# Analysis I

James C. Robinson

Autumn Term 2022

# Contents

# Preliminaries

## Notation

We will use the standard notation for various sets of numbers.

- $\mathbb{N}$ denotes the natural numbers: $\mathbb{N} = \{1, 2, 3, 4, \cdots\}$. (In this module $0 \notin \mathbb{N}$.)
- $\mathbb{Z}$ denotes the integers $\{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$.
- $\mathbb{Q}$ denotes the rational numbers - all numbers of the form $p/q$, where $p$ and $q$ are integers.
- $\mathbb{R}$ denotes the set of all real numbers.
- *Intervals* are sets of real numbers written as $(-\infty, b)$, $(-\infty, b]$, $(a, b)$, $(a, b]$, $[a, b)$, $[a, b]$, $(a, \infty)$, $[a, \infty)$, or $\mathbb{R}$. We use a round bracket for a strict inequality and a square bracket for a non-strict inequality, e.g.

  $$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}.$$

  Intervals of the form $(a, b)$ are called *open* and of the form $[a, b]$ are called *closed*.

We will often use the Greek letter epsilon, $\varepsilon$, to denote a (small) positive number. [There are two standard pronunciations in British English, epsílon and epsīlon.] It is not to be confused with the 'is a member of' sign $\in$ (which can happen since epsilon it is sometimes typeset as $\epsilon$), or with a capital sigma ($\Sigma$, which we will use later to represent the sum) written very small. When we move on to discuss continuity, we will also be using the Greek letter delta, $\delta$, to be another (often small) positive number.

The binomial theorem: if $n \in \mathbb{N}$ then

$$(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j, \quad \text{where} \quad \binom{n}{j} = \frac{n!}{(n-j)!j!},$$

with the convention that $0! = 1$.

## Axioms for the real numbers: arithmetic and order

Addition

(A1) if $x, y \in \mathbb{R}$ then $x + y \in \mathbb{R}$ (closure under addition)

(A2) if $x, y, z \in \mathbb{R}$ then $(x + y) + z = x + (y + z)$ (addition is associative)

(A3) if $x, y \in \mathbb{R}$ then $x + y = y + x$ (addition is commutative)

(A4) there exists $0 \in \mathbb{R}$ such that for any $x \in \mathbb{R}$, $x + 0 = 0 + x = x$ (existence of an additive identity)

(A5) if $x \in \mathbb{R}$ there exists $-x \in \mathbb{R}$ such that $x + (-x) = (-x) + x = 0$ (existence of an additive inverse)

Multiplication

(M1) if $x, y \in \mathbb{R}$ then $xy \in \mathbb{R}$ (closure under multiplication)

(M2) if $x, y, z \in \mathbb{R}$ then $(xy)z = x(yz)$ (multiplication is associative)

(M3) if $x, y \in \mathbb{R}$ then $xy = yx$ (multiplication is commutative)

(M4) there exists $1 \in \mathbb{R}$, such that for any $x \in \mathbb{R}$, $1.x = x.1 = x$ (existence of multiplicative identity)

(M5) if $x \in \mathbb{R}$ with $x \neq 0$, there exists $x^{-1} \in \mathbb{R}$ such that $x.x^{-1} = x^{-1}.x = 1$ (existence of multiplicative inverses)

Interaction between addition and multiplication

(D) if $x, y, z \in \mathbb{R}$, $x(y + z) = xy + xz$ (distributive property)

Order properties. We use $y > x$ as an equivalent for $x < y$, and $x \leq y$ means that either $x < y$ or $x = y$.

(O1) if $x, y \in \mathbb{R}$ exactly one of the statements $x < y$, $x = y$, $x > y$ is true (trichotomy)

(O2) if $x, y, z \in \mathbb{R}$ then if $x < y$ and $y < z$, $x < z$ (transitivity)

Interaction between order and arithmetic

(O3) if $x, y, z \in \mathbb{R}$ and $x < y$, then $x + z < y + z$

(O4) if $x, y, z \in \mathbb{R}$, $x < y$, and $z > 0$, then $xz < yz$

There are some familiar properties of arithmetic that seem to be missing from the above list, but you can prove them using these axioms and some ingenuity. These sort of arguments are a bit tricksy, and either fun or painful depending on your taste. It's very rare to have to argue in this sort of 'micro-logical' way.

- there is only one element $0 \in \mathbb{R}$.

  Suppose that there are two elements, 0 and $0'$, that both satisfy (A4). Then $0 = 0 + 0' = 0'$ (first use (A4) for $0'$ and then use (A4) for 0).

- if $x, y, z \in \mathbb{R}$, $x + z = y + z$, then $x = y$. [Additive cancellation.]

  Add $-z$ to both sides, use (A2), then (A5), then (A4):

  $x = x+0 = x+(z+(-z)) = (x+z)+(-z) = (y+z)+(-z) = y+(z+(-z)) = y+0 = y.$

- for any $x \in \mathbb{R}$, $0 \cdot x = 0$. (See Examples 1.)

- for any $x \in \mathbb{R}$, $-x = (-1) \cdot x$.

  From the previous fact and (A5)

  $$0 = (1 + (-1))x = x + (-1)x;$$

  use (A5) again and add $-x$ to both sides to give (using (A2) and (A5)) $-x = (-1)x$.

- If $ab = ac$ and $a \neq 0$ then $b = c$. [Multiplicative cancellation.]

  We use (M5), (M2), and (M4):

  $$b = 1 \cdot b = (a^{-1} \cdot a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = 1 \cdot c = c.$$

- $(-a)b = -(ab)$. (See Examples 1.)

- $(-a)(-b) = ab$. Use the previous fact along with $-(-x) = x$.

- if $x \in \mathbb{R}$ with $x > 0$ then $-x < 0$.

  Start with $x > 0$, and add $-x$ to both sides: using (A5) and (O3) gives $0 = (-x) + x > -x$, and $0 > -x$ is the same as $-x < 0$.

- if $x \in \mathbb{R}$ with $x < 0$ then $-x > 0$. (Essentially the same argument.)

- if $x < y$ and $z < 0$ then $xz > yz$.

  Since $-z > 0$, we have $(-z)x < (-z)y$, which is $-zx < -zy$. Now add $zx + zy$ to both sides to obtain $zy < zx$.

- if $0 < a < b$ and $0 < c < d$ then $ac < bd$.

  Since $c > 0$ and $a < b$ we have $ac < bc$ (O4). Since $b > 0$ and $c < d$ we have $bc < bd$ (O4) again. Now we use (O2) [transitivity]: $ac < bc$ and $bc < bd$ implies that $ac < bd$.

- $x^2 > 0$ for every $x \neq 0$. (See Examples 1.)

# *Fields and ordered fields

Suppose that we replace $\mathbb{R}$ on the previous page with some abstract set $\mathbb{F}$, and define two 'binary operations' $+$ and $\times$ that both take two elements of $\mathbb{F}$ and combine them to give another element of $\mathbb{F}$ (we could write $+ \colon \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ and $\times : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$).

If $\mathbb{F}$, along with $+$ and $\times$ satisfy the properties for Addition, Multiplication, and their interaction on the previous page [(A1–5), (M1–5), and (D)] then $\mathbb{F}$ is called a *field.*

So $\mathbb{R}$ is a field, and so is $\mathbb{Q}$ (we will check this shortly). In Analysis, the two most important fields are $\mathbb{R}$ and $\mathbb{C}$ (the complex numbers).

But there are other more abstract examples that you will meet, primarily in Algebra. A fairly simple example is to choose a prime number $p$, take $\mathbb{F} = \{0, 1, \ldots, p-1\}$ (integers from 0 to $p - 1$) and define $+$ and $\times$ to be addition and multiplication *modulo* $p$ (i.e. you keep the remainder when you divide by $p$).

If you can also define an order on $\mathbb{F}$ (i.e. find a way of giving meaning to the statement $x < y$ for $x, y \in \mathbb{F}$) that satisfies the order properties on the previous page and interacts with $+$ and $\times$ in the right way, then $\mathbb{F}$ is an *ordered field.*

Both $\mathbb{R}$ ('of course') and $\mathbb{Q}$ (easy to chek) are ordered fields. However, you can prove from the basic rules that in any ordered field $x^2 > 0$ for every $x \neq 0$ (see Examples), so $\mathbb{C}$ is not an ordered field (since $\mathrm{i}^2 = 1- < 0$).

An interesting example of an ordered field (in that it provides a useful example to demonstrate that the assumptions on page v are not enough to tell us everything we want about the real numbers) is the collection of all functions of the form

$$\frac{P(x)}{Q(x)},$$

where $P$ and $Q$ are both polynomials with real coefficients and $Q \neq 0$; these are called 'rational functions'. We can define addition and multiplication in the usual way, and a consistent definition of an order is to say that if $p(x) = p_0 x^n + \cdots$ and $q(x) = q_0 x^m + \cdots$ then $p(x)/q(x) > 0$ whenever $p_0/q_0 > 0$.

## Powers and inequalities

We will prove the following very useful result, and along the way demonstrate the difference between direct proof, proof by contradiction, and proof by the contrapositive.

**Lemma.** *If $n \in \mathbb{N}$ and $x, y > 0$ then*

$$\begin{aligned} x < y \quad &\Leftrightarrow \quad x^n < y^n \\ x \le y \quad &\Leftrightarrow \quad x^n \le y^n \\ x = y \quad &\Leftrightarrow \quad x^n = y^n. \end{aligned}$$

*Proof.* The argument is similar for the first two cases; we prove the first.

We use induction to prove the forward implication. We start the induction with $x < y$. Now suppose that $0 < x^k < y^k$; then since $0 < x < y$ we can use the penultimate fact on page iv[1] to deduce that $0 < x^{k+1} < y^{k+1}$. Induction now shows that $x < y$ implies that $x^n < y^n$ for every $n \in \mathbb{N}$.

Noting that if $x = y$ then clearly $x^n = y^n$, we have also obtained $x \le y \Rightarrow x^n \le y^n$.

To prove the reverse implication we can use contradiction or a 'contrapositive' argument.

1 - Contradiction

Suppose that $x^n < y^n$ but that $x$ is not less than $y$, i.e. that $x \ge y$. But we just proved that if $x \ge y$ then $x^n \ge y^n$, contradicting our initial assumption (that $x^n < y^n$). So we must have $x < y$.

2 - Contrapositive

To show that $x^n < y^n$ implies that $x < y$ it is enough to prove the contrapositive: that $x \ge y$ implies that $x^n \ge y^n$. But this is part of the previous lemma.

The $x = y \Leftrightarrow x^n = y^n$ is proved on Examples 1. □

(Proof by the contrapositive (you will see this further in Foundations): to show that $P \Rightarrow Q$, you can also show that not $Q \Rightarrow$ not $P$. This is a hidden contradiction argument (as in the first proof above): suppose that $P$ is true and that $Q$ is not true. If you then show that 'not $Q$' implies 'not $P$' you get a contradiction, so $Q$ must be true.)

---

1 Without this we can transpose the same argument to this particular case: since $x > 0$ we have $x^{k+1} = x \cdot x^k < xy^k$ and since $x < y$ and $y^k > 0$ we have $xy^k < y \cdot y^k = y^{k+1}$; we use the transitivity property (O2) to deduce that $x^{k+1} < y^{k+1}$.

# Part I

# The real numbers

# Chapter 1

# Rational and irrational numbers

## 1.1 $\mathbb{Q}$ contains no number whose square is 2.

All of the axioms on page v hold if we replace the real numbers by the rational numbers. They are still closed under addition and multiplication, and the additive of multiplicative inverses of rational numbers are again rational.

