(y) - f^{-1}(\eta)}{y - \eta} - \frac{1}{f'\big(f^{-1}(\eta)\big)} \right| = \left| \frac{x - \xi}{f(x) - f(\xi)} - \frac{1}{f'(\xi)} \right| < \varepsilon.$$

The continuity from the left at c and the right at d is easier. Let $\varepsilon > 0$. Let $\delta = f\big(\min\{a+\varepsilon, b\}\big) - f(a)$. Then whenever $c < y < c + \delta$ we have

$$a = f^{-1}(c) < f^{-1}(y) < f^{-1}(c+\delta) = f^{-1}(f(a)+\delta) \le a + \varepsilon = f^{-1}(c) + \varepsilon$$

which deals with c. d is similar. □

ex:nine5 **Example 9.5.** *We advert to the function* log *introduced in Definition 7.5 as the inverse function of* exp. *We showed in Example 9.3 that* $\exp(x)$ *is differentiable and has derivative* $\exp(x)$. *Hence, by Theorem 9.4, when* $\eta > 0$

$$\log'(\eta) = \frac{1}{\exp\big(\log(\eta)\big)} = \frac{1}{\eta}.$$

9.1.1 Exercises

subsec:nine1

1. Let $f : \mathbb{R}^+ \mapsto \mathbb{R}^+ : x \mapsto x^2$. Show that f has an inverse function and that for $y > 0$

$$(f^{-1})'(y) = \tfrac{1}{2} y^{-1/2}.$$

9.2 Extrema and Mean Value Theorems

sec:ten1

We can now say more about maxima and minima.

def:ten1 **Definition 9.2.** *1. Suppose that* $a < b$ *and* $f : (a,b) \mapsto \mathbb{R}$. *If there is a* $\xi \in (a,b)$ *and a* $\delta > 0$ *so that whenever* $x \in (a,b)(\xi - \delta, \xi + \delta)$ *we have* $f(x) \le f(\xi)$, *then we say that* f *has a* **local maximum** *at* ξ. *2. Likewise if there is a* $\xi \in (a,b)$ *and a* $\delta > 0$ *so that whenever* $x \in (a,b) \cap (\xi - \delta, \xi + \delta)$ *we have* $f(x) \ge f(\xi)$, *then we say that* f *has a* **local minimum** *at* ξ.

thm:ten1 **Theorem 9.5.** *Suppose that* $a < b$, $f : (a,b) \mapsto \mathbb{R}$, f *is differentiable on* (a,b) *and there is a* $\xi \in (a,b)$ *such that* f *has a local maximum or minimum at* ξ. *Then*

$$f'(\xi) = 0.$$

Proof. We treat the case of a local maximum. The case of a local minimum follows on replacing f by $-f$.

Let δ be as in the definition. Thus

$$f(x) - f(\xi) \le 0$$

whenever $|x-\xi|<\delta$. Therefore, when $\xi-\delta<x<\xi$ we have

$$\frac{f(x)-f(\xi)}{x-\xi} \geq 0$$

and so

$$\lim_{x\to 0-} \frac{f(x)-f(\xi)}{x-\xi} \geq 0.$$

But when $\xi<x<\xi+\delta$ we have

$$\frac{f(x)-f(\xi)}{x-\xi} \leq 0$$

and so

$$\lim_{x\to 0+} \frac{f(x)-f(\xi)}{x-\xi} \leq 0.$$

But since the limit exists the only possible common value is 0. □

It is fundamental to differentiable functions that we can use the derivative to estimate how a function varies near a particular point. As a first step we have

thm:ten2 **Theorem 9.6** (Rolle's Theorem)**.** *Suppose that f is continuous on $[a,b]$, differentiable on (a,b) and $f(a)=f(b)$. Then there is a $\xi\in(a,b)$ so that*

$$f'(\xi)=0.$$

Proof. If $f(x)=f(a)$ for every $x\in[a,b]$, then at once $f'(\xi)=0$ for every $\xi\in(a,b)$. Thus we can suppose that there is an $x\in(a,b)$ so that $f(x)\neq f(a)$. As $f(x)$ is continuous it attains its extrema and so has a maximum or minimum at some $\xi\in(a,b)$. Hence, by Theorem 9.5, $f'(\xi)=0$. □

thm:ten3 **Theorem 9.7** (The Mean Value Theorem)**.** *Suppose that f is continuous on $[a,b]$ and differentiable on (a,b). Then there is a $\xi\in(a,b)$ so that*

$$f(b)-f(a)=(b-a)f'(\xi).$$

Proof. Let

$$g(x)=(b-a)(f(x)-f(a))-(x-a)(f(b)-f(a)).$$

Then

$$g(a)=g(b)=0.$$

Hence, by Theorem 9.6, there is a $\xi\in(a,b)$ so that $g'(\xi)=0$. Moreover

$$g'(x)=(b-a)f'(x)-(f(b)-f(a)).$$

□

There is a more elaborate version of this

thm:ten4 **Theorem 9.8** (Cauchy's Mean Value Theorem)**.** *Suppose that $a < b$ and f and g are continuous on $[a, b]$, differentiable on (a, b) and $g'(x) \neq 0$ for every $x \in (a, b)$. Then there is a $\xi \in (a, b)$ so that*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. By Theorem 9.6, $g(b) \neq g(a)$ since otherwise there would be a $x \in (a, b)$ such that $g'(x) = 0$. Let

$$h(x) = (g(x) - g(a))(f(b) - f(a)) - (g(b) - g(a))(f(x) - f(a)).$$

Then $h(a) = h(b) = 0$. Thus by Theorem 9.6 there is $\xi \in (a, b)$ such that $h'(\xi) = 0$. But

$$h'(x) = g'(x)(f(b) - f(a)) - (g(b) - g(a))f'(x).$$

□

Now we can establish every student's favourite theorem.

thm:ten5 **Theorem 9.9** (l'Hôpital's Rule)**.** *Suppose that $a < b$ and f and g are continuous on [a,b], differentiable on (a, b), $f(a) = g(a) = 0$, $g'(x) \neq 0$ for $x \in (a, b)$ and*

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)}$$

exists and $= \ell$. Then

$$\lim_{x \to a+} \frac{f(x)}{g(x)}$$

exists and $= \ell$.

There is a corresponding theorem for limits from below, and thus two sided limits.

Proof. Let $\varepsilon > 0$. Choose $\delta > 0$ so that whenever $a < x < a + \delta$ we have

$$\left| \frac{f'(x)}{g'(x)} - \ell \right| < \varepsilon.$$

Then, whenever $y \in (a, a + \delta)$ we have, by Theorem 9.8,

$$\frac{f(y)}{g(y)} = \frac{f'(\xi)}{g'(\xi)}$$

for some ξ with $a < \xi < y < a + \delta$. Thus

$$\left| \frac{f(y)}{g(y)} - \ell \right| = \left| \frac{f'(\xi)}{g'(\xi)} - \ell \right| < \varepsilon.$$

□

There is a generalization of the Mean Value Theorem which can be used to obtain power series expansions of interesting functions.

thm:ten8 **Theorem 9.10** (Taylor's Theorem with a Remainder)**.** *Suppose that $a < \xi < b$ and $f : (a,b) \mapsto \mathbb{R}$ is n times differentiable on (a,b). Suppose also that x, $\xi \in (a,b)$ and that p is a real number with $0 < p \leq n$. Then there is an η between ξ and x such that*

$$f(x) = \sum_{k=0}^{n-1} \frac{(x-\xi)^k}{k!} f^{(k)}(\xi) + \frac{(x-\xi)^p (x-\eta)^{n-p}}{(n-1)!p} f^{(n)}(\eta).$$

The case $p = 1$ gives Cauchy's form of the Remainder and the case $p = n$ gives Lagrange's form of the remainder. Note that different p likely give different η.

Proof. Consider for $y \in (a,b)$

$$\psi(y) = (x-\xi)^p \left(-f(x) + \sum_{k=0}^{n-1} \frac{(x-y)^k}{k!} f^{(k)}(y) \right) + (x-y)^p \left(f(x) - \sum_{k=0}^{n-1} \frac{(x-\xi)^k}{k!} f^{(k)}(\xi) \right).$$

Then $\psi(x) = \psi(\xi) = 0$ and ψ has a derivative on the closed interval with ξ and x as endpoints. Hence, by Theorem 9.6, Rolle's theorem, there is an η between ξ and x so that $\psi'(\eta) = 0$, and

$$\psi'(y) = -(x-\xi)^p \frac{(x-y)^{n-1}}{(n-1)!} f^{(n)}(y) + p(x-y)^{p-1} \left(f(x) - \sum_{k=0}^{n-1} \frac{(x-\xi)^k}{k!} f^{(k)}(\xi) \right).$$

□

9.2.1 Exercises

sec:ten2

1. **The Newton-Raphson method.** Suppose that $a < b$ and $f : (a,b) \mapsto \mathbb{R}$ is twice differentiable on (a,b) and f'' is continuous on $[a,b]$. Suppose further that $f'(x) \neq 0$ for $x \in (a,b)$ and that there is an $x_0 \in (a,b)$ so that $f(x_0) = 0$.