**Lemma 1.1.** *The sum and product of any two rational numbers is again rational.*

*Proof.* Take two rational numbers $p/q$ and $p'/q'$; then

$$\frac{p}{q} + \frac{p'}{q'} = \frac{pq' + p'q}{qq'}$$

and

$$\frac{p}{q} \times \frac{p'}{q'} = \frac{pp'}{qq'}.$$

□

It is easy to check that the additive and multiplicative inverses of rational numbers are again rational: $-(p/q) = -p/q$ and $(p/q)^{-1} = q/p$.

There are many things we cannot do if we restrict to rational numbers: for example, we cannot in general take square roots.

**Proposition 1.2.** *There is no rational number whose square is* 2.

It is tempting to write this proposition as '$\sqrt{2}$ is irrational', but since we are trying to demonstrate that rational numbers 'aren't everything' it would be a little strange to have a formal statement that contained a number ($\sqrt{2}$) that as yet we can't really put our fingers on.

*Proof.* Suppose that there is a rational number, say $p/q$, whose square is 2:

$$\left(\frac{p}{q}\right)^2 = 2. \tag{1.1}$$

We will assume that $p/q$ is a fraction in its lowest terms, i.e. $p$ and $q$ have no common factors, and then deduce a contradiction. So our initial assumption must be wrong.

Our assumption (1.1) means that

$$\frac{p^2}{q^2} = 2 \qquad \Rightarrow \qquad p^2 = 2q^2,$$

and $p^2$ must be even. The only way that $p^2$ can be even is that $p$ is even, so $p = 2m$ for some integer $m$. But then

$$(2m)^2 = 4m^2 = 2q^2 \qquad \Rightarrow \qquad q^2 = 2m^2,$$

so now $q^2$ is even. Arguing as before, $q$ must be even, $q = 2n$. This contradicts our initial assumption that $p$ and $q$ have no common factors, so our initial assumption must be wrong. □

(Similarly, if $p$ is an integer and $n$ is an integer that is not a perfect $p$th power, there is no rational number $r$ with $r^p = n$. This uses the uniqueness of prime factorisations, which you will cover in Foundations.)

## 1.2 Least upper bound axiom

How can we be sure that there is a number whose square is 2? We need another axiom, and this one (unlike all of those on page v) is *not* obvious.

Suppose that $S$ is a set of real numbers.

An *upper bound* for a set $S$ is a real number $r$ such that

$$s \leq r \qquad \text{for each } s \in S.$$

If a set has an upper bound then it is *bounded above.* [Any real number is an upper bound for the empty set, since the definition holds.]

There is a similar definition for what it means for a set to be 'bounded below' and of a 'lower bound'. Note that if $r$ is an upper bound for $S$ then $-r$ is a lower bound for the set $-S := \{-s : \ s \in S\}$ (and vice versa). So we can easily switch between statements for upper bounds and lower bounds.

If a set has one upper bound, it has many (infinitely many, in fact): if $s \leq r$ for every $s \in S$ then every $s \in S$ also satisfies $s \leq r'$ for any $r' \geq r$.

Before we look at some more interesting examples, we start with the simple observation that all of the intervals $(a, b)$, $(a, b]$, $[a, b)$, and $[a, b]$ are bounded below by $a$ and bounded above by $b$.

It is often not too hard to find upper and lower bounds; although sometimes it can be tricky (see (v) below).

(i) the set $\{1/n : \ n \in \mathbb{N}\}$ is bounded above by 1 and bounded below by 0.

(ii) the set[1] $\{\sin n : \ n \in \mathbb{N}\}$ is bounded below by $-1$ and bounded above by 1.

(iii) the interval $(-1, \infty)$ is bounded below by $-1$ and is not bounded above.

(iv) the set $\{1/2, 2/3, \cdots, n/(n+1) : \ n \in \mathbb{N}\}$ is bounded below by $1/2$ and bounded above by 1.

(v) the set $\{(1+\frac{1}{n})^n : \ n = 1, 2, \ldots\}$ is bounded below by 2 and above by 3 (in fact this set is bounded above by e, see Proposition **??**). We use the binomial expansion for both. The lower bound is easy:

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

*The upper bound requires a bit more thought:

$$\begin{aligned}
\left(1 + \frac{1}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{1}{n^k} \\
&= 1 + 1 + \sum_{k=2}^{n} \frac{1}{k!} \left(\frac{n(n-1)\cdots(n-k+1)}{n^k}\right) \\
&= 2 + \sum_{k=2}^{n} \frac{1}{k!} \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}\right) \\
&\leq 2 + \sum_{k=2}^{n} \frac{1}{k!} \\
&\leq 2 + \sum_{k=2}^{n} \frac{1}{k(k-1)} = 2 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k}\right) = 3 - \frac{1}{n} < 3.
\end{aligned}$$

1 What is the function $\sin x$? You already ‘know’ what it is, but how is it actually defined? We have suddenly jumped a level (or two) of complexity here. For now, rely on what you ‘know’ about the trig functions; we will revisit them later when we start to discuss continuity, and you will see a ‘proper’ definition in terms of power series in Analysis 2 (or perhaps Analysis 3!).

Is the set $\mathbb{N}$ of all natural numbers bounded above? 'Of course not' – but given the 'straightforward' rules of arithmetic and order we have introduced on page v there is no way to prove that this is not true[2]. We need another property of the real numbers, the *Least Upper Bound Axiom*, which will also enable us to 'fill it the gaps' and show that there is a real number whose square is 2.

As remarked above, if a set has an upper bound then it has many upper bounds: the 'best' upper bound would be the smallest possible, which we call the *least upper bound*, or *supremum*. A real number $l$ is the least upper bound for a set $S$ if

(i) it is an upper bound for $S$ ($s \leq l$ for every $s \in S$) and

(ii) if $L$ is any upper bound for $S$, $l \leq L$ (so $l$ is the least upper bound).

Note that there can only be one least upper bound: if $l_1$ and $l_2$ are two least upper bounds, then by property (ii) we must have $l_1 \leq l_2$ ($l_1$ is a least upper bound and $l_2$ is an upper bound) and $l_2 \leq l_1$ (vice versa). Given this uniqueness, we can write $l = \sup S$.

Before we give some examples we make an observation that will be extremely useful.

**Lemma 1.3.** *We have $l = \sup S$ if and only if $l$ is an upper bound for $S$ and for every $t < l$ then there exists $s \in S$ with $s > t$.*

*In particular, $l = \sup S$ if and only if $l$ is an upper bound for $S$ and for every $\varepsilon > 0$ there exists $s \in S$ with*

$$l - \varepsilon < s \leq l.$$

*Proof.* For the first part, suppose that $l = \sup S$, $t < l$, and there is no $s \in S$ with $s > t$. Then every $s \in S$ must satisfy $s \leq t$, and so $t$ is an upper bound for $S$ that is strictly less than $l$, which contradicts property (ii).

Now suppose that $l$ is an upper bound and for every $t < l$ there exists $s \in S$ with $s > t$. Then no $t < l$ can be an upper bound for $S$, so $l$ is the least upper bound.

For the second part, note that every $s \in S$ satisfies $s \leq l$ (because $l$ is an upper bound) and if we take $t = l - \epsilon$ we have just shown that there exists $s \in S$ with $s > t$. $\square$

2 How do you show that something *cannot* be proved given a collection of assumptions? Such questions of mathematical logic tend to be very difficult, and in general constructing counterexamples is hard. In this case, though, the ordered field of rational functions satisfies all the axioms on page v, and contains a 'copy' of the natural numbers as the constants $n$ (just as with the reals, $\mathbb{N}$ is what you get by adding the multiplicative identity to itself repeatedly. But in this example we have $x > n$ for every $n$, so in this case $\mathbb{N}$ *is* 'bounded above'.

Similarly, a lower bound for a set $S$ is a real number $r$ such that

$$s \geq r \qquad \text{for each } s \in S.$$

A set with a lower bound is *bounded below.* The *greatest lower bound* or *infimum* $l$ for a set $S$ is such that

$$l \leq s \text{ for each } s \in S \qquad \text{and} \qquad \text{if } t \text{ is a lower bound for } S \text{ then } t \leq l.$$

Arguing as in Lemma 1.3, if $l = \inf S$ then for every $\varepsilon > 0$ there exists $s \in S$ with $l \leq s < l + \varepsilon$.

Note that we can switch between least upper bounds and greatest lower bounds by considering the set $-S$ (see Examples 1).

If a set has a supremum, the supremum may or not be in the set: if the supremum is contained in the set then we call it the maximum (or minimum in the case of an infimum that is contained in the set). We look first at the case of intervals, and then revisit examples (i)–(v) from abov.

Both the open interval $(a, b)$ and the closed interval $[a, b]$ have infimum $a$ and supremum $b$. These are not contained in the open interval (so $(a, b)$ does not have a maximum or minimum element), but are elements of the closed interval (so $[a, b]$ has minimum $a$ and maximum $b$). In a little more detail, let's consider the supremum: we know that $b$ is an upper bound for both intervals, and if we take any $t < b$ then there exists $x \in (a, b)$ with $x > b$ (take $x = (t + b)/2$), so there is no smaller

(i) the supremum of the set $S = \{1/n : \ n \in \mathbb{N}\}$ is 1 (and $1 \in S$) and the infimum is 0 (and $0 \notin S$): 0 is a lower bound, while if we take any $l > 0$, there exists[3] $n$ such that $1/n < l$ and $l$ is not a lower bound.

(ii) the supremum of the set $S = \{\sin n\}$ is 1 and the infimum is $-1$ (this is not at all obvious: we will prove it later in Section X), but neither are in $S$. [We only have $|\sin \theta| = 1$ when $\theta = k\pi + (\pi/2)$, and since $\pi$ is irrational (see Analysis 2) the right-hand side here is never an integer.]

(iii) the infimum of $(-1, \infty)$ is $-1$ (which is not in $S$) and it has no supremum, since it is not bounded above.

(iv) the infimum of $\{1/2, 2/3, \cdots, n/(n + 1) : \ n \in \mathbb{N}\}$ is $1/2$ (which is in the set) and the supremum is 1 (which is not in the set). [1 is an upper bound, and for any $l < 1$ we can find $n$ such that $n/(n + 1) > l$.]

3 Careful! Along with example (iv) this actually uses the Archimedean Property of the real numbers ($\mathbb{N}$ is not bounded above in $\mathbb{R}$), so strictly we should delay this 'simple' example).

(v) the infimum of the set $E := \{(1+\frac{1}{n})^n : \; n = 1, 2, \ldots\}$ is 2: we will prove later that the sequence $x_n$ defined by setting $x_n = (1+\frac{1}{n})^n$ is increasing (i.e. $x_{n+1} \geq x_n$), and $x_1 = 2$. We will also see that the supremum is e (and that this is not contained in $E$).

**Least upper bound axiom.** *Any non-empty set of real numbers that is bounded above has a least upper bound.*

This is the one formulation of the so-called 'Completeness Axiom' for the real numbers[4]. A second version is almost immediate (given a non-empty set $S$ that is bounded below, consider $-S$ and apply the LUB Axiom), and equivalent (see Examples 1).

**Greatest lower bound axiom.** *Any non-empty set of real numbers that is bounded below has a greatest lower bound.*

## 1.3 There is a real number whose square is 2.

Using the Least Upper Bound Axiom we can show that '$\sqrt{2}$' is in fact a real number. In the proof we will use the observation that for any $x > 0$,

$$(1+x)^2 = 1 + 2x + x^2 \geq 1 + 2x.$$

[This is a particularly easy case of Bernoulli's inequality

$$(1+x)^k \geq 1 + kx, \qquad x > -1,$$

which can be proved by induction (see Examples 1).]