(i) Given $\xi \in (a,b)$ define

$$x = \xi - \frac{f(\xi)}{f'(\xi)}$$

and suppose that $x \in (a,b)$. Prove that there is an η between x and ξ so that

$$x - x_0 = \frac{f''(\eta)}{2f'(\xi)} (x_0 - \xi)^2.$$

(ii) Let

$$\lambda = \frac{b-a}{2} \left(\sup\{|f''(x)| : x \in [a,b]\}\right) \left(\sup\{|f'(x)|^{-1} : x \in [a,b]\}\right)$$

and suppose that $\lambda \le 1$. Further define the sequence $\langle x_n \rangle$ by choosing $x_1 \in (a, b)$ and then defining iteratively

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Prove that

$$|x_{n+1} - x_0| \le \lambda |x_n - x_0|^2 \le \lambda^n |x_1 - x_0|^{2^n}.$$

Thus a good initial guess x_1 for x_0 leads to fantastically fast convergence.

(iii) Prove that if $f(x) = x^2 - 2$ and $a = 1$, $b = 3$, and $x_1 = 2$, then we have

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right),$$

the sequence $\langle x_n \rangle$ introduced in Example 4.2 and analysed in Example 5.2. Crucially there we managed to arrange that $|x_1 - x_0| = 2 - \sqrt{2} < 1$.

2. Find an approximation to the real root of $x^3 - 3x + 3$ to eight decimal places.

9.3 Derivatives of Power Series

We have already seen that there are interesting functions which can be defined as power series, and generally power series are really well behaved. Thus it will be no surprise that they can be differentiated.

thm:ten6 **Theorem 9.11.** *Let $\langle a_n \rangle$ be a sequence of real numbers and suppose that the corresponding power series*

$$\sum_{n=0}^{\infty} a_n x^n \tag{9.4}$$ eq:ten1

has positive radius of convergence R. When $|x| < R$ let $A(x)$ denote the sum of the series. Then $A'(x)$ exists,

$$A'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{m=0}^{\infty} (m+1) a_{m+1} x^m \tag{9.5}$$ eq:ten2

and this series also has radius of convergence R.

In other words we can obtain the derivative by term-by-term differentiation of the series. That is, by interchanging the limiting operations. The following example shows that in other situations this is not always possible.

Example 9.6. *We have*

$$\lim_{x \to 1-} \lim_{n \to \infty} x^n = 0$$

but

$$\lim_{n \to \infty} \lim_{x \to 1-} x^n = 1.$$

Proof of Theorem 9.11. When $a_n = 0$ for every $n \geq 2$ the proof is easy, so we may assume that for some $n \geq 2$ we have $a_n \neq 0$. We have already seen in Exercise 6.5.1.2 that the series in (9.5) and the series

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}$$

have radius of convergence identical to that in (9.4). Also it is immediate from Theorem 6.12 that all three series converge *absolutely* for $|x| < R$.

The proof is now an elaboration of that of Example 9.3. Suppose that $|x| < R$. Let $\varepsilon > 0$. Choose δ_0 so that $0 < \delta_0 < \frac{1}{2}(R - |x|)$ and then choose δ so that $\delta < \delta_0/2$ and

$$0 < \delta < \varepsilon \left(\sum_{n=2}^{\infty} n(n-1)|a_n|(R-\delta_0)^{n-2} \right)^{-1}.$$

Note that our assumption that $a_n \neq 0$ for some $n \geq 2$ ensures that the series here is positive. Suppose that $0 < |h| < \delta$. Then

$$\frac{A(x+h) - A(x)}{h} - \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=1}^{\infty} a_n \left(\frac{(x+h)^n - x^n}{h} - n x^{n-1} \right).$$

Then it is easily checked that

$$\begin{aligned} \frac{(x+h)^n - x^n}{h} - n x^{n-1} &= \sum_{m=0}^{n-1} \big((x+h)^m - x^m \big) x^{n-1-m} \\ &= h \sum_{m=0}^{n-1} \sum_{k=0}^{m-1} (x+h)^k x^{n-2-k}. \end{aligned}$$

Just apply twice the formula for a sum of the terms of a geometric progression. Thus

$$\left| \frac{(x+h)^n - x^n}{h} - n x^{n-1} \right| \leq |h| n(n-1)(R-\delta_0)^{n-2}$$

and so

$$\left| \frac{A(x+h) - A(x)}{h} - \sum_{n=1}^{\infty} n a_n x^{n-1} \right| < |h| \sum_{n=2}^{\infty} n(n-1)|a_n|(R-\delta_0)^{n-2} < \varepsilon.$$

□

thm:ten7 **Theorem 9.12** (The identity theorem for power series)**.** *Suppose $\langle a_n \rangle$ and $\langle a_n \rangle$ are real sequences, that $R > 0$, and that*

$$A(x) = \sum_{n=0}^{\infty} a_n x^n \qquad (9.6)$$ eq:ten3

and

$$B(x) = \sum_{n=0}^{\infty} b_n x^n \tag{9.7}$$

eq:ten4

are both convergent and satisfy $A(x) = B(x)$ for $|x| < R$. Then $a_n = b_n$ for every non-negative integer n.

Proof. It suffices to suppose that $A(x) = 0$ for $|x| < R$. Then $a_0 = A(0) = 0$. Now suppose that $a_0 = \cdots = a_{m-1} = 0$ for some $m \in \mathbb{N}$. Then

$$\sum_{n=m}^{\infty} a_n x^n = 0 \quad (|x| < R)$$

and so

$$A_m(x) = \sum_{n=m}^{\infty} a_n x^{n-m} = 0 \quad (0 < |x| < R).$$

This is also a power series, and so has radius of convergence $\geq R$. Thus by continuity

$$a_m = A_m(0) = 0.$$

□

9.3.1 Exercises

subsec:ten1

1. Prove that the functions $\sin(x)$, $\cos(x)$, $\sinh(x)$, $\cosh(x)$, defined by (6.10), (6.11) and Exercise 6.5.1.3 satisfy for every $\xi \in \mathbb{R}$

$$\sin'(\xi) = \cos(\xi),\ \cos'(\xi) = -\sin(\xi),\ \sinh'(\xi) = \cosh(\xi),\ \cosh'(\xi) = \sinh(\xi).$$

2. Prove that $\sin(x)$ has an inverse function on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and that

$$(\sin^{-1})'(y) = \frac{1}{\sqrt{1-y^2}}.$$

3. Find inverse functions for tan, cosh and sinh on suitable domains, and their corresponding derivatives.

9.4 Notes

sec:ten8

For an overview of the many possible generalizations of the derivative see `https://en.wikipedia.org/wiki/Derivative`

Chapter 10

The Riemann Integral

ch:eleven

10.1 Upper and Lower Sums

sec:eleven1

An integral is a means of measuring certain mathematical objects. Its construction is motivated by the need to measure the area under a curve, but it can also be considered as a kind of average value of a function. Since many functions we come across are not continuous at every point of their domain we want to include as wide a range of functions as possible. A laudable attempt in this direction is the Riemann integral. There are more sophisticated integrals, such as the Lebesgue integral, but the Riemann integral is adequate for many purposes. The motivation for its construction is the idea that a good approximation to the area under a curve is a collection of rectangles on a narrow interval base whose height approximates the function on that interval. We work on an interval $[a, b]$ and we partition this into smaller intervals.

def:eleven1

Definition 10.1 (Partition of an Interval)**.** *Suppose that $a \leq b$. A* **partition** Δ *of the interval $[a, b]$ is a finite increasing sequence of distinct points in $[a, b]$ including a and b. We let $n + 1 = \operatorname{card} \Delta$ and order the sequence as $x_0, x_1, \ldots x_n$ with*

$$a = x_0 \leq x_1 \leq \ldots \leq x_{n-1} \leq x_n = b.$$

We call n the **length** *of the partition. Note that Δ dissects $[a, b]$ into n subintervals. We then define $\mathcal{D}[a, b]$ to be the set of all such Δ of any length.*

ex:eleven1

Example 10.1. *Each of the finite sequences*

$$x_j = a + \frac{b-a}{n} j \quad (0 \leq j \leq n),$$

where $n \in \mathbb{N}$, are in $\mathcal{D}[a, b]$.

The next definition attempts to approximate the area under a given curve $y = f(x)$.

def:eleven2 **Definition 10.2** (Upper and Lower Sums)**.** *Suppose that $a \le b$ and $f : [a,b] \mapsto \mathbb{R}$ is bounded. Given an arbitrary partition $\Delta \in \mathcal{D}[a,b]$ we define upper and lower sums $\overline{S}$ and $\underline{S}$ by*

$$\underline{S}(f,\Delta) = \sum_{j=1}^{n}(x_j - x_{j-1})\inf\{f(x) : x \in [x_{j-1}, x_j]\},$$
$$\overline{S}(f,\Delta) = \sum_{j=1}^{n}(x_j - x_{j-1})\sup\{f(x) : x \in [x_{j-1}, x_j]\}.$$

These sums ought to minorise and majorise the area under the curve $y = f(x)$. The next definition attempts to squeeze as closely as possible to this area (if it exists!).

def:eleven3 **Definition 10.3** (Upper and Lower Integrals)**.** *Suppose that $f : [a,b] \mapsto \mathbb{R}$ is bounded. Then we define upper and lower integrals by*

$$\underline{\int_a^b} f(x)dx = \sup\{\underline{S}(f,\Delta) : \Delta \in \mathcal{D}[a,b]\},$$
$$\overline{\int_a^b} f(x)dx = \inf\{\overline{S}(f,\Delta) : \Delta \in \mathcal{D}[a,b]\}.$$

def:eleven4 **Definition 10.4** (The Riemann Integral)**.** *When $f : [a,b] \mapsto \mathbb{R}$ is bounded and f is such that*

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx,$$

then we say that f is **Riemann Integrable** *on $[a,b]$ and write*

$$\int_a^b f(x)dx$$

for the common value. It is convenient to take $\mathcal{R}[a,b]$ to be the set of Riemann integrable functions on $[a,b]$. When $b < a$ and $f \in \mathcal{R}[b,a]$ we extend the definition by taking

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

ex:eleven0 **Example 10.2.** *Suppose that $f(x)$ is a constant, that is $a \le b$, $f : [a,b] \mapsto \mathbb{R} : f(x) = c$. Then $f \in \mathcal{R}[a,b]$ and*

$$\int_a^b f(x)dx = (b-a)c.$$

Proof. Of course, for any partition of $[a,b]$ we have $\inf\{f(x) : x \in [x_{j-1}, x_j]\} = c$ and $\sup\{f(x) : x \in [x_{j-1}, x_j]\} = c$. Hence

$$\underline{S}(f,\Delta) = \sum_{j=1}^{n}(x_j - x_{j-1})c = (b-a)c.$$

and likewise $\overline{S}(f,\Delta) = (b-a)c$. $\square$

ex:example0+ **Example 10.3.** *Suppose that $a < b$ and define $f : [a, b] \mapsto \mathbb{R}$ by*

$$f(x) = \begin{cases} 0 & x \in \mathbb{R} \setminus \mathbb{Q}, \\ 1 & x \in \mathbb{Q}. \end{cases}$$

Then

$$\underline{\int_a^b} f(x)dx = 0$$

and

$$\overline{\int_a^b} f(x)dx = 1$$

so f is **not** *Riemann integrable.*

The point is that every interval $[x_{j-1}, x_j]$ with $x_{j-1} < x_j$ contains both a rational number and an irrational number.

There are various useful relationships that we need to establish between upper and lower sums and upper and lower integrals. In particular we will need to combine partitions.

thm:eleven1 **Theorem 10.1.** *Suppose that $a \leq b$, $f : [a, b] \mapsto \mathbb{R}$, f is bounded and $\Delta_j \in \mathcal{D}[a, b]$ $(j = 1, 2)$. Let $\Delta^* = \Delta_1 \cup \Delta_2$. Then*

$$\Delta^* \in \mathcal{D}[a, b]. \tag{10.1}$$ eq:eleven1

Moreover

$$\underline{S}(f, \Delta_1) \leq \underline{S}(f, \Delta^*) \leq \overline{S}(f, \Delta^*) \leq \overline{S}(f, \Delta_2) \tag{10.2}$$ eq:eleven2

and for any $\Delta \in \mathcal{D}[a, b]$ we have

$$\underline{S}(f, \Delta) \leq \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \leq \overline{S}(f, \Delta). \tag{10.3}$$ eq:eleven3

thm:eleven1a **Corollary 10.2.** *Suppose that $\mathfrak{m}$ and $\mathfrak{M}$ are real numbers such that for every $x \in [a, b]$ we have*

$$\mathfrak{m} \leq f(x) \leq \mathfrak{M}.$$

Then

$$(b - a)\mathfrak{m} \leq \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \leq (b - a)\mathfrak{M}.$$

Proof of Theorem 10.1. The equation (10.1) is immediate on ordering the elements of Δ^*. Suppose that $a \leq u < v < w \leq b$. Then

$$\begin{aligned} &\inf\{f(x) : x \in [u, w]\} \\ &\qquad \leq \inf\{f(x) : x \in [u, v]\}, \inf\{f(x) : x \in [u, w]\} \\ &\qquad\qquad \leq \inf\{f(x) : x \in [v, w]\}. \end{aligned}$$

Hence

$$\begin{aligned}(w-u)\inf\{f(x):x\in[u,w]\}\\ \le (v-u)\inf\{f(x):x\in[u,v]\}+(w-v)\inf\{f(x):x\in[u,w]\},\end{aligned}$$

and likewise we have

$$\begin{aligned}(v-u)\sup\{f(x):x\in[u,v]\}+(w-v)\sup\{f(x):x\in[u,w]\}\\ \le (w-u)\sup\{f(x):x\in[u,w]\}.\end{aligned}$$

Repeated use of these inequalities establishes (10.2)

By (10.2) $\overline{S}(f,\Delta_2)$ is an upper bound for

$$\{S(f,\Delta_1):\Delta_1\in\mathcal{D}[a,b]\}$$

and hence for its supremum

$$\underline{\int_a^b} f(x)dx.$$

But then this is a lower bound for $\overline{S}(f,\Delta_2)$ for any $\Delta_2\in\mathcal{D}[a,b]$ and hence for

$$\inf\{\overline{S}(f,\Delta_2):\Delta_2\in\mathcal{D}[a,b]\}=\overline{\int_a^b} f(x)dx.$$

This establishes the middle inequality in (10.3) and the outer ones follow by the definition of upper and lower integrals. □

thm:eleven5 **Theorem 10.3.** *Suppose that $a\le b$ and $f\in\mathcal{R}[a,b]$. Then $|f|\in\mathcal{R}[a,b]$ and*

$$\left|\int_a^b f(x)dx\right|\le\int_a^b|f(x)|dx.$$

Proof. Let $\varepsilon>0$ and choose the partition Δ so that

$$\overline{S}(f,\Delta)-\underline{S}(f,\Delta)<\varepsilon.$$

Consider a particular interval $[x_{j-1},x_j]$ of the partition. Then there are three possibilities.

1. $\inf\{f(x):x\in[x_{j-1},x_j]\}\le\sup\{f(x):x\in[x_{j-1},x_j]\}\le 0$.
2. $\inf\{f(x):x\in[x_{j-1},x_j]\}\le 0<\sup\{f(x):x\in[x_{j-1},x_j]\}$.
3. $0\le\inf\{f(x):x\in[x_{j-1},x_j]\}\le\sup\{f(x):x\in[x_{j-1},x_j]\}$.

In case 1. we have

$$\inf\{|f(x)|:x\in[x_{j-1},x_j]\}=-\sup\{f(x):x\in[x_{j-1},x_j]\}$$

and

$$\sup\{|f(x)|:x\in[x_{j-1},x_j]\}=-\inf\{f(x):x\in[x_{j-1},x_j]\}$$

so that

$$\begin{aligned}\sup\{|f(x)| : x \in [x_{j-1}, x_j]\} &- \inf\{|f(x)| : x \in [x_{j-1}, x_j]\}\\ &\le \sup\{f(x) : x \in [x_{j-1}, x_j]\} - \inf\{f(x) : x \in [x_{j-1}, x_j]\} \quad (10.4)\end{aligned}$$ eq:eleven|

and this also holds in case 3.