**Proposition 1.4.** *There is a unique positive real number $r$ such that $r^2 = 2$.*

*Proof.* We consider the set

$$S = \{x \in \mathbb{R} : \; x^2 < 2\}.$$

This set is non-empty, since $1 \in S$, and is bounded above since if $x > 2$ we have $x^2 > 4$, so $x \notin S$: it follows that every $x \in S$ satisfies $x \leq 2$.

By the Completeness Axiom the set $S$ has a least upper bound; let's set

$$r := \sup\{x \in \mathbb{R} : \; x^2 < 2\}.$$

There are three possibilities: (i) $r^2 < 2$; (ii) $r^2 = 2$; (iii) $r^2 > 2$. If we can exclude (i) and (iii) then all that is left is (ii), that $r^2 = 2$.

4 In a way that can be made precise, the real numbers are the unique ordered field that satisfies the Completeness Axiom, in whichever of the many equivalent forms we assume it.

So suppose case (i) holds and $r^2 < 2$. We show that there is a positive $\varepsilon$ such that $r + \varepsilon \in S$, which means that $r$ is not an upper bound for $S$.

Consider
$$(r + \varepsilon)^2 = r^2 + 2\varepsilon r + \varepsilon^2.$$
If $\varepsilon < 1$ then $\varepsilon^2 < \varepsilon$, so in this case
$$(r + \varepsilon)^2 < r^2 + 2\varepsilon r + \varepsilon = r^2 + \varepsilon(2r + 1).$$
Now if we choose
$$0 < \varepsilon < \frac{2 - r^2}{2r + 1},$$
then
$$(r + \varepsilon)^2 < r^2 + \frac{2 - r^2}{2r + 1}(2r + 1) = 2,$$
so $r + \varepsilon \in S$. But $r + \varepsilon > r$, so $r$ cannot be an upper bound for $S$: possibility (i) cannot be true.

Now suppose that case (ii) holds and $r^2 > 2$. We will show that although $r$ is an upper bound for $S$, it is not the *least* upper bound, because $r - \varepsilon$ is also an upper bound (for some positive $\varepsilon$).

This time we consider
$$(r - \varepsilon)^2 = r^2 - 2r\varepsilon + \varepsilon^2 > r^2 - 2r\varepsilon.$$
If we take $0 < \varepsilon < (r^2 - 2)/2r$ then
$$(r - \varepsilon)^2 > 2.$$
It follows that $r - \varepsilon$ is an upper bound for $S$, since for any $x \in S$ we have
$$x^2 < 2 < (r - \varepsilon)^2 :$$
this implies that $x < r - \varepsilon$. But this contradicts the fact that $r$ was supposed to be the least upper bound, so case (iii) cannot hold.

The only possibility remaining is case (ii): $r^2 = 2$.

To show that there is no other real number with this property, suppose that $s^2 = 2$. Then either (i) $s < r$, (ii) $s = r$, or (iii) $s > r$. If $s < r$ then $s^2 < r^2 = 2$; if $s > r$ then $s^2 > r^2 = 2$; so we must have $s = r$. □

There is nothing special about the number 2 here, in either place it occurs: we could use a similar argument to prove that there is a unique positive number $y$ such that $y^q = x$ for any $q \in \mathbb{N}$ and any positive real number $x$: see Examples 1.

If $x > 0$, the unique positive $y$ for which $y^q = x$ we write as $x^{1/q}$. We can now define rational powers of any positive real number: if $\alpha = p/q$ with $p \in \mathbb{Z}$, $q \in \mathbb{N}$, we set

$$x^\alpha = (x^p)^{1/q}\,. \tag{1.2}$$

Note that we don't yet have a way of defining irrational powers, e.g. what does $2^\pi$ mean? (We will return to this briefly later, but you will see the 'correct' answer in Analysis 2.)

## *1.4 Rational powers

We need to check that if we define rational powers as in (1.2) then this definition obeys all the rules we are used to: this involves checking that two expressions involving integer powers or integer roots coincide. (This section is not starred because it is 'interesting and advanced'; it is starred because it is largely tedious.)

We showed earlier that $x^k = y^k$ if and only if $x = y$, and we will use this repeatedly: if $a$ and $b$ are both positive, to show that $a = b$ we can take powers (since $a^k = b^k$ implies that $a = b$) or roots (since $a = b$ implies that $a^k = b^k$) to try to end up with an obvious equality.

The rules we know for integer powers come essentially from what we know about addition and multiplication: if $x, y > 0$ and $n, m \in \mathbb{Z}$ then

$$x^{n+m} = x^n x^m, \qquad (x^n)^m = x^{nm} = (x^m)^n, \qquad (xy)^n = x^n y^n.$$

[Given that $x^0 = 1$, the first of these gives the rule $x^{-n} = 1/x^n$.] We want to extend these to rational powers, when these are defined as in (1.2), given our new rule

$$(x^{1/q})^q = x,$$

which is in fact our definition of what $x^{1/q}$ means.

- First we have to check that however we choose to write the rational $\alpha$, we get the same value for $x^\alpha$ using the definition in (1.2), i.e. that $x^\alpha$ is 'well defined'. We need to check that if $\alpha = p/q = p'/q'$ then

  $$(x^p)^{1/q} = (x^{p'})^{1/q'}.$$

  If we take the $qq'$ power of both sides we get

  $$(x^p)^{q'} = (x^{p'})^q \qquad \text{which is} \qquad x^{pq'} = x^{p'q},$$

  and we have $pq' = p'q$ since $p/q = p'/q'$, so LHS=RHS as required.

- We want to extend the rules for integer powers to rational powers. Take $\alpha = p/q$ and $\beta = r/s$, so then, since $\alpha + \beta = (ps + qr)/qs$, we want to show that

$$(x^p)^{1/q}(x^r)^{1/s} = (x^{ps+qr})^{1/qs}.$$

Take the $qs$ power of the LHS we get

$$[(x^p)^{1/q}]^{qs}[(x^r)^{1/s}]^{qs} = (x^p)^s(x^r)^q = x^{ps}x^{rq} = x^{ps+qr},$$

which is the $qs$ power of the RHS.

- Take $\alpha = p/q$ and $\beta = r/s$; then $\alpha\beta = pr/qs$. We want to show that $(x^\alpha)^\beta = x^{\alpha\beta}$, which using or definition means we need to show that

$$\{[(x^p)^{1/q}]^r\}^{1/s} = (x^{pr})^{1/qs}.$$

If we take the $qs$ power of the LHS then, since $y^{qs} = (y^s)^q$, we get

$$(\{[(x^p)^{1/q}]^r\}^{1/s})^{qs} = \{[(x^p)^{1/q}]^r\}^q = \{[(x^p)^{1/q}]^q\}^r = (x^p)^r = x^{pr},$$

which is the same as the $qs$ power of the RHS.

- If $x, y > 0$ and $\alpha = p/q$, to check that $(xy)^\alpha = x^\alpha y^\alpha$ we need to check that

$$(x^p)^{1/q}(y^p)^{1/q} = [(xy)^p]^{1/q}.$$

If we raise the LHS to the power of $q$ we obtain $x^p y^p = (xy)^p$, which is what we get raising the RHS to the power of $q$.

- Before we prove our final property, we show that $(x^{1/q})^p = (x^p)^{1/q}$, so we could have defined $x^{p/q}$ by taking the $q$th root before the $p$th power and got the same answer. Raising the LHS to the power $q$ we get

$$[(x^{1/q})^p]^q = [(x^{1/q})^q]^p = x^p$$

and on the RHS we get

$$[(x^p)^{1/q}]^q = x^p.$$

- And finally, we want to be able to take roots of rational powers: we want to show that $(x^{1/\alpha})^\alpha = x$; when $\alpha = p/q$ the left-hand side is

$$\{[(x^q)^{1/p}]^p\}^{1/q} = \{x^q\}^{1/q} = x,$$

using the 'either way' result we just proved.

## 1.5 Rational and irrational numbers are interleaved

We have shown that the real numbers contain at least one irrational number ($\sqrt{2}$). In fact there are many: given any irrational number $x$ and any rational number $p/q$, $x + (p/q)$ and $xp/q$ are both irrational. To see this, we argue by contradiction: if

$$x + \frac{p}{q} = \frac{p'}{q'}, \qquad \text{then} \qquad x = \frac{p'}{q'} - \frac{p}{q} = \frac{p'q - pq'}{qq'} \in \mathbb{Q},$$

contradicting the fact that $x$ is irrational; and similarly if

$$x\frac{p}{q} = \frac{p'}{q'}, \qquad \text{then} \qquad x = \frac{p'q}{pq'} \in \mathbb{Q}.$$

To investigate a little more how rational and irrational numbers are intertwined we will use the following surprising consequence of the Least Upper Bound Axiom: $\mathbb{N}$ is not bounded above in $\mathbb{R}$. [The result is not surprising - but it is surprising that its proof uses the Least Upper Bound Axiom.]

The following theorem offers a more useful formulation of this fact.

**Theorem 1.5** (Archimedean property of the real numbers)**.** *For any real number $r$, there exists $n \in \mathbb{N}$ such that $n > r$.*

*Proof.* If this does not hold then there is some real number $r$ such that $n < r$ for all $n \in \mathbb{N}$, i.e. the set $\mathbb{N}$ of all natural numbers would be bounded above. In this case, as a non-empty subset of the real numbers that is bounded above, it would have a least upper bound $R$.

Given any $n \in \mathbb{N}$, $n + 1$ is also a natural number, so for every natural number $n$,

$$n + 1 \leq R.$$

But this implies that $n \leq R - 1$ for every natural number $n$. So $R - 1$ is an upper bound for $\mathbb{N}$, contradicting the fact that $R$ is the least upper bound. □

This means, in particular, that are no 'infinitesimal' real numbers.

**Lemma 1.6.** *Suppose that $x \in \mathbb{R}$ with $x \geq 0$ and $x < 1/n$ for every $n \in \mathbb{N}$. Then $x = 0$.*

*Proof.* Suppose that $x \neq 0$. Then $x > 0$, so using the Archimedean property, we can find $m \in \mathbb{N}$ such that $m > 1/x$. Then $1/m < x$, contradicting our hypothesis. So we must have $x = 0$. initial assumption. So $x = 0$. □

We often apply this when we have shown that $0 \le x < \varepsilon$ for every $\varepsilon > 0$ to deduce that $x = 0$. Note that the proof contains a key argument: given any real number $x > 0$, we can find $n \in \mathbb{N}$ such that $1/n < x$.

**Lemma 1.7.** *Between any two real numbers $a < b$ there is a rational number and an irrational number.*

*Proof.* First we find a rational number between $a$ and $b$.

If $a < 0 < b$ then we can take $x = 0$. If the result holds for $0 \le a < b$ then it also holds for $a < b \le 0$ (find $x$ with $-b < x < -a$ and then take $-x$ with $a < -x < b$).

So we take $0 \le a < b$ with $a, b \in \mathbb{R}$. Choose $n \in \mathbb{N}$ such that $1/n < b - a$, and let $m$ be the smallest integer[5] such that $m > an$. Then

$$m - 1 \le an \qquad \text{and} \qquad an < m,$$

which means that

$$a < \frac{m}{n} \le a + \frac{1}{n} < a + (b - a) = b,$$

since we choose $n$ to satisfy $1/n < b - a$. So $m/n$ is a rational number that lies between $a$ and $b$.