In case 2. we have

$$\begin{aligned}\sup\{|f(x)| &: x \in [x_{j-1}, x_j]\}\\ &= \max\big\{ \sup\{f(x) : x \in [x_{j-1}, x_j]\}, -\inf\{f(x) : x \in [x_{j-1}, x_j]\}\big\}\\ &\le \sup\{f(x) : x \in [x_{j-1}, x_j]\} - \inf\{f(x) : x \in [x_{j-1}, x_j]\}\end{aligned}$$

and

$$\inf\{|f(x)| : x \in [x_{j-1}, x_j]\} \ge 0$$

so that (10.4) again holds. Therefore

$$0 \le \overline{S}(|f|, \Delta) - \underline{S}(|f|, \Delta) \le \overline{S}(f, \Delta) - \underline{S}(f, \Delta) < \varepsilon,$$

whence

$$0 \le \overline{\int_a^b} |f(x)|dx - \underline{\int_a^b} |f(x)|dx < \varepsilon$$

and this holds for every $\varepsilon > 0$.

The final inequality follows from the observations

$$\overline{S}(-f, \Delta) \le \overline{S}(|f|, \Delta) \text{ and } \overline{S}(f, \Delta) \le \overline{S}(|f|, \Delta)$$

□

We have a combination theorem for Riemann integrable functions.

thm:eleven2b **Theorem 10.4** (Combination Theorem for Integrals)**.** *Suppose that $a \le b$, $f, g \in \mathcal{R}[a, b]$ and $\lambda, \mu \in \mathbb{R}$. Then $\lambda f + \mu g \in \mathcal{R}[a, b]$ and*

$$\int_a^b \big(\lambda f(x) + \mu g(x)\big)dx = \lambda \int_a^b f(x)dx + \mu \int_a^b g(x)dx.$$

Note that there is no formula for the integral of fg in terms of the integrals of f and g.

Proof. If $\lambda \ge 0$ we have

$$\inf\{\lambda f(x) : x \in [x_{j-1}, x_j]\} = \lambda \inf\{f(x) : x \in [x_{j-1}, x_j]\}$$

and likewise for inf replaced by sup. Hence $\lambda f(x) \in \mathcal{R}[a,b]$ and

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx.$$

When $\lambda < 0$ we have

$$\inf\{\lambda f(x) : x \in [x_{j-1}, x_j]\} = \lambda \sup\{f(x) : x \in [x_{j-1}, x_j]\}$$

and

$$\sup\{\lambda f(x) : x \in [x_{j-1}, x_j]\} = \lambda \inf\{f(x) : x \in [x_{j-1}, x_j]\}.$$

Moreover, because multiplication by negatives flips inequalities we get

$$\underline{\int_a^b} \lambda f(x)dx = \lambda \overline{\int_a^b} f(x)dx = \lambda \int_a^b f(x)dx$$

and

$$\overline{\int_a^b} \lambda f(x)dx = \lambda \underline{\int_a^b} f(x)dx = \lambda \int_a^b f(x)dx.$$

Thus we can now assume that $\lambda = \mu = 1$.

Let $\varepsilon > 0$ and choose partitions Δ_1 and Δ_2 of $[a,b]$ so that

$$\int_a^b f(x)dx - \frac{\varepsilon}{2} < \underline{S}(f, \Delta_1)$$

and

$$\int_a^b g(x)dx - \frac{\varepsilon}{2} < \underline{S}(g, \Delta_2).$$

As in Theorem 10.1 let $\Delta^* = \Delta_1 \cup \Delta_2$. Then

$$\underline{S}(f, \Delta_1) \le \underline{S}(f, \Delta^*),\ \underline{S}(g, \Delta_2) \le \underline{S}(g, \Delta^*)$$

and

$$\underline{S}(f, \Delta^*) + \underline{S}(g, \Delta^*) \le \underline{S}(f+g, \Delta^*).$$

Therefore

$$\int_a^b f(x)dx + \int_a^b g(x)dx - \varepsilon < \underline{\int_a^b} \big(f(x) + g(x)\big)dx.$$

Likewise

$$\overline{\int_a^b} \big(f(x) + g(x)\big)dx < \int_a^b f(x)dx + \int_a^b g(x)dx + \varepsilon.$$

Since this holds for every $\varepsilon > 0$, by the middle inequality in (10.3) with f replaced by $f + g$, we have

$$\underline{\int_a^b} \big(f(x) + g(x)\big)dx = \overline{\int_a^b} \big(f(x) + g(x)\big)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

$\square$

ex:eleven0a **Example 10.4.** *Suppose that $a \le b$, $f, g \in \mathcal{R}[a,b]$ and $f(x) \le g(x)$ for every $x \in [a,b]$. Then*

$$\int_a^b f(x)dx \le \int_a^b g(x)dx.$$

Proof. By the combination theorem $g - f \in \mathcal{R}[a,b]$. Moreover, $g(x) - f(x) \ge 0$ for every $x \in [a,b]$. Hence, by Corollary 10.2

$$0 \le \int_a^b \big(g(x) - f(x)\big)dx.$$

The result then follows by the combination theorem. □

thm:eleven2c **Theorem 10.5.** *Suppose that $f \in \mathcal{R}[a,b]$, and $g : [a,b] \mapsto \mathbb{R}$ differs from f at only a finite number of values of x. Then $g \in \mathcal{R}[a,b]$ and*

$$\int_a^b g(x)dx = \int_a^b f(x)dx.$$

Proof. It suffices to prove the theorem when f and g only differ at one place, for then we can apply the theorem a finite number of times. Suppose the difference occurs at $x = \xi$, so that for some real number d we have

$$g(x) = \begin{cases} f(x) & (x \ne \xi), \\ f(x) + d & (x = \xi). \end{cases}$$

Let $\varepsilon > 0$ and let Δ be the partition of $[a,b]$ given by

$$a, \xi - \varepsilon, \xi + \varepsilon, b$$

with an obvious adjustment when ξ is within a distance ε of a or b. Then

$$-2|d|\varepsilon \le \underline{S}(g - f, \Delta) \le \overline{S}(g - f, \Delta) < 2\varepsilon|d|.$$

and so

$$-2|d|\varepsilon \le \underline{\int_a^b}\big(g(x) - f(x)\big)dx \le \overline{\int_a^b}\big(g(x) - f(x)\big)dx \le< 2\varepsilon|d|.$$

Since this holds for every $\varepsilon > 0$ we have $g - f \in \mathcal{R}[a,b]$ and

$$\int_a^b \big(g(x) - f(x)\big)dx = 0.$$

The result then follows from the combination theorem. □

thm:eleven2+ **Theorem 10.6.** *Suppose that* $a \le b < c$, $f : [a,c] \mapsto \mathbb{R}$ *and* f *is bounded. Then* $f \in \mathcal{R}[a,c]$ *if and only if* $f \in \mathcal{R}[a,b]$ *and* $f \in \mathcal{R}[b,c]$, *and in either case*

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx.$$

Proof. First suppose that $f \in \mathcal{R}[a,c]$. Let $\varepsilon > 0$. Then there is a partition Δ of $[a,c]$ so that

$$0 \le \overline{S}(f,\Delta) - \underline{S}(f,\Delta) < \varepsilon. \tag{10.5}$$ eq:eleven+

Let $\Delta^* = \Delta \cup \{a,b,c\}$. Then, by Theorem 10.1, (10.5) holds with Δ replaced by Δ^*. Let $\Delta_1 = \Delta^* \cap [a,b]$ and $\Delta_2 = \Delta^* \cap [b,c]$ so that Δ_1 and Δ_2 are partitions of $[a,b]$ and $[b,c]$ respectively. Then by (10.5) with Δ replaced by Δ^* we have

$$0 \le \overline{S}(f,\Delta_1) - \underline{S}(f,\Delta_1) + \overline{S}(f,\Delta_2) - \underline{S}(f,\Delta_2) = \overline{S}(f,\Delta^*) - \underline{S}(f,\Delta^*) < \varepsilon.$$

and hence for $j = 1, 2$

$$0 \le \overline{S}(f,\Delta_j) - \underline{S}(f,\Delta_j) < \varepsilon.$$

It follows that $f \in \mathcal{R}[a,b]$ and $\mathcal{R}[b,c]$ and that

$$-\varepsilon < \int_a^c f(x)dx - \int_a^b f(x)dx - \int_b^c f(x)dx < \varepsilon$$

and this holds for every $\varepsilon > 0$.

If we suppose on the other hand that $f \in \mathcal{R}[a,b]$ and $f \in \mathcal{R}[b,c]$, then for each $\varepsilon > 0$ there will be partitions of Δ_1 and Δ_2 of $[a,b]$ and $[b,c]$ respectively so that (10.5) holds with Δ replaced by Δ_1 and Δ_2. Let $\Delta = \Delta_1 \cup \Delta_2$. Then (10.5) will hold with ε replaced by 2ε, and it follows in a similar way to the above that $f \in \mathcal{R}[a,c]$ and the conclusion of the theorem holds once more. □

10.1.1 Exercises

susec:eleven3

1. Suppose that $a < b$ and $f : [a,b] \mapsto \mathbb{R}$ is defined by $f(a/q) = 1/q$ when $q \in \mathbb{N}$, $a \in \mathbb{Z}$, and a and q have no common factors > 1, and $f(x) = 0$ when $x \in \mathbb{R} \setminus \mathbb{Q}$. Prove that $f \in \mathcal{R}[a,b]$.