Now we show that there is an irrational number between $a$ and $b$. We use the previous result to find a rational number $x$ such that

$$\frac{a}{\sqrt{2}} < x < \frac{b}{\sqrt{2}}.$$

We showed above that if $x$ is rational and $y$ is irrational then $xy$ is irrational, so $x\sqrt{2}$ is irrational and $a < x\sqrt{2} < b$, as required. □

**Corollary 1.8.** *Any real number can be approximated arbitrarily closely by a rational number: if $r$ is a real number, then for any $\varepsilon > 0$, there exists a rational number $x$ such that*

$$r - \varepsilon < x < r + \varepsilon. \tag{1.3}$$

*Proof.* Take $r \in \mathbb{R}$. There is only something to prove if $r$ is irrational. Take any $\varepsilon > 0$ and consider $a = r - \varepsilon$, $b = r + \varepsilon$. By Lemma 1.7 there is a rational number $x$ with $r - \varepsilon = a < x < b = r + \varepsilon$. □

5 In fact the existence of such an $n$ relies on the Greatest Lower Bound Property (or, if you prefer, the Well Ordering Principle from Examples 1/Foundations). Let

$$S = \{k \in \mathbb{N} : \ k > an\}$$

This set is a non-empty subset of $\mathbb{R}$ (by the Archimedean property) and is bounded below by $an$. So we can set $r = \inf S$. Since $an$ is a lower bound, $r \ge an$. Since $r$ is the greatest lower bound, there exists $m \in S$ such that $m < r + 1$. This implies that $m - 1 < r$, so $m - 1 \notin S$ ($r$ is a lower bound for $S$). So $m - 1 \le an$. We have found $m$ with $an < m \le an + 1$.

## 1.6 The triangle inequality

We can define the absolute value function as follows: for any real number $x \in \mathbb{R}$ we write

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases} \qquad \text{or, alternatively,} \quad |x| = \max(x, -x).$$

Note that $|-x| = |x|$; that $|xy| = |x||y|$ and $|x/y| = |x|/|y|$ (when $y \neq 0$).

Most importantly for what follows, observe that $|x| < r$ means that $-r < x < r$. So if $|x - a| < \varepsilon$ then the distance between $x$ and $a$ is strictly less than $\varepsilon$ (and this is the way to think of the notation): we have

$$-\varepsilon < x - a < \varepsilon$$

and so

$$a - \varepsilon < x < a + \varepsilon.$$

For example, we could write (1.3) much more concisely and clearly as $|x - r| < \varepsilon$.

**Lemma 1.9** (Triangle inequality)**.** *For any two real numbers $a$ and $b$,*

*(i)* $|a + b| \leq |a| + |b|$ *(triangle inequality);*

*(ii)* $|a - b| \geq ||a| - |b||$ *(reverse triangle inequality).*

*Proof.* (i) For any $a \in \mathbb{R}$ we have $-|a| \leq a \leq |a|$. We have the same for $b$, $-|b| \leq b \leq |b|$. It follows that

$$-|a| - |b| \leq a + b \leq |a| + |b|;$$

since the left-hand side is $-(|a| + |b|)$, this shows that $|a + b| \leq |a| + |b|$ as required.

(There are other proofs: a tedious case-by-case analysis, or one can consider $|a + b|^2 = (a + b)^2$.)

(ii) We apply the triangle inequality twice. First to $(a - b) + b$, to give

$$|a| = |(a - b) + b| \leq |a - b| + |b| \qquad \Rightarrow \qquad |a| - |b| \leq |a - b|;$$

and then to $(b - a) + a$ to give

$$|b| = |(b - a) + a| \leq |b - a| + |a| \qquad \Rightarrow \qquad |b| - |a| \leq |a - b|$$

(since $|a - b| = |b - a|$). So

$$|a - b| \geq \max(|a| - |b|, |b| - |a|) = ||a| - |b||\,. \qquad \square$$

You can easily get two more inequalities by replacing $b$ with $-b$ in either (i) or (ii):

$$|a-b| \leq |a| + |b| \qquad \text{and} \qquad |a+b| \geq ||a| - |b||\,.$$

[You only need to remember (i) and (ii).]

Why the 'triangle inequality'? If we put $a = x - z$ and $b = z - y$ in the triangle inequality $|a + b| \leq |a| + |b|$ then we obtain

$$|x-y| = |(x-z) + (z-y)| \leq |x-z| + |z-y|;$$

if you imagine a triangle in the plane with corners at $x$, $y$, and $z$ then this is what you would expect. (This is the reason for the terminology, but isn't an ideal explanation, since at the moment we're only using $|\cdot|$ to measure the distance between real numbers, not points on the plane.)

## *1.7 'Good approximation' by rationals

In the next result we will use the integer part (the 'floor') of a rational number: we write

$$\lfloor x \rfloor = \text{largest integer } n \text{ such that } n \leq x;$$

note that

$$x - 1 < \lfloor x \rfloor \leq x. \tag{1.4}$$

See Examples 1 for a proof that this notion is well defined.

**Theorem 1.10** (Dirichlet Approximation Theorem)**.** *Choose any $\alpha \in \mathbb{R}$. Then there are infinitely many choices $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}. \tag{1.5}$$

*Proof.* Note that if $\alpha \in \mathbb{Q}$ with $\alpha = a/b$ then we can take $p = na$, $q = nb$ for any $n \in \mathbb{N}$ and the left-hand side in (1.5) is zero, so the theorem is trivially true.

So we can take $\alpha$ to be irrational. Multiplying (1.5) by $q$, it is enough to show that there are infinitely many choices of $p, q$ such that

$$|q\alpha - p| < \frac{1}{q}. \tag{1.6}$$

We use the 'pigeonhole principle'. Fix $N \geq 1$, and consider the fractional parts of the numbers $0, \alpha, \ldots, N\alpha$ (i.e. these numbers modulo 1): this must give $N + 1$ distinct numbers. To see this, observe that if not then we can find integers $m$ and $n$ with $0 \leq m < n \leq N$ such that

$$n\alpha - \lfloor n\alpha \rfloor = m\alpha - \lfloor m\alpha \rfloor,$$

and so

$$(m-n)\alpha = \lfloor m\alpha \rfloor - \lfloor n\alpha \rfloor \qquad \Rightarrow \qquad \alpha = \frac{\lfloor m\alpha \rfloor - \lfloor n\alpha \rfloor}{m-n},$$

which contradicts the fact that $\alpha$ is irrational.

Now divide the interval $[0, 1)$ into $N$ equal parts,

$$\left[0, \frac{1}{N}\right), \quad \left[\frac{1}{N}, \frac{2}{N}\right), \quad \cdots, \quad \left[\frac{N-1}{N}, 1\right);$$

then there must be one of these $N$ intervals that contains at least two of the $N + 1$ 'fractional parts' we just found.

That means that for some $m, n \in \{0, \ldots, N\}$, $m > n$ the factional parts of $m\alpha$ and $n\alpha$ are no more than $1/N$ apart. It follows that there is an integer $p$ such that

$$|(m\alpha - n\alpha) - p| = |\alpha(m-n) - p| < \frac{1}{N}.$$

Noting that $q := m - n \in \mathbb{N}$ with $0 < m - n \leq N$ we have

$$|q\alpha - p| < \frac{1}{N} \leq \frac{1}{q}.$$

This gives one choice $(p_1, q_1)$ that satisfies (1.6) and hence (1.5). Now if we start again but choose a larger value of $N_2$ such that $1/N_2 < |q_1\alpha - p_1|$ we can follow exactly the same argument until we end up with $(p_2, q_2)$ such that

$$|q_2\alpha - p_2| < \frac{1}{N_2} \leq \frac{1}{q_2}.$$

Note that $(p_2, q_2)$ is different from $(p_1, q_1)$, since $\frac{1}{N_2} < |q_1\alpha - p_1|$.

We can continue in this way indefinitely to find infinitely many different values of $(p, q)$ that satisfy (1.5). $\square$

# Part II

# Sequences and convergence

# Chapter 2

# Sequences and convergence

A sequence is an infinite ordered list of real numbers:

$$a_1, a_2, a_3, a_4, \ldots.$$

We use the notation $(a_n)$ or $(a_n)_{n=1}^{\infty}$ for a sequence. [This distinguishes a sequence, which requires us to know the order of the elements, from the set $\{a_n\}_{n=1}^{\infty}$, which is just a collection of numbers. But you may also find the notation $\{a_n\}$ for a sequence in some textbooks.]

Particular sequences can be specified by a formula, a 'recipe' (an involved construction), or even a picture.

Examples: (i) specified by an explicit formula: $a_n = 1/n$; $a_n = (1 + \frac{1}{n})^n$, $x_n = \sin n$, $z_n = 2^{-n} + (-1)^n 3^{-n/2}$, ...

(ii) specified by a recipe: $p_n =$ the $n$th prime number (most easily specified this way, but in fact there is also an impractical formula for the $n$th prime[1]); $f_1 = f_2 = 1$, $f_{n+1} = f_n + f_{n-1}$ (so $f_n$ is the $n$th Fibonnacci number; there is a formula for $f_n$ too); $\pi_n = n$th digit in the decimal expansion of $\pi$; $r_1 = x$, $r_{n+1} = 1 - 1/r_n$.

(iii) indicated by a sketch: if we think of a sequence as a function from $\mathbb{N}$ to $\mathbb{R}$ then we can sketch its graph, which is often a useful way to think about definitions/proofs involving sequences.

1 The formula is due to Willans, see `https://www.youtube.com/watch?v=j5s0h42GfvM`; we have

$$p_n = 1 + \sum_{i=1}^{2^n} \left\lfloor \left( \frac{n}{\sum_{j=1}^{i} \left\lfloor \left( \cos \pi \frac{(j-1)!+1}{j} \right)^2 \right\rfloor} \right)^{1/n} \right\rfloor.$$

FIGURE HERE

**Definition 2.1.** A sequence $(a_n)$ converges to a limit $l \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|a_n - l| < \varepsilon \qquad \text{for every } n \geq N. \tag{2.1}$$

In this case we write $a_n \to l$ as $n \to \infty$.

[When it's clear we will often omit the 'as $n \to \infty$' and just write $a_n \to l$. The assumption that $l \in \mathbb{R}$ means that we cannot take '$l = \infty$' here - we will look at this case later.]

We could also write (2.1) as

$$n \geq N \qquad \Rightarrow \qquad |a_n - l| < \varepsilon. \tag{2.1$'$}$$

CONVERGENCE FIGURE

The definition says that you will be as close as you want to $l$ if you wait long enough. The symbols $\forall$ ('for all' - which means 'for each') and $\exists$ ('there exists') can be useful to make some statements look shorter [and they are much quicker to write on the blackboard], but be careful! You don't want to end up playing with symbols without understanding what is really going on. Still, the following might help you to remember the convergence definition: the English (while a little clumsy) is the key, not the upside-down A and backwards E.

$$\underbrace{\forall\, \varepsilon > 0}_{\text{no matter how close you want to get to } l,} \qquad \underbrace{\exists\, N \text{ s.t } n \geq N}_{\text{if you wait long enough,}} \quad \Rightarrow \quad \underbrace{|a_n - l| < \varepsilon.}_{\text{you will be there}}$$

To show that a sequence converges, you have to show that for any $\varepsilon > 0$ you can find an appropriate $N$ so that (2.1) [or (2.1')] holds.

## 2.1 Some simple examples of convergent sequences

• A constant sequence. Suppose that $a_n = l$ for every $n \in \mathbb{N}$. Then $a_n \to l$: given any $\varepsilon > 0$, choose your favourite nartual number $N$; then for all $n \geq N$,

$$|a_n - l| = |l - l| = 0 < \varepsilon.$$

**Lemma 2.2.**

$$\frac{1}{n} \to 0 \quad as \quad n \to \infty.$$

*Proof.* Take $\varepsilon > 0$: we need to show that we can find $N$ such that $n \geq N$ implies that

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

If we take $N > 1/\varepsilon$ (we can do this by the Archimedean Property) then for $n \geq N$ we have

$$\frac{1}{n} \leq \frac{1}{N} < \varepsilon. \qquad \square$$

Simple exercise: show that for any $\alpha > 0$, $\alpha \in \mathbb{Q}$, $1/n^\alpha \to 0$ as $n \to \infty$.

**Lemma 2.3.** *If* $|x| < 1$ *then*

$$x^n \to 0 \quad as \quad n \to \infty.$$

*Proof.* Take $\varepsilon > 0$; we need to find an $N$ such that $n \geq N$ implies that

$$|x^n - 0| = |x^n| = |x|^n < \varepsilon.$$

We only need to consider $0 < x < 1$.

Let's write $x = \frac{1}{1+h}$ $(h = (1/x) - 1 > 0)$. Then

$$x^n = \left(\frac{1}{1+h}\right)^n = \frac{1}{(1+h)^n}.$$

Using the inequality $(1+h)^n \geq 1 + nh$ it follows that

$$x^n \leq \frac{1}{1+nh}.$$

Now choose $N$ sufficiently large so that

$$\frac{1}{1+Nh} < \varepsilon$$

(any $N$ with $N > \frac{1}{n}\left(\frac{1}{\varepsilon} - 1\right)$ will do); then for any $n \geq N$ we have

$$|x^n - 0| \leq \frac{1}{1+nh} \leq \frac{1}{1+Nh} < \varepsilon. \qquad \square$$

There are, of course, sequences that don't converge, e.g. $a_n = (-1)^n$ (see Examples 1); $a_n = \sin n$. [That the first doesn't converge is clear, that the second doesn't converge is not surprising, but is not quite so easy to show.]

## 2.2 Some properties of limits

First, we check that a sequence can have only one limit.

**Lemma 2.4** (Uniqueness of limits)**.** *A sequence can have at most one limit.*

This means that we can write $\lim_{n\to\infty} a_n = l$ if $a_n \to l$.

*Proof.* We suppose that $a_n \to l_1$ and $a_n \to l_2$ and show that we must have $l_1 = l_2$.

Proof 1. Assume that $l_1 \neq l_2$ and deduce a contradiction.

Suppose that $l_2 > l_1$. Take $\varepsilon = (l_2 - l_1)/2$. Since $a_{\to} l_1$, there exists $N_1$ such that

$$|a_n - l_1| < \varepsilon \qquad \text{for every } n \geq N_1;$$

in particular, $a_n > l_1 - (l_2 - l_1)/2 = (l_1 + l_2)/2$ for all $n \geq N_1$.

Since $a_n \to l_2$, there exists $N_2$ such that

$$|a_n - l_2| < \varepsilon \qquad \text{for every } n \geq N_2;$$

in particular, $a_n < l_2 + (l_2 - l_1)/2 = (l_1 + l_2)/2$ for all $n \geq N_2$.

If we take $n \geq \max(N_1, N_2)$ it follows that we must have both $a_n > (l_1 + l_2)/2$ and $a_n < (l_1 + l_2)/2$, which is impossible; so we must have $l_1 = l_2$.

Proof 2. Direct proof: show that $|l_1 - l_2| < \varepsilon$ for all $\varepsilon > 0$.

Pick $\varepsilon > 0$. Then, since $a_n \to l_1$, there exists $N_1$ such that

$$|a_n - l_1| < \varepsilon \qquad \text{for every } n \geq N_1.$$

Since $a_n \to l_2$, there exists $N_2$ such that

$$|a_n - l_2| < \varepsilon \qquad \text{for every } n \geq N_2.$$

Now, if we choose some $n > \max(N_1, N_2)$ then we have

$$|a_n - l_1| < \varepsilon \qquad \text{and} \qquad |a_n - l_2| < \varepsilon.$$

Using the triangle inequality,

$$|l_1 - l_2| \leq |l_1 - a_n| + |a_n - l_2| < \varepsilon + \varepsilon = 2\varepsilon$$

We have shown that for any $\varepsilon > 0$ we have $|l_1 - l_2| < 2\varepsilon$, which implies (by Lemma 1.6) that $l_1 = l_2$. □

Both of the proofs above use the key idea that we can find various ‘threshold values’ (here $N_1$ and $N_2$) so that something we want happens if $n \geq N_1$ and if $n \geq N_2$. We then take $n \geq \max(N_1, N_2)$ to ensure that for this range of $n$ both things happen at the same time. We will use this same technique repeatedly.

——

We have various ways of saying that a sequence converges. The following statements mean the same thing:

- the sequence $(a_n)$ converges/tends to $l$ [as $n \to \infty$];
- $a_n$ converges/tends to $l$ [as $n \to \infty$];
- $a_n \to l$ [as $n \to \infty$];
- $\lim_{n\to\infty} a_n = l$.

We often omit ‘as $n \to \infty$’ if it is obvious from to context. But sometimes it is required, e.g. taking $k \to \infty$ and $n \to \infty$ in $k^{-n}$ describe different limits (are they always the same?).

We can write (as above) things like $1/n \to 0$; we do not have to recast things by defining $a_n = 1/n$ and then saying that $a_n \to 0$.

——

We now prove some more basic properties of convergent sequences.

**Lemma 2.5** (Shift rule)**.** *For any fixed $k \in \mathbb{N}$, $a_n \to l$ as $n \to \infty$ if and only if $a_{n+k} \to l$ as $n \to \infty$.*

*Proof.* Given $\varepsilon > 0$, choose $N$ such that $|a_n - l| < \varepsilon$ for all $n \geq N$. Then, if we take $n > N$, we have $n + k > n > N$, so $|a_{n+k} - l| < \varepsilon$ also. Conversely, given $\varepsilon > 0$, find $M$ such that

$$|a_{m+k} - l| < \varepsilon$$

for all $m \geq M$. Then if we take $N = M + k$ we have

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

for all $n \geq N$. ∎

Examples: $1/(10 + n) \to 0$ as $n \to \infty$; if $|x| < 1$ then $|x|^{n-5} \to 0$ as $n \to \infty$.

We say that a sequence $(a_n)$ is bounded if the set $\{a_n : n \in \mathbb{N}\}$ is bounded, i.e. there exist $A \in \mathbb{R}$, $A > 0$, such that $|a_n| \leq A$ for every $n \in \mathbb{N}$ (so the set is bounded above by $A$ and below by $-A$).

**Lemma 2.6.** *Any convergent sequence is bounded.*

*Proof.* Take a sequence $(a_n)$ such that $a_n \to a$. Then, taking $\varepsilon = 1$ in the definition of convergence, there exists $N$ such that

$$|a_n - a| < 1 \qquad n \geq N.$$

So we have

$$|a_n| = |(a_n - a) + a| \leq 1 + |a| \qquad n \geq N.$$

Set

$$A = \max(|a_1|, \dots, |a_{N-1}|, |a| + 1)$$

and then $|a_n| \leq A$ for all $n \in \mathbb{N}$. □

This gives one quick test for *non-convergence*. Note that a sequence is unbounded (= not bounded) if for every $A > 0$ there exists $n \in \mathbb{N}$ such that $|a_n| > A$. (However large you choose $A$, the sequence contains a term with modulus larger than $A$.) E.g. the sequence $(n)$ does not converge.

**Lemma 2.7.** *If $a_n \to a$ then $|a_n| \to |a|$.*

*Proof.* Using the reverse triangle inequality

$$||a_n| - |a|| \leq |a_n - a|.$$

Given any $\varepsilon > 0$ choose $N$ such that $|a_n - a| < \varepsilon$ for all $n \geq N$; then $||a_n| - |a|| < \varepsilon$ for all $n \geq N$ too. □

The following result is on Examples 1.

**Lemma 2.8.** *Suppose that $|a_n - l| \leq |b_n|$, where $b_n \to 0$ as $n \to \infty$. Then $a_n \to l$.*

### 2.2.1 Algebra of limits

We don't want to have to calculate limits afresh every time; we have a natural 'algebra of limits' that can help.

**Lemma 2.9.** *Suppose that $a_n \to a$ and $b_n \to b$. Then*

*(i)* $a_n + b_n \to a + b$;

*(ii)* $a_n b_n \to ab$;

*(iii)* *if* $b \neq 0$ *then* $a_n/b_n \to a/b$.

*Proof.* (i) First, notice that

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|.$$

We choose $N$ so that for $n \geq N$, both $|a_n - a| < \varepsilon/2$ and $|b_n - b| < \varepsilon/2$ are small.

Pick $\varepsilon > 0$. Then, since $a_n \to a$, there exits $N_1$ such that

$$|a_n - a| < \varepsilon/2 \qquad \text{for every } n \geq N_1;$$

since $b_n \to b$, there exists $N_2$ such that

$$|b_n - b| < \varepsilon/2 \qquad \text{for every } n \geq N_2.$$

So for every $n \geq N := \max(N_1, N_2)$ we have

$$|a_n - a| < \varepsilon/2 \qquad \text{and} \qquad |b_n - b| < \varepsilon/2.$$

So

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

We have shown that given any $\varepsilon > 0$, there exists $N$ such that $n \geq N$ implies that

$$|(a_n + b_n) - (a + b)| < \varepsilon;$$

so $a_n + b_n \to a + b$ as claimed.

(ii) First, notice that

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \\ &\leq |a_n b_n - a_n b| + |a_n b - ab| \\ &= |a_n||b_n - b| + |b||a_n - a|. \end{aligned}$$

Since $(a_n)$ converges, by Lemma 2.6, there exists $A > 0$ such that $|a_n| \leq A$ for all $n$.

We choose $N$ so that for $n \geq N$, we have $|b_n - b| < \varepsilon/2A$, and $|a_n - a| < \varepsilon/2(|b| + 1)$.

Since $a_n \to a$ we can find $N_1$ such that

$$|a_n - a| < \frac{\varepsilon}{2(|b| + 1)} \qquad \text{for all } n \geq N_1;$$

and since $b_n \to b$ we can find $N_2$ such that

$$|b_n - b| < \frac{\varepsilon}{2A} \qquad \text{for all } n \geq N_2.$$

If we take $n \geq N := \max(N_1, N_2)$ then we have

$$|a_n - a| < \frac{\varepsilon}{2(|b|+1)}, \qquad \text{and} \qquad |b_n - b| < \frac{\varepsilon}{2A},$$

and so for $n \geq N$

$$\begin{aligned} |a_n b_n - ab| &\leq |a_n||b_n - b| + |b||a_n - a| \\ &< A\frac{\varepsilon}{2A} + |b|\frac{\varepsilon}{2|b|+1} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

We have shown that given any $\varepsilon > 0$ there exists $N$ such that $n \geq N$ implies that

$$|a_n b_n - ab| < \varepsilon;$$

so $a_n b_n \to ab$ as claimed.

(iii) We will show that if $b_n \to b$ and $b \neq 0$ then $1/b_n \to 1/b$. Part (iii) of the theorem then follows using part (ii).

So we observe that
$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b_n - b}{b_n b}\right| = \frac{|b_n - b|}{|b_n||b|}.$$
Since $b \neq 0$, $|b| \neq 0$.

Choose $\varepsilon > 0$. We choose $N$ large enough so that for $n \geq N$ we have $|b_n| \geq |b|/2$ and $|b_n - b| < |b|^2\varepsilon/2$.