2. Suppose that $a \le b$ and $f : [a,b] \mapsto \mathbb{R}$ is monotonic on $[a,b]$. Prove that $f \in \mathcal{R}[a,b]$.

10.2 Step Functions

sec:eleven3

There are several places where we can usefully approximate a Riemann integrable function by a step function.

def:eleven5 **Definition 10.5.** *Suppose that* $a \le b$. *By a* **step function** F *on* $[a,b]$ *we mean that there is a partition* Δ
$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$
of $[a,b]$ *and a finite sequence*
$$c_0, c_1, c_2, \ldots, c_n$$
of real numbers and the F *satisfies*
$$F(x) = c_j \quad \big(x \in (x_{j-1}, x_j)\big) \quad (1 \le j \le n)$$
and
$$F(x_j) = c_j \text{ or } c_{j+1} \quad (0 \le j \le n).$$

By Theorem 10.5 and Example 10.2, for each j with $1 \le j \le n$, $f \in \mathcal{R}[x_{j-1}, x_j]$ and
$$\int_{x_{j-1}}^{x_j} F(x)dx = (x_j - x_{j-1})c_j.$$
Thus, by Theorem 10.6, $F \in \mathcal{R}[a,b]$. Moreover, if we are given $f \in \mathcal{R}[a,b]$ and take
$$c_j = \sup\{f(x) : x \in [x_{j-1}, x_j]\},$$
then
$$\int_a^b F(x)dx = \overline{S}(f, \Delta).$$
Likewise if we take
$$c_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\},$$
then
$$\int_a^b F(x)dx = \underline{S}(f, \Delta).$$

thm:eleven2z **Theorem 10.7.** *Suppose that* $a \le b$, $f, g \in \mathcal{R}[a,b]$. *Then* $fg \in \mathcal{R}[a,b]$.

Proof. We write this as
$$fg = \left(\frac{f+g}{2}\right)^2 - \left(\frac{f-g}{2}\right)^2.$$
Since $\frac{f \pm g}{2} \in \mathcal{R}[a,b]$ it suffices to deal with f^2, that is, the special case $g = f$. Moreover, $f^2 = |f|^2$ and by Theorem 10.3 we know that $|f| \in \mathcal{R}[a,b]$.

Let $\varepsilon > 0$ and choose a partition Δ so that
$$\int_a^b |f(x)|dx \le \overline{S}(|f|, \Delta) < \int_a^b |f(x)|dx + \varepsilon.$$
and define the step function F_ε by
$$F_\varepsilon(x) = \sup\{|f(x)| : x \in [x_{j-1}, x_j]\} \text{ when } x \in (x_{j-1}, x_j) \quad (1 \le j \le n)$$

and

$$F_\varepsilon(x) = |f(x_j)| \quad (0 \le j \le n).$$

Then $|f(x)| \le F_\varepsilon(x)$ for every $x \in [a,b]$, and F_ε^2 is also a step function, so by Example 10.2 and Theorem 10.6, $F_\varepsilon \in \mathcal{R}[a,b]$, $F_\varepsilon^2 \in \mathcal{R}[a,b]$ and

$$\overline{S}(|f|,\Delta) = \int_a^b F_\varepsilon(x)dx.$$

Hence, by Example 10.4,

$$0 \le \int_a^b \big(F_\varepsilon(x) - |f(x)|\big)dx < \varepsilon.$$

Let $\mathfrak{M} = \sup\{|f(x)| : x \in [a,b]\}$. Then

$$\begin{aligned} 0 &\le \overline{\int_a^b} \big(F_\varepsilon(x)^2 - f(x)^2\big)dx \\ &\le 2\mathfrak{M}\overline{\int_a^b} \big(F_\varepsilon(x) - |f(x)|\big)dx \\ &< 2\mathfrak{M}\varepsilon. \end{aligned}$$

Therefore

$$-2\mathfrak{M}\varepsilon + \underline{\int_a^b} F_\varepsilon^2(x)dx < \underline{\int_a^b} f(x)^2dx \le \overline{\int_a^b} f(x)^2dx \le \overline{\int_a^b} F_\varepsilon^2(x)dx + 2\mathfrak{M}\varepsilon.$$

But

$$\underline{\int_a^b} F_\varepsilon^2(x)dx = \overline{\int_a^b} F_\varepsilon^2(x)dx.$$

Thus

$$0 \le \overline{\int_a^b} f^2(x)dx - \underline{\int_a^b} f(x)^2dx < 4\mathfrak{M}\varepsilon$$

and this holds for every $\varepsilon > 0$. □

The following theorem is important in the theory of Fourier series and Fourier transforms.

thm:eleven2d **Theorem 10.8** (The Riemann-Lebesgue Lemma). *Suppose that $f \in \mathcal{R}[a,b]$. Then*

$$\begin{aligned} \lim_{\lambda\to\infty} \int_a^b f(x)\cos(\lambda x)dx &= 0, \\ \lim_{\lambda\to\infty} \int_a^b f(x)\sin(\lambda x)dx &= 0. \end{aligned}$$

Note that $\cos(\lambda x)$ and $\sin(\lambda x)$ are continuous functions of x so, by Theorem 10.7 the integrands above are in $\mathcal{R}[a, b]$.

Proof. Let $\varepsilon > 0$. Choose the partition Δ of $[a, b]$ so that

$$\overline{S}(f, \Delta) - \varepsilon < \int_a^b f(x)dx \leq \overline{S}(f, \Delta). \tag{10.6}$$ eq:eleven8

Let F_ε be a step function associated with Δ as above, so that

$$\int_a^b F_\varepsilon(x)dx = \overline{S}(f, \Delta). \tag{10.7}$$ eq:eleven9

Note that $f(x) \leq F_\varepsilon(x)$, and n and Δ may well depend on ε.

Now

$$\begin{aligned}\int_a^b F_\varepsilon(x)\cos(\lambda x)dx &= \sum_{j=1}^{n} \sup\{f(x) : x \in [x_{j-1}, x_j]\} \int_{x_{j-1}}^{x_j} \cos(\lambda x) \\ &= \sum_{j=1}^{n} \sup\{f(x) : x \in [x_{j-1}, x_j]\} \frac{\sin(\lambda x_j) - \sin(\lambda x_{j-1})}{\lambda}.\end{aligned}$$

Therefore

$$\left| \sum_{j=1}^{n} \sup\{f(x) : x \in [x_{j-1}, x_j]\} \int_a^b F_\varepsilon(x)\cos(\lambda x)dx \right| \leq 2n\mathfrak{M}\lambda^{-1}$$

where $\mathfrak{M} = \sup\{|f(x)| : x \in [a, b]\}$, whence if $\lambda > \Lambda(\varepsilon)$, then we have

$$\left| \int_a^b F_\varepsilon(x)\cos(\lambda x)dx \right| < \varepsilon.$$

Hence, by Theorem 10.3 and Example 10.4,

$$\begin{aligned}\left| \int_a^b f(x)\cos(\lambda x)dx \right| &\leq \left| \int_a^b F_\varepsilon(x)\cos(\lambda x)dx \right| + \int_a^b |f(x) - F_\varepsilon(x)|dx \\ &< \varepsilon + \int_a^b \big(F_\varepsilon(x) - f(x)\big)dx.\end{aligned}$$

Thus, by (10.6) and (10.7)

$$\left| \int_a^b f(x)\cos(\lambda x)dx \right| < 2\varepsilon.$$

$\square$

10.2.1 Exercises

bsec:eleven1

1. Suppose that $a \leq b$ and $f \in \mathcal{R}[a,b]$. Show that for every $k \in \mathbb{N}$ we have $f^k \in \mathcal{R}[a,b]$.