From the triangle inequality we have

$$|b| = |b - b_n + b_n| \leq |b - b_n| + |b_n|, \qquad \text{and so} \qquad |b_n| \geq |b| - |b - b_n|. \tag{2.2}$$

Since $b_n \to b$ there exists $N_1$ such that for all $n \geq N_1$ we have

$$|b_n - b| < \frac{|b|}{2},$$

and so for all $n \geq N_1$ from (2.2) we get

$$|b_n| \geq |b| - |b - b_n| > |b| - \frac{|b|}{2} = \frac{|b|}{2}.$$

Now, since $b_n \to b$ we can find $N_2$ such that for all $n \geq N_2$ we have

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

So for $n \geq N := \max(N_1, N_2)$ we have

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b_n||b|} < \frac{\varepsilon|b|^2}{2} \times \frac{1}{|b|^2/2} = \varepsilon.$$

So $1/b_n \to 1/b$. Now we use part (ii): $a_n/b_n = a_n(1/b_n) \to a(1/b) = a/b$. $\square$

Some examples

1. Find the limit as $n \to \infty$ of

$$a_n = \frac{n^3 - n}{2n^3 + n^2}.$$

Rewrite $a_n$ as

$$a_n = \frac{1 - \frac{1}{n^2}}{2 + \frac{1}{n}}.$$

Since $1/n^2 \to 0$ and $1/n \to 0$, by the addition rule the numerator converges to 1 and the denominator converges to 2. Since $2 \neq 0$, by the quotient rule the whole expression converges to $1/2$.

2. We use this on a more interesting example. The key point here is that sometimes, by assuming that a sequence does converge, we can find out what the limit should be, and then prove that the sequence does actually converge to this limit. We will see some similar examples later.

Suppose that $x_1 = 1$ and

$$x_{n+1} = \frac{x_n + 4}{x_n + 1}.$$

If $x_n$ converges to $s$ then we would have

$$s = \frac{s+4}{s+1} \qquad s(s+1) = s + 4 \qquad s^2 = 4 \qquad s = \pm 2.$$

Since the terms are positive, we guess that $x_n \to 2$.

Let's look at $x_n - 2$; then

$$x_{n+1} - 2 = \frac{x_n + 4}{x_n + 1} - 2 = \frac{x_n + 4 - 2x_n - 2}{x_n + 1} = \frac{2 - x_n}{x_n + 1} = -\frac{x_n - 2}{x_n + 1}.$$

We can prove that $x_n \geq 1$ for all $n$ by induction: true for $n = 1$. If $x_n \geq 1$ then

$$x_{n+1} = \frac{x_n + 4}{x_n + 1} \geq 1.$$

It follows that

$$|x_{n+1} - 2| = \frac{|x_n - 2|}{|x_n + 1|} \leq \frac{1}{2}|x_n - 2|$$

so

$$|x_{n+1} - 2| \leq 2^{-n}|x_1 - 2| = 2^{-n} = (1/2)^n.$$

Therefore $|x_n - 2| \to 0$ as $n \to \infty$, so $x_n \to 2$.

3. Show that $\sin n$ does not converge to any limit as $n \to \infty$.

First we suppose that $\sin n \to l$ with $l \neq 0$ as $n \to \infty$ and deduce a contradiction. If we use the trig identity

$$\sin(a + b) = \sin a \cos b + \sin b \cos a$$

then we have

$$\begin{aligned} \sin(n + 1) &= \sin n \cos 1 + \sin 1 \cos n \\ \sin(n - 1) &= \sin n \cos 1 - \sin 1 \cos n, \end{aligned}$$

(since $\cos(-1) = \cos 1$ and $\sin(-1) = -\sin 1$) and so, adding these two

$$\sin(n + 1) + \sin(n - 1) = 2 \cos 1 \sin n.$$

If $\sin n \to l$, then by the shift rule $\sin(n + 1) \to l$ and $\sin(n - 1) \to l$. By the sum rule, the LHS converges to $2l$, while by the product rule, the RHS converges to $(2 \cos 1)l$. Since $\cos 1 \neq 1$ we obtain a contradiction.

So[2] the only possibility left is that $\sin n \to 0$. But then we have

$$\cos n = \frac{\sin(n + 1) - \sin n \cos 1}{\sin 1} \to 0,$$

using the sum rule. If $\sin n \to 0$ then $\sin^2 n \to 0$ by the product rule (and similarly for $\cos^2 n$), so taking limits in

$$\cos^2 n + \sin^2 n = 1$$

we obtain $0 = 1$, a contradiction. So we cannot have $\sin n \to l$ for any $l \in \mathbb{R}$.

### 2.2.2 Limits and inequalities

Strict inequalities are not preserved by limits, e.g. take $a_n = 1 - 1/n$, then $a_n < 1$ for all $n$ but $a_n \to 1$.

**Lemma 2.10.** *If $a_n \leq b_n$ for all $n$, $a_n \to a$ and $b_n \to b$, then $a \leq b$.*

[We can always take $a_n = a$ for all $n$ or $b_n = b$ for all $n$.]

2 Many thanks to the student who pointed out the need for this case - and the solution - in the lecture!

*Proof.* Suppose that $a > b$. Choose $\varepsilon = (a-b)/2$; then there exists $N$ such that for all $n \geq N$,

$$|a_n - a| < \frac{a-b}{2} \qquad \text{and} \qquad |b_n - b| < \frac{a-b}{2}.$$

So then

$$a - a_n \leq |a_n - a| < \frac{a-b}{2} \qquad \Rightarrow \qquad a_n > a - \frac{a-b}{2} = \frac{a+b}{2}$$

and

$$b_n - b \leq |b_n - b| < \frac{a-b}{2} \qquad \Rightarrow \qquad b_n < b + \frac{a-b}{2} = \frac{a+b}{2}.$$

This contradicts the assumption that $a_n \leq b_n$ for all $n$. □

**Corollary 2.11.** *If $a \leq b_n \leq c$ and $b_n \to b$; then $a \leq b \leq c$.*

*Proof.* Apply Lemma 2.10 with $a_n = a$ and $b_n = b_n$, so that $a \leq b$. Then apply Lemma 2.10 with $a_n = b_n$ and $b_n = c$, so that $b \leq c$. □

The following lemma is very useful to show that sequences converge.

**Lemma 2.12** (Sandwich rule)**.** *Suppose that*

$$a_n \leq b_n \leq c_n,$$

*where $a_n \to l$ and $c_n \to l$. Then $b_n \to l$.*

*Proof.* Take $\varepsilon > 0$.

Since $a_n \to l$ there exists $N_1$ such that for all $n \geq N_1$,

$$l - \varepsilon < a_n < l + \varepsilon;$$

since $c_n \to l$ there exists $N_2$ such that for all $n \geq N_2$,

$$l - \varepsilon < c_n < l + \varepsilon.$$

If we take $n \geq N := \max(N_1, N_2)$ then

$$l - \varepsilon < a_n \leq b_n \leq c_n < l + \varepsilon,$$

i.e. $|b_n - l| < \varepsilon$, which shows that $b_n \to l$. □

## 2.3 Examples using the sandwich rule

We start with a simple example: $\frac{1}{n}\sin(n) \to 0$ as $n \to \infty$. Just note that we have

$$-\frac{1}{n} \le \frac{1}{n}\sin(n) \le \frac{1}{n};$$

and we know that $-1/n \to 0$ and $1/n \to 0$.

We now look at two more interesting limits.

**Lemma 2.13.** *For any $x > 0$, $x^{1/n} \to 1$ as $n \to \infty$.*

*Proof.* First we take $x > 1$ and write $x^{1/n} = 1 + h_n$. Since $x^{1/n} > 1$ when $x > 1$, we know that $h_n > 0$.

So we have

$$x = (1 + h_n)^n \ge 1 + nh_n > nh_n \qquad \Rightarrow \qquad h_n < \frac{x}{n}. \tag{2.3}$$

Since we have $0 < h_n < \frac{x}{n}$, it follows by the Sandwich Rule that $h_n \to 0$ as $n \to \infty$, and so (by sum rule for limits, since $x^{1/n} = 1 + h_n$) $x^{1/n} \to 1$ as $n \to \infty$.

For $x < 1$ we can write $x^{1/n} = 1/(1/x)^{1/n}$, and use the fact that $(1/x)^{1/n} \to 1$ as $n \to \infty$ and the quotient rule for limits. □

**Lemma 2.14.** *$n^{1/n} \to 1$ as $n \to \infty$.*

If we follow the above proof replacing $x$ by $n$ then in (2.3) we would obtain $h_n < 1$, which doesn't give us the right limit in the 'sandwich'. Instead, we look at $n^{1/2n} = [n^{1/2}]^{1/n}$, which will make a similar proof work.

*Proof.* We will show first that $n^{1/2n} \to 1$. Arguing as before, since $n \ge 1$ we have $n^{1/2n} \ge 1$. Let's write $n^{1/2n} = 1 + h_n$, and then

$$\begin{aligned}\sqrt{n} = (n^{1/2n})^n &= (1 + h_n)^n \\ &\ge 1 + nh_n > nh_n,\end{aligned}$$

using Bernoulli's inequality/binomial expansion. Rearranging we have

$$0 < h_n \le \frac{1}{\sqrt{n}},$$

and so by the Sandwich Rule, $h_n \to 0$, which shows that (sum rule again) $n^{1/2n} \to 1$. It then follows that $n^{1/n} = (n^{1/2n})^2 \to 1$ by the product rule for limits. □

## 2.4 Powers and factorials

We say that $a_n \to \infty$ as $n \to \infty$ if for every $R \in \mathbb{R}$ there exists $N$ such that

$$a_n > R \qquad \text{for every } n \geq N.$$

[There is a corresponding definition for $a_n \to -\infty$: for every $R \in \mathbb{R}$ there exists $N$ such that $a_n < R$ for all $n \geq N$.]

Some very simple examples:

- $n \to \infty$ as $n \to \infty$ (!): give $R \in \mathbb{R}$, take $N > R$.
- $\sqrt{n} \to \infty$ as $n \to \infty$: given $R \in \mathbb{R}$, take $N > R^2$; then for $n \geq N$ we have $\sqrt{n} \geq \sqrt{N} > R$.
- if $x > 1$ then $x^n \to \infty$ as $n \to \infty$. Write $x = 1 + h$; then by binomial/Bernoulli we have
$$x^n = (1 + h)^n \geq 1 + nh.$$
Given $R \in \mathbb{R}$, pick $N$ such that $1 + Nh > R$. Then for $n \geq N$,
$$x^n \geq 1 + nh \geq 1 + Nh > r.$$

**Lemma 2.15.** *If $a_n \to \infty$ as $n \to \infty$ then $1/a_n \to 0$ as $n \to \infty$.*

*Proof.* Given $\varepsilon > 0$, since $a_n \to \infty$ there exists $N$ such that $a_n > 1/\varepsilon$ for all $n \geq N$. Then
$$0 \leq 1/a_n < \varepsilon$$
for all $n \geq N$, so in particular $|1/a_n| < \varepsilon$ for all $n \geq N$ and so $1/a_n \to 0$. $\square$

There is a partial converse, see Examples 2.

**Lemma 2.16** (Comparison test for sequences)**.** *If $a_n \geq b_n$ and $b_n \to \infty$ as $n \to \infty$, then $a_n \to \infty$ as $n \to \infty$.*

The proof is a simple exercise using the definition, see Examples 2.

We now look at some standard limits that mix powers and factorials; the following lemma will be useful here (and more generally). Remember that we have already shown that $x^n \to 0$ if $0 \leq x < 1$ and that $x^n \to \infty$ if $x > 1$.

**Lemma 2.17** (Ratio test for limits)**.** *Suppose that $(a_n)$ is a sequence of positive terms, and that*
$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = r.$$
*If $r < 1$ then $a_n \to 0$, and if $r > 1$ then $a_n \to \infty$.*

If $r = 1$ then we cannot deduce anything: consider the three sequences $a_n = 1/n$ for each $n$, $a_n = 1$ for all $n$, and $a_n = n$ for each $n$. In all three cases $a_{n+1}/a_n \to 1$ as $n \to \infty$. (We will see similar issues later.)

*Proof.* Suppose first that $r < 1$. Choose $\varepsilon > 0$ such that $r + \varepsilon < 1$; there exists $N$ such that

$$\frac{a_{n+1}}{a_n} < r + \varepsilon, \qquad \text{for all } n \geq N.$$

So then $a_{N+k} \leq (r+\varepsilon)^k a_N$ for all $k \geq 1$. Since $|r + \varepsilon| < 1$, $(r+\varepsilon)^k \to 0$; so $a_{N+k} \to 0$ as $k \to 0$, which shows that $a_n \to 0$ as $n \to \infty$ (by the shift rule).

On the other hand, if $r > 1$ then we choose $\varepsilon > 0$ such that $r - \varepsilon > 1$; there exists $N$ such that

$$\frac{a_{n+1}}{a_n} > r - \varepsilon, \qquad \text{for all } n \geq N,$$

and $a_{N+k} > (r-\varepsilon)^k a_N$ for all $k \geq 1$. Since $r - \varepsilon > 1$, $(r-\varepsilon)^k \to \infty$ as $k \to \infty$; so $a_{N+k} \to \infty$ as $k \to \infty$, which shows that $a_n \to \infty$ as $n \to \infty$. □

We now prove some 'standard limits' that mix powers and factorials.

**Proposition 2.18.** *We have the following limits as $n \to \infty$:*

(i) $|x|^n \to 0$ *if* $|x| < 1$ *and* $|x|^n \to \infty$ *if* $|x| > 1$*;*

(ii) *for* $\alpha \in \mathbb{Q}$, $n^\alpha \to \infty$ *if* $\alpha > 0$ *and* $n^\alpha \to 0$ *if* $\alpha < 0$*;*

(iii) $x^n/n^k \to \infty$ *if* $x > 1$ *(and converges to* $0$ *if* $0 \leq x <\leq 1$*);*

(iv) $x^n/n! \to 0$*; and*

(v) $n!/n^n \to 0$

Note in (i) that if $x < 0$ then the sequence $x^n$ has alternate signs, so we get convergence of $x^n$ (to zero) as $n \to \infty$ if $-1 < x < 0$, but no convergence if $x \leq -1$. In (ii) we assume that $\alpha \in \mathbb{Q}$, since we have not defined powers otherwise.

*Proof.* We have already proved the various parts of (i).

(ii) Given $R \in \mathbb{R}$, choose $N$ such that $N > R^{1/\alpha}$; then if $n \geq N$, $n^\alpha \geq N^\alpha > R$. So $n^\alpha \to \infty$. For $\alpha < 0$ we can use the question on Examples 1, or apply Lemma 2.15.

(iii) Consider $a_n = x^n/n^k$, then

$$\frac{a_{n+1}}{a_n} = \frac{x^{n+1}}{n^k} \times \frac{(n+1)^k}{x^n} = x\left(1 + \frac{1}{n}\right)^k.$$

Since $1 + \frac{1}{n} \to 1$, by the product rule applied repeatedly we know that $(1 + \frac{1}{n})^k \to 1$ as $n \to \infty$ as well. So $a_{n+1}/a_n \to x$ as $n \to \infty$, so if $x > 1$ it follows from Lemma 2.17 that $x^n/n^k \to \infty$.

If $0 \le x \le 1$ then $x^n/n^k \le n^{-k}$, which tends to zero by (ii).

(iv) Set $a_n = x^n/n!$. Then

$$a_{n+1}/a_n = \frac{x^{n+1}}{(n+1)!} \times \frac{n!}{x^n} = \frac{x}{n+1}.$$

Since $x/(n+1) \to 0$ as $n \to \infty$, it follows from Lemma 2.17 that $x^n/n! \to 0$.

(v) We have

$$0 < \frac{n!}{n^n} = \frac{n(n-1)\cdots 1}{n^n} = \frac{n}{n}\frac{n-1}{n}\cdots\frac{1}{n} < \frac{1}{n} \to 0,$$

so $n!/n^n \to 0$ by the sandwich rule.

$\square$

If $x > 1$, $k \in \mathbb{N}$, then, as $n \to \infty$,

$$n^k \qquad \ll \qquad x^n \qquad \ll \qquad n! \qquad \ll \qquad n^n,$$

using $a_n \ll b_n$ to mean that $a_n/b_n \to 0$ as $n \to \infty$. [Factorials beat exponentials beat powers.]

# Chapter 3

# Monotonic sequences; the Bolzano–Weierstrass Theorem; Cauchy sequences

In some cases we can prove that a sequence converges even if we do not know the limit.

A sequence $(a_n)$ is *increasing* if

$$a_1 \leq a_2 \leq a_3 \leq \cdots$$

(in general $a_{n+1} \geq a_n$) - the terms increase or stay the same. [If $a_{n+1} > a_n$ we say that the sequence is *strictly increasing*.]

A sequence $(a_n)$ is *decreasing* if

$$a_1 \geq a_2 \geq a_3 \geq \cdots$$

(i.e. $a_{n+1} \leq a_n$); the terms decrease or stay the same. [If $a_{n+1} < a_n$ for all $n$ then we say that the sequence is *strictly decreasing*.]

A sequence is (strictly) monotonic if it is (strictly) increasing or (strictly) decreasing. [Note that a constant sequence is both 'increasing' and 'decreasing'.]

Note that it is a consequence of the definitions that any constant sequence is both increasing and decreasing (but neither strictly increasing nor strictly decreasing).

A sequence $(a_n)$ is *bounded above* if there exists $M \in \mathbb{R}$ such that $a_n \leq M$ for all $n$, and *bounded below* if there exists $K \in \mathbb{R}$ such that $a_n \geq K$ for all $n$. [This is the same as the set $\{a_n\}$ being bounded above or below.]

**Theorem 3.1.** *If $(a_n)$ is increasing then (i) if it is bounded above then it converges (ii) if it is not bounded above then $a_n \to \infty$.*

*Proof.* Consider the set
$$A = \{a_n : \ n \in \mathbb{N}\}.$$
This is non-empty and bounded above (by assumption), so (by the Least Upper Bound Axiom) it has a least upper bound; we set $l = \sup A$. We will show that $a_n \to l$ as $n \to \infty$.

Since $l$ is an upper bound for all elements of $A$,
$$a_n \leq l \qquad \text{for every } n \in \mathbb{N}. \tag{3.1}$$
Now take $\varepsilon > 0$. Since $l$ is the least upper bound, there exists an element $a$ of $A$ such that $a > l - \varepsilon$, i.e. there exists $N$ such that
$$a_N > l - \varepsilon.$$
Since $(a_n)$ is increasing, we know that if $n \geq N$ then
$$a_n \geq a_N > l - \varepsilon.$$

Combining this with (3.1) it follows that for all $n \geq N$ we have $|a_n - l| < \varepsilon$, which shows that $a_n \to l$ as $n \to \infty$.

(ii) Choose $R \in \mathbb{R}$. Since $(a_n)$ is not bounded above there exists $N$ such that $a_N > R$. Now since $(a_n)$ is increasing, it follows that $a_n > R$ for all $n \geq N$; so $a_n \to \infty$. $\square$

[We could take part (i) of this result as the Axiom of Completeness for $\mathbb{R}$: if we assume this then we can deduce that every non-empty set that is bounded above has a supremum, so they are equivalent, see Examples 2.]

**Corollary 3.2.** *If $(a_n)$ is a decreasing sequence then (i) if it is bounded below it converges (ii) if it is not bounded below then $a_n \to -\infty$ as $n \to \infty$.*

## 3.1 Examples

We can sometimes use the fact that a limit exists to find out what the limit is.

Example 1

Consider the sequence $(x_n)$, where
$$x_1 = 1 \qquad \text{and} \qquad x_{n+1} = 3 - \frac{1}{x_n}, \ n \geq 1.$$

To show that $x_n$ is increasing, observe that $x_2 = 2 > x_1$, and now use induction: if $x_n \geq x_{n-1}$ then

$$x_{n+1} = 3 - \frac{1}{x_n} \qquad \text{and} \qquad x_n = 3 - \frac{1}{x_{n-1}}$$

which implies that $x_{n+1} \geq x_n$. Once we know that $(x_n)$ is increasing, it follows that $x_n \geq 1$ for all $n$, so $x_{n+1} \leq 3$ for all $n$.

Since $(x_n)$ is increasing an bounded above, it must converge. Suppose that $x_n \to s$ for some $s \in \mathbb{R}$. Taking limits on both sides of

$$x_{n+1} = 3 - \frac{1}{x_n}$$

we obtain

$$s = 3 - \frac{1}{s} \qquad \Rightarrow \qquad s^2 - 3s + 1 = 0.$$

This has roots $s = 3 \pm \sqrt{5}/2$. Since $x_n$ is increasing a $x_n \geq 2$ for all $n \geq 2$, we know that we must have $s \geq 2$, so $s = (3 + \sqrt{5})/2$.

Example 2

Take $a_1 = 1$, $a_{n+1} = \sqrt{a_n + 2}$. We can prove by induction that $1 \leq a_n \leq 2$ for all $n$; so then

$$a_n^2 - a_n - 2 = (a_n - 2)(a_n + 1) \leq 0 \quad \Rightarrow \quad a_n^2 \leq a_n + 2 \quad \Rightarrow \quad a_n \leq \sqrt{a_n + 2} = a_{n+1}.$$

So $(a_n)$ is increasing and bounded above and must converge to a limit, $l$, say. Since

$$a_{n+1}^2 = a_n + 2$$

then we can take limits on both sides [using the product rule and sum rule for limits] to obtain $l^2 = l + 2$. The two solutions of this are $l = 2$ and $l = -1$. Since $a_n \in [1, 2]$, we must have $l \in [1, 2]$, so the limit must be $l = 2$.

Example 3: e in disguise

Consider

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

Note that, by Bernoulli's inequality/binomial expansion, $x_n \geq 1 + 1 = 2$ for all $x$. So $x_n \not\to 1$.

This is increasing, since

$$\begin{aligned}\frac{x_n}{x_{n-1}} &= \frac{(1+\frac{1}{n})^n}{(1+\frac{1}{n-1})^{n-1}} \\ &= \left(1+\frac{1}{n}\right)\left(\frac{(n+1)(n-1)}{n^2}\right)^{n-1} \\ &= \left(1+\frac{1}{n}\right)\left(1-\frac{1}{n^2}\right)^{n-1} \\ &\geq \left(1+\frac{1}{n}\right)\left(1-\frac{n-1}{n^2}\right) \\ &= 1+\frac{1}{n^3}.\end{aligned}$$

We saw before the $x_n$ is bounded above by 3. Here is another (neater) proof (of a worse upper bound). Rewrite $x_n$ as follows

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

Now observe that

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

So $x_{2n} \leq 4$. Since $(x_n)$ is increasing, we have $x_{2n-1} \leq x_{2n} \leq 4$. So $x_n \leq 4$ for all $n$.