2 (The Cauchy-Schwarz inequality for integrals). Prove that if $a \leq b$ and $f, g \in \mathcal{R}[a,b]$, then

$$\left(\int_a^b f(x)g(x)dx \right)^2 \leq \left(\int_a^b f(x)^2 dx \right) \left(\int_a^b g(x)^2 dx \right).$$

10.3 Integration of Continuous Functions

sec:eleven2

The next theorem shows that there is a plentiful supply of Riemann integrable functions.

thm:eleven2

Theorem 10.9. *Suppose $a \leq b$ and that $f : [a,b] \mapsto \mathbb{R}$ is continuous on $[a,b]$. Then $f \in \mathcal{R}[a,b]$.*

Proof. We can certainly suppose that $a < b$. Let Δ be an arbitrary partition in $\mathcal{D}[a,b]$. Then

$$\begin{aligned} \underline{S}([a,b],\Delta) &= \sum_{j=1}^{n} (x_j - x_{j-1}) \inf\{f(x) : x \in [x_{j-1}, x_j]\} \\ &\leq \sum_{j=1}^{n} (x_j - x_{j-1}) \sup\{f(x) : x \in [x_{j-1}, x_j]\} = \overline{S}([a,b],\Delta). \end{aligned}$$

Hence, by (10.3) we have

$$\begin{aligned} 0 &\leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq \\ &\sum_{j=1}^{n} (x_j - x_{j-1}) \big(\sup\{f(x) : x \in [x_{j-1}, x_j]\} - \inf\{f(x) : x \in [x_{j-1}, x_j]\} \big). \end{aligned} \tag{10.8}$$

eq:eleven3a

Let $\varepsilon > 0$. By Theorem 8.10 and Definition 8.4 there is a $\delta > 0$ so that whenever $x, y \in [a,b]$ and $|x - y| < \delta$ we have

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}. \tag{10.9}$$

eq:eleven4

Choose $n > (b-a)/\delta$ and $x_j = a + \frac{b-a}{n} j$. Since f is continuous on $[x_{j-1}, x_j]$ there are $x, y \in [x_{j-1}, x_j]$ so that $\sup\{f(x) : x \in [x_{j-1}, x_j]\} = f(y)$ and $\inf\{f(x) : x \in [x_{j-1}, x_j]\} = f(x)$. Therefore

$$|x - y| \leq \frac{b-a}{n} < \delta$$

and so (10.9) holds. Thus

$$\sum_{j=1}^{n}(x_j - x_{j-1})\big(\sup\{f(x) : x \in [x_{j-1}, x_j]\} - \inf\{f(x) : x \in [x_{j-1}, x_j]\}\big)$$
$$< \sum_{j=1}^{n}(x_j - x_{j-1})\frac{\varepsilon}{b-a} = \varepsilon.$$

Hence, by (10.8)

$$0 \le \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx < \varepsilon$$

and since this holds for every $\varepsilon > 0$ we have equality. □

`def:eleven6` **Definition 10.6.** *Suppose that $a < b$ and $f : [a,b] \mapsto \mathbb{R}$, that $F : [a,b] \mapsto \mathbb{R}$ is differentiable on (a,b), continuous on $[a,b]$ and $F'(x) = f(x)$ for $x \in (a,b)$. Then F is called a* **primitive** *of f on $[a,b]$.*

One of the benefits of a continuous integrand is that one may well be able to spot a primitive for the integrand which enables us to easily perform the integration.

`thm:eleven7` **Theorem 10.10** (The Fundamental Theorem of Calculus)**.** *Suppose that $a \le b$ and f is continuous on $[a,b]$, and for $a \le y \le b$ let $I(y)$ be defined by*

$$I(y) = \int_a^y f(x)dx.$$

Then I is a primitive for f on $[a,b]$. Moreover, suppose that $F : [a,b] \mapsto \mathbb{R}$ is a primitive of f on [a,b]. Then for $a \le y \le b$ we have

$$\int_a^y f(x)dx = F(y) - F(a).$$

Proof. Suppose $a < y < b$. Let $\varepsilon > 0$. By the definition of continuity there is a $\delta > 0$ so that when $a \le y - \delta < x < y + \delta \le b$ we have

$$|f(x) - f(y)| < \varepsilon/2.$$

Suppose that $|h| < \delta$. Then

$$\frac{I(y+h) - I(y)}{h} - f(y) = \frac{1}{h}\int_y^{y+h} \big(f(x) - f(y)\big)$$

and by Theorem 10.5 and Corollary 10.2 we have

$$\left|\frac{1}{h}\int_y^{y+h} \big(f(x) - f(y)\big)dx\right| \le \frac{1}{|h|}\left|\int_y^{y+h} \big|f(x) - f(y)\big|dx\right| \le \varepsilon/2 < \varepsilon$$

(note that we might have $h < 0$) which gives the first part of the theorem.

For the second part note first that the conclusion is immediate when $y = a$. Hence we may suppose that $a < y \le b$. Then we have $F'(y) - I'(y) = 0$. Therefore, by the Mean Value Theorem, Theorem 9.7, we have

$$\big(F(y) - I(y)\big) - \big(F(a) - I(a)\big) = (y-a)\big(F'(\xi) - I'(\xi)\big) = 0$$

for some $\xi \in (a, y)$. Thus

$$F(y) - F(a) = I(y) - I(a) = I(y).$$

$\square$

From this and the chain rule we have this.

thm:eleven8 **Theorem 10.11** (The substitution rule)**.** *Suppose that $a < b$, $g : I \mapsto \mathbb{R}$ where I is an open interval containing $[a, b]$, g is differentiable on I and g' is continuous on $[a, b]$. Suppose also that $f : J \mapsto \mathbb{R}$ where J is an open interval containing $g([a, b])$ and f is continuous on J. Then*

$$\int_a^b f\big(g(x)\big)g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt.$$

Proof. For $a \le y \le b$ $u \in T$ let

$$F(y) = \int_a^y f\big(g(x)\big)g'(x)dx, \quad G(u) = \int_{g(a)}^{u} f(t)dt.$$

Thus, by the fundamental theorem of calculus

$$F'(y) = f\big(g(y)\big)g'(y), \quad G'(u) = f(u).$$

Hence $G\big(g(x)\big)$ is a primitive for $f\big(g(x)\big)g'(x)$ and so

$$\int_a^b f\big(g(x)\big)g'(x)dx = G\big(g(b)\big) - G\big(g(a)\big) = G\big(g(b)\big).$$

$\square$

Another important consequence of the Fundamental Theorem is integration by parts.

thm:eleven8a **Theorem 10.12.** *Suppose that $a < b$ and $f, g : [a, b] \mapsto \mathbb{R}$, that f and g are differentiable on (a, b), and $f(x)g'(x)$ and $f'(x)g(x)$ are continuous on $[a, b]$. Then*

$$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx.$$

Proof. This follows on observing that $f(x)g(x)$ is a primitive of $f(x)g'(x)+f'(x)g(x)$. $\square$

thm:eleven8b **Corollary 10.13.** *Suppose that $a \leq b$ and f is differentiable on (a,b) and that f and g are continuous on $[a,b]$. Let G be a primitive of g on $[a,b]$. Then*

$$\int_a^b f(x)g(x)dx = f(b)G(b) - f(a)G(a) - \int_a^b f'(x)G(x)dx.$$

On combining the first part of Theorem 10.10 with the Mean Value Theorem, Theorem 9.7 we have the following.

thm:eleven9 **Theorem 10.14** (The Mean Value Theorem for Integrals)**.** *Suppose that $a \leq b$ and f is continuous on $[a,b]$. Then there is a $\xi \in (a,b)$ so that*

$$\int_a^b f(x)dx = (b-a)f(\xi).$$

10.3.1 Exercises

bsec:eleven2

1. Let for $x \in [a,b]$ and $k \in \mathbb{N}$ define

$$P(x) = \sum_{j=0}^{k} c_j x^j.$$

Prove that

$$\int_a^b P(x)dx = \sum_{j=0}^{k} c_j \frac{b^{j+1}-a^{j+1}}{j+1}.$$

2. Suppose

$$A(x) = \sum_{n=0}^{\infty} a_n x^n$$

is a power series with radius of convergence $0 < R$.

(i) Show that the power series

$$B(x) = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1}$$

also has radius of convergence R.

(ii) Suppose that $a, b \in \mathbb{R}$ with $|a| < R$ and $|b| < R$. Prove that B is a primitive of A for $|x| < R$ and that

$$\int_a^b A(x)dx = B(b) - B(a) = \sum_{n=0}^{\infty} a_n \frac{b^{n+1}-a^{n+1}}{n+1}$$

3. Suppose that $-1 < a < 1$, that $a \le t \le 1$ and $n \in \mathbb{N}$.
   (i) Prove that

$$\int_0^t \frac{1-(-x)^n}{1+x} = \sum_{m=1}^{n} \frac{(-1)^{m-1}}{m} t^m$$

   (ii) Prove that

$$\left| \log(1+t) - \sum_{m=1}^{n} \frac{(-1)^{m-1}}{m} t^m \right| \le \frac{t^{n+1}}{n+1}.$$

(iii) Deduce that if $-1 < t \le 1$ we have

$$\log(1+t) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} t^m$$

and in particular that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

10.4 Improper Riemann Integrals

sec:eleven4

There are a number of situations where for technical reasons the Riemann integral does not exist but we can adapt it to give a reasonable version of an integral. Typically these concern either the integrand being unbounded, or the desire for one or both of the limits a, b to be $\pm\infty$ respectively.

def:eleven7 **Definition 10.7.** *Suppose that $a < b$, that for every $a \le \xi < b$ we have $f \in \mathcal{R}[a, \xi]$ and that*

$$\ell = \lim_{\xi \to b-} \int_a^\xi f(x)dx$$

exists. Then we define the first kind of **improper Riemann integral** *by*

$$\int_a^b f(x)dx = \ell.$$

Likewise when $a < \eta \le b$ and $f \in \mathcal{R}[\eta, b]$ we define

$$\int_a^b f(x)dx = \lim_{a \to +} \int_\eta^b f(x)dx$$

when the limit exists.

def:eleven8 **Definition 10.8.** *Let a be given. Suppose that for every $b \ge a$ we have $f \in \mathcal{R}[a, b]$ and that*

$$\ell = \lim_{b \to \infty} \int_a^b f(x)dx$$

exists. Then we define the second kind of **improper Riemann integral** *by*

$$\int_a^\infty f(x)dx = \ell.$$

Likewise when given b, for every $a \leq b$ we have $f \in \mathcal{R}[a,b]$ we define

$$\int_{-\infty}^b f(x)dx = \lim_{a\to-\infty}\int_a^b f(x)dx$$

when the limit exists.

There is a variation of this kind which can occur, namely when there is something "singular" happening at a point c between a and b.

def:eleven9 **Definition 10.9.** *Suppose that $a < c < b$ be given, that for every η and ξ with $a \leq \eta < c < \xi \leq b$ we have $f \in \mathcal{R}[a,\eta]$ and $f \in \mathcal{R}[\xi,b]$, and that*

$$\int_a^c f(x)dx \text{ and } \int_c^b f(x)dx$$

both exist as improper Riemann integrals of the first kind. Then we can define

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

However it can happen that the two improper integrals above do not exist but nevertheless somehow the effects of the singularity from the left cancel those from the right.

def:eleven10 **Definition 10.10.** *Suppose that $a < c < b$ and for each $\eta > 0$ with $a \leq c-\eta$ and $c+\eta \leq b$ we have $f \in \mathcal{R}[a,c-\eta]$ and $f \in \mathcal{R}[c+\eta,b]$. Suppose also that*

$$\ell = \lim_{\eta\to 0+}\left(\int_a^{c-\eta} f(x)dx + \int_{c+\eta}^b f(x)dx\right)$$

exists. Then we define the **Cauchy Principal Value (CPV)** *by*

$$(CPV)\int_a^b f(x)dx = \ell.$$

There is a further variation on this.

def:eleven11 **Definition 10.11.** *Suppose that for every $b > 0$ we have $f \in \mathcal{R}[-b,b]$ and neither*

$$\int_0^\infty f(x)dx \text{ nor } \int_{-\infty}^0 f(x)dx$$

exist as improper Riemann integrals of the second kind, but nevertheless

$$\ell = \lim_{b\to\infty}\int_{-b}^b f(x)dx$$

exists. Then we define the **Cauchy Principal Value (CPV)** *by*

$$(CPV)\int_{-\infty}^\infty f(x)dx = \ell.$$

10.4.1 Exercises

bsec:eleven4

1. Suppose $f \in \mathcal{R}[a,b]$ for every $b \geq a$ and

$$\int_a^\infty |f(x)| dx$$

exist. Prove that so does

$$\int_a^\infty f(x) dx, \tag{10.10}$$

eq:eleven10

and

$$\left| \int_a^\infty f(x) dx \right| \leq \int_a^\infty |f(x)| dx,$$

In this circumstance we say that the integral in 10.10 converges absolutely.

2. **The Integral Test for convergence of series.** Suppose that $f : [1, \infty) \mapsto \mathbb{R}^+$ is a decreasing function. Then prove that

$$\sum_{n=1}^\infty f(n)$$

converges if and only if the improper Riemann integral

$$\int_1^\infty f(x) dx$$

exists.

3. (**Euler.**) Prove that

$$\lim_{n \to \infty} \left(\sum_{m=1}^n \frac{1}{m} - \log n \right)$$

exists. The limit is usually denoted by γ and is known as Euler's constant. Its first 50 decimal places are

$$0.57721566490153286060651209008240243104215933593992...$$

Euler calculated the first 19 (by hand, of course!).

4. Prove that

$$\int_0^\infty \frac{\sin x}{x} dx$$

exists as a (double!) improper integral.

10.5 Notes

sec:elevenN

The version of the Riemann integral as presented here is as modified by Darboux. In Riemann's time the concept of infimum and supremum had not been introduced and instead Riemann defined lower and upper sums by choosing values of f in the relevant interval.

Appendix A

The Complex Numbers

ap:A

A.1 Construction

sec:A1

The complex numbers can be defined in terms of real numbers by a similar process to that used to construct the rational numbers from the integers.

def:A1 **Definition A.1.** *The Complex Numbers We define the complex numbers $\mathbb{C}$ as the set of ordered pairs of real numbers (x, y) together with two operations $+$ and $\times$ which satisfy the following axioms, and which hold for any $x, y, u, v \in \mathbb{R}$.*

1. *C1* $(x, y) + (u, v) = (x + u, y + v)$,
2. *C2* $u(x, y) = (ux, uy)$,
3. *C3* $(x, y) \times (u, v) = (xu - yv, xv + yu)$.

Closure, Commutativity, Associativity, Distributivity *follow from the corresponding properties of* $\mathbb{R}$.

It is readily checked that

$$(x, y) + (0, 0) = (x, y) = (0, 0) + (x, y),\ (1, 0) \times (x, y) = (x, y) = (x, y) \times (1, 0),$$

so $(0, 0)$ *and* $(1, 0)$ *act as* **identities**. *Since*

$$(x, y) + (-x, -y) = (-x, -y) + (x, y) = (0, 0)$$

the ordered pair $(-x, -y)$ *acts as an* **additive inverse**, *and it is convenient to write* $-(x, y)$ *for* $(-x, -y) = (-1)(x, y)$. *We further define the* **modulus** *of* (x, y) *by*

$$|(x, y)| = (x^2 + y^2)^{1/2}$$

and then when $(x, y) \neq (0, 0)$ *we have*

$$(x, y) \times \left(\frac{x}{|(x, y)|^2}, \frac{-y}{|(x, y)|^2} \right) = (1, 0).$$

so $\left(\frac{x}{|(x,y)|^2}, \frac{-y}{|(x,y)|^2} \right)$ *acts as a* **multiplicative inverse**.

Thus the arithmetic axioms for a field are satisfied. We also have

$$(0,1)^2 = (-1,0) = (-1)(1,0) = -(1,0).$$

At this point it is usual to discard the cumbrous notation and use 0 for $(0,0)$, 1 for $(1,0)$ and i for $(0,1)$ and then one can see that $(x,y) = x(1,0) + y(0,1) = x + iy$ and we typically write $z = x + iy$. For historical reasons, x is called the *real* part of z and y the *imaginary* part of z and we write $x = \operatorname{Re} z$, $y = \operatorname{Im} z$.

One thing that immediately needs to be laid to rest. Unlike the real numbers the complex numbers cannot satisfy the order relations O1,...,O4 of Definition 2.2. For suppose they do. Clearly $i \neq 0$, so is $i > 0$ or < 0? Suppose the former. Then, by O4, $0 = 0.i < i^2 = (0,1) \times (0,1) = (-1,0) = -(1,0) = -1$, Then $0 = 0.i < (-1)i = -i$, $i = 0 + i < -i + i = 0$, so we just proved that $0 < i < 0$! The same thing would happen if we supposed that $i < 0$.

One other piece of notation. Given $z = x + iy$ it is usual to write $\overline{z}$ for $x - iy$ and call this the **conjugate** of z. Then $z + \overline{z}$ is real and one can check that

$$|z|^2 = z\overline{z} \text{ and when } z \neq 0 \text{ we have } z^{-1} = \frac{\overline{z}}{|z|^2}.$$

By definition $|z| = (x^2 + y^2)^{1/2} \geq (x^2)^{1/2} = |x| \geq x = \operatorname{Re} z$. Thus we can adapt the proof of the triangle inequality from Theorem 2.8.

thm:app1 **Theorem A.1.** *Suppose that $z, w \in \mathbb{C}$. Then*

$$|zw| = |z||w| \text{ and } |z + w| \leq |z| + |w|.$$

Proof. Put $z = x + iy$, $w = u + iv$. Then $zw = zu - yv + i(xv + yu)$ and it is readily verified that $|zw|^2 = (xu - yv)^2 + (xv + yu)^2 = (x^2 + y^2)(u^2 + v^2) = |z|^2|w|^2$ which establishes $|zw| = |z||w|$.

When $w \neq 0$ we also have $|z| = |(z/w)w| = |z/w||w|$ so that

$$\left|\frac{z}{w}\right| = \frac{|z|}{|w|}. \tag{A.1}$$ eq:appzw

For the second assertion, note first that it is immediate when $z + w = 0$. so we may suppose that $z + w \neq 0$. Then, by (A.1)

$$\frac{|z| + w|}{|z + w|} = \frac{|z|}{|z + w|} + \frac{|w|}{|z + w|} = \left|\frac{z}{z + w}\right| + \left|\frac{w}{z + w}\right| \geq \operatorname{Re}\frac{z}{z + w} + \operatorname{Re}\frac{w}{z + w} = 1.$$

$\square$

subsec:A1

A.1.1 Exercises

1. Prove that if $z, w \in \mathbb{C}$, then $\big||z| - |w|\big| \leq |z - w|$.

2. Let $A, C \in \mathbb{R}$ and $B \in \mathbb{C}$. If we represent $\mathbb{C}$ by $\mathbb{R}^2$ by associating $z = x + iy$ with (x, y), show that every circle in $\mathbb{R}^2$ can be represented by an equation of the form

$$Az\overline{z} + Bz + \overline{B}\overline{z} + C = 0$$

with $A \neq 0$ and every line in $\mathbb{R}^2$ can be represented by an equation of the form

$$Bz + \overline{B}\overline{z} + C = 0$$

3. Suppose that $a, b, c, d \in \mathbb{C}$ with $ad \neq bc$ and $c \neq 0$, and let $\mathcal{A} = \mathbb{C} \setminus \{-d/c\}$. Suppose $f : \mathcal{A} \mapsto \mathbb{C}$ is given by

$$f(z) = \frac{az + b}{cz + d}.$$

(i) Prove that $f(\mathcal{A}) = \mathcal{B}$ where $\mathcal{B} = \mathbb{C} \setminus \{a/c\}$ and that the function is bijective.
(ii) Prove that

$$f^{-1}(w) = \frac{dw - b}{-cw + a}.$$

These are **Möbius transformations** and they have an interesting structure which leads to a major area of research, namely modular forms.

sec:app2

A.2 Complex Sequences

Much of the theory we developed for real sequences can be ported over to complex sequences, that is sequences $\langle a_n \rangle$ where $a_n \in \mathbb{C}$. The definition of convergence is identical, namely that there is an $\ell \in \mathbb{C}$ such that for every $\varepsilon > 0$ there exists an N such that whenever $n > N$ we have

$$|a_n - \ell| < \varepsilon.$$

Whilst the above remarks show that monotonicity does not make sense for complex sequences the Bolzano-Weierstrass still holds. The reason is that if $\langle a_n \rangle$ is bounded, then so are $\langle \operatorname{Re} a_n \rangle$ and $\langle \operatorname{Im} a_n \rangle$. Thus $\langle \operatorname{Re} a_n \rangle$ has a convergent subsequence $\langle \operatorname{Re} a_{m_n} \rangle$, and then by the same token $\langle \operatorname{Im} a_{m_n} \rangle$ has a convergent subsequence $\langle \operatorname{Im} a_{m_{k_n}} \rangle$. Hence $\langle a_{m_{k_n}} \rangle$ gives a convergent subsequence of $\langle a_n \rangle$.

subsec:A2

A.2.1 Exercises

1. Let $z \in \mathbb{Z}$ and define $a_n = (1 + z/n)^n$. Prove that

$$\lim_{n\to\infty} a_n = \exp(z).$$

A.3 Complex Series and Integrals

sec:app3

Just as for sequences the theory we developed for series can mostly be generalised, if necessary by separating out the real and imaginary parts. Amongst the tests for convergence only the Leibnitz test cannot be generalised, although it could be applied separately to the real and imaginary parts.

The situation for power series

$$\sum_{n=1}^{\infty} a_n z^n$$

is interesting. For real series (with $z = x$) we introduced the radius convergence R and showed that the series converges absolutely on the oven interval $(-R, R)$. For complex series we have a similar conclusion except that now we have convergence on the open disc $|z| < R$.

For functions $f : [a, b] \mapsto \mathbb{C}$, when $a \leq b$ we can set up a theory of integration by supposing that both $\operatorname{Re} f$ and $\operatorname{Im} f \in \mathcal{R}[a, b]$ and then defining

$$\int_a^b f(x)dx = \int_a^b \operatorname{Re} f(x)dx + i \int_a^b \operatorname{Im} f(x)dx.$$

We can also extend the concept to the various kinds of improper integrals. It can also be extended to paths in the complex plane. Suppose that there is a path $\mathcal{P}$ from w_1 to w_2 which can be parameterised by taking a (complex valued) function $g(t)$ defined on the real interval $[a, b]$ and which has the properties $g(a) = w_1$, $g(b) = w^2$ and as t varies from a to b the function $g(t)$ describes the path $\mathcal{P}$ from a to b. Suppose also that g is differentiable, i.e. the real and imaginary parts are differentiable. Then we can define

$$\int_{\mathcal{P}} f(z)dz = \int_a^b f\big(g(t)\big)g'(t)dt,$$

which is exactly what we would anticipate from the substitution rule in the real case.

The generalizations of continuity, differentiation and integration when the underlying functions are defined on more general subsets of $\mathbb{C}$ are best studied systematically in a course of complex analysis where the consequences of considering functions $f : \mathcal{A} \mapsto \mathcal{C}$ where $\mathcal{A} \subset \mathbb{C}$ are explored systematically. Underpinning it is a rich, powerful and very beautiful theory which is one of the gems of mathematics. In that direction we can define for $z \in \mathbb{C}$

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Then as in the real case this radius of convergence ∞ and

$$\exp(z + w) = \exp(z)\exp(w).$$

In particular

$$\exp(x + iy) = \exp(x)\exp(iy).$$

We can likewise define

$$\cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

and

$$\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$

and it is readily checked that

$$\exp(iz) = \cos(z) + i\sin(z),$$

$$\cos(-z) = \cos(z),\ \sin(-z) = -\sin(z),\ \cos(0) = 1,\ \sin(0) = 0,$$

$$\cos(z) = \frac{1}{2}\big(\exp(iz) + \exp(-iz)\big),\ \sin(z) = \frac{1}{2i}\big(\exp(iz) - \exp(-iz)\big).$$

In particular, we have Euler's formula

$$\exp(x + iy) = \exp(x)\big(\cos(y) + i\sin(y)\big).$$

A.3.1 Exercises

subsec:A3

1. Suppose $\langle a_n \rangle$ and $\langle a_n \rangle$ are complex sequences, that $R > 0$, and that

$$A(z) = \sum_{n=0}^{\infty} a_n z^n \tag{A.2}$$

eq:ANA3

and

$$B(z) = \sum_{n=0}^{\infty} b_n z^n \tag{A.3}$$

eq:ANA4

are both convergent and satisfy $A(z) = B(z)$ for $|z| < R$. Then Prove that $a_n = b_n$ for every non-negative integer n.

2. (i) Prove that if $\operatorname{Re} z > 1$, then the series

$$\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$$

converges absolutely and so does the improper integral

$$\int_1^{\infty} x^{-z} dx.$$

(ii) For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the greatest integer not exceeding x, namely

$$\lfloor x \rfloor = \max\{n : n \in \mathbb{Z} \text{ and } n \leq x\}.$$

Prove that if $\operatorname{Re} z > 0$, then

$$I(z) = \int_1^\infty (x - \lfloor x \rfloor) x^{-z-1} dx$$

converges absolutely for $\operatorname{Re} z > 0$.

(iii) Prove that if $\operatorname{Re} z > 1$, then we have

$$\zeta(z) = \frac{z}{z-1} + zI(z).$$

Observe that the expression on the right can be used to give a definition to $\zeta(z)$ for all $\operatorname{Re} z > 0$ with $z \neq 1$.

If you can prove that all the zeros ρ of $\zeta(z)$ with $\operatorname{Re} z > 0$ have $\operatorname{Re} z = \frac{1}{2}$, then you will become very famous!

A.4 Notes

sec:eleven8

For an account of the history of the development of complex numbers, see https://en.wikipedia.org/wiki/Complex_number