Since $x_n$ is increasing and bounded above, it converges. We will see later that it converges to e.

## 3.2 Decimal expansions

The theorem that we have just proved (an increasing sequence that is bounded above tends to a limit) has an immediate corollary that any decimal corresponds to a real number.

Given a decimal expansion

$$0 \cdot d_1 d_2 d_3 d_4 d_5 \cdots,$$

where $d_i \in \{0, \ldots, 0\}$ for each $i$, we can consider the sequence

$$0 \cdot d_1, \quad 0 \cdot d_1 d_2, \quad 0 \cdot d_1 d_2 d_3, \quad \ldots,$$

i.e. the sequence $(x_n)$ with

$$x_n = \sum_{j=1}^{n} d_j 10^{-j}.$$

If the decimal expansion is finite (so that $d_n = 0$ for all $n \geq N$) then this sum is finite. Otherwise, $(x_n)$ is an increasing sequence that is bounded above by 1. So it converges to some real number, and this is what we mean by saying

$$0 \cdot d_1 d_2 d_3 d_4 d_5 \cdots = \sum_{j=1}^{\infty} d_j 10^{-j} = x.$$

(We will talk more about series - these sort of infinite sums - in Part II.)

Conversely, we can write any $x \in [0, 1]$ as a decimal, by setting

$$\begin{aligned}
d_1 &= \max\left\{ j : \frac{j}{10} \leq x \right\} \\
d_2 &= \max\left\{ j : \frac{d_1}{10} + \frac{j}{10^2} \leq x \right\} \\
&\vdots \\
d_n &= \max\left\{ j : \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{n-1}}{10^{n-1}} \leq x \right\}.
\end{aligned}$$

Each digit is contained in $\{0, \ldots, 9\}$, and after $n$ digits we have

$$x - \frac{1}{10^n} < \sum_{k=1}^{n} d_k 10^{-k} \leq x,$$

so the expansion converges to $x$ by the Sandwich Rule.

**Proposition 3.3.** *A number $x \in [0, 1]$ is rational if and only if its decimal expansion terminates or recurs.*

*Proof.* If $x$ is given by a terminating decimal

$$x = 0 \cdot a_1 a_2 \ldots a_n$$

then $x = a_1 a_2 \ldots a_n / 10^n \in \mathbb{Q}$.

Now suppose that $x$ is a recurring decimal:

$$x = 0 \cdot a_1 a_2 \ldots a_n \dot{d}_1 d_2 \ldots \dot{d}_m.$$

Then

$$10^n x = a_1 a_2 \ldots a_n \cdot \dot{d}_1 d_2 \ldots \dot{d}_m$$

and

$$10^{n+m}x = a_1a_2\dots a_nd_1d_2\dots d_m\cdot\dot{d}_1d_2\dots\dot{d}_m.$$

Subtracting we obtain

$$(10^{n+m}-10^n)x = a_1a_2\dots a_nd_1d_2\dots d_m - a_1a_2\dots a_n$$

from which

$$x = \frac{a_1a_2\dots a_nd_1d_2\dots d_m - a_1a_2\dots a_n}{10^{n+m}-10^n}\in\mathbb{Q}.$$

To see the converse, let $x = p/q$. We find the decimal expansion by long division. Either the division terminates or repeats, since there are at most $q$ possible remainders on division by $q$, and eventually these must repeat. This is easier to see by example than a complete formal proof, e.g. find the decimal expansion of $1/7$ (which is, as we all know, $0\cdot\dot{1}4285\dot{7}$). □

## 3.3 Subsequences and The Bolzano–Weierstrass Theorem

If $(a_n)$ is a sequence that does not converge, it may still have a subsequence that converges.

A subsequence means 'pick out some terms from the sequence'. The notation is awkward, but the idea should be fairly straightforward.

Strictly, given $(a_n)_{n=1}^{\infty}$, we get a subsequence by choosing a sequence of indices $(n_j)_{j=1}^{\infty}$, where $n_j\in\mathbb{N}$ and $n_{j+1}>n_j$ for every $j\in\mathbb{N}$: the subsequence is $(a_{n_j})_{j=1}^{\infty}$.

For example, the constant sequence of all 1s and the constant sequence of all $-1$s are both subsequences of $a_n = (-1)^n$. One subsequence of

$$0,0,1,0,0,2,0,0,3,0,0,4,0,0,5,0,0,6,0,0,7,\cdots$$

is the constant sequence 0; another is $a_{3n}$, which is $1,2,3,4,\dots$.

Consider the sequence $(a_n)$ with terms

$$\frac{1}{1},\frac{1}{2},\frac{2}{2},\frac{1}{3},\frac{2}{3},\frac{3}{3},\frac{1}{4},\frac{2}{4},\frac{3}{4},\frac{4}{4},\dots$$

(this lists all the rational numbers in the interval $(0,1]$). Then given any $x\in(0,1]$, there is a subsequence that converges to $x$. See Examples 2. For another example, see Section **??**.

**Lemma 3.4.** *If $a_n\to l$, then any subsequence converges to $l$.*

*Proof.* Let $(a_{n_j})_j$ be a subsequence of $(a_n)$. Take $\varepsilon > 0$. Since $a_n \to l$, there exists $N$ such that $|a_n - l| < \varepsilon$ for all $n \geq N$. Now since $n_j \to \infty$, there exists $J$ such that $n_j \geq N$ for all $j \geq J$. It follows that

$$|a_{n_j} - l| < \varepsilon$$

for all $j \geq J$, and so $a_{n_j} \to l$ as $j \to \infty$, as claimed. $\square$

This gives a much neater proof of the fact that $a_n = (-1)^n$ does not converge.

The following result at first may seem simply a curiosity, but its consequence (the Bolzano–Weierstrass Theorem) is extremely powerful.

**Proposition 3.5.** *Any sequence of (real numbers) contains a monotonic subsequence.*

*Proof.* Let $(a_n)$ be our sequence. We'll call an element $a_j$ of the sequence 'high' if it is greater than (or equal to) any subsequent element, i.e.

$$a_j \geq a_n \qquad \text{for all } n \geq j.$$

Suppose that there are an infinite number of high elements, with indices $(n_j)_{j=1}^{\infty}$. Then, by definition,

$$a_{n_{j+1}} \leq a_{n_j} \qquad \text{for every } j,$$

so $(a_{n_j})_{j=1}^{\infty}$ is a decreasing subsequence.

If there are only a finite number of high elements, then there is an index $N$ such that $a_j$ is not high for every $j \geq N$.

Take $n_1 = N$. Since $a_{n_1}$ is not high, there exists $n_2 > n_1$ such that $a_{n_2} > a_{n_1}$. Since $a_{n_2}$ is not high, there exists $n_3 > n_2$ such that $a_{n_3} > a_{n_2}$. Continuing in this way we find indices $n_{j+1} > n_j$ such that $a_{n_{j+1}} > a_{n_j}$. So now we have an increasing subsequence $(a_{n_j})_{j=1}^{\infty}$. $\square$

**Theorem 3.6** (Bolzano–Weierstrass Theorem)**.** *Any bounded sequence of real numbers contains a convergent subsequence.*

*Proof.* If $(a_n)$ is bounded then it contains a monotonic subsequence; this is bounded both above and below (since $(a_n)$ is bounded) - so it is increasing and bounded above or decreasing and bounded below, so (by Theorem 3.1 or its corollary) it must converge. $\square$

[This is another statement that is equivalent to the Completeness Axiom.]

This is the simplest example of a (sequential) *compactness* result: there are many situations in which 'bounds' on a sequence allow one to extract a convergent subsequence.

These sequences can be of vectors in $\mathbb{R}^n$, or even of functions. The reason these results are so useful can be sketched as follows.

Suppose that we want to find a solution $x$ of a problem $P$. Perhaps we can approximate $P$ by a sequence of problems $P_n$, which are easier to solve: we can find solutions $x_n$ of these approximate problems. We would then like to show that $x_n \to x$ and that $x$ is a solution of $P$. Compactness results like the Bolzano–Weierstrass Theorem mean that if we can show that $x_n$ is bounded then we can at least find a subsequence $x_{n_j}$ that converges to some $x$; we then hope to be able to show that $x$ solves our initial problem $P$.

## 3.4 Cauchy sequences

How do we see if a sequence converges when we don't know the limit?

Suppose that $a_n \to l$. Then for any choice of $\varepsilon > 0$ there exists $N$ such that $|a_n - l| < \varepsilon/2$ for all $n \geq N$. Let's take $m, n \geq N$. Then

$$|a_n - a_m| \leq |a_n - l| + |l - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $m, n \geq N$, by the triangle inequality.

This is a property of convergent sequence that does not feature the limit $l$.

**Definition 3.7.** A sequence $(a_n)$ is *Cauchy* if for each $\varepsilon > 0$ there exists an $N$ such that

$$|a_n - a_m| < \varepsilon \qquad \text{for } m, n \geq N.$$

In fact, if a sequence is Cauchy then it must converge. The proof requires an easy lemma.

**Lemma 3.8.** *Any Cauchy sequence is bounded.*

(If any Cauchy sequence converges this had better be true, since we showed that any convergent sequence must be bounded.)

*Proof.* Take $\varepsilon = 1$ in the definition of a Cauchy sequence. Then there exists $N$ such that for all $m, n \geq N$,

$$|a_m - a_n| < 1.$$

In particular, if we take $m = N$ then we have $|a_N - a_n| < 1$ for all $n \geq N$. So for $n \geq N$ we have

$$|a_n| \leq |a_n - a_N| + |a_N| < 1 + |a_N|.$$

Therefore we have

$$|a_n| \le \max(|a_1|, \ldots, |a_{N-1}|, |a_N| + 1)$$

for all $n \in \mathbb{N}$ and $(a_n)$ is bounded. $\square$

We can now prove our final key result about sequences.

**Theorem 3.9** (General Principle of Convergence)**.** *A sequence of real numbers converges if and only if it is Cauchy.*

*Proof.* We have already shown that any convergent sequence is Cauchy.

So suppose that $(a_n)$ is a Cauchy sequence. Since $(a_n)$ is bounded, it must have a convergent subsequence: $a_{n_j} \to l$ as $j \to \infty$, for some $l \in \mathbb{R}$. We just have to show that $a_n \to l$.

Pick $\varepsilon > 0$. Since $(a_n)$ is Cauchy, there exists $N$ such that

$$|a_n - a_m| < \varepsilon/2 \qquad \text{for all } m, n \ge N.$$

Since the subsequence $a_{n_j} \to l$ as $j \to \infty$, we can pick $J$ such that $n_J \ge N$ and such that

$$|a_{n_j} - l| < \varepsilon/2 \qquad \text{for all } j \ge J.$$

If $k \ge n_J$ then, since $n_J \ge N$, we have $k, n_J \ge N$ and so

$$|a_k - a_{n_J}| < \varepsilon/2;$$

we also know that $|a_{n_J} - l| < \varepsilon/2$, so

$$|a_k - l| = |a_k - a_{n_J} + a_{n_J} - l| \le |a_k - a_{n_J}| + |a_{n_J} - l| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which shows that $a_k \to l$ as $k \to \infty$. $\square$

As an example, suppose that $(a_n)$ is a sequence such that

$$|a_{n+1} - a_n| \le R\alpha^n \qquad n \ge M,$$

for some $R > 0$, $M \in \mathbb{N}$, $\alpha \in (0, 1)$. Then for all $m \ge n \ge N \ge M$ we have

$$\begin{aligned} |a_m - a_n| &\le |a_m - a_{m_1}| + \cdots + |a_{n+1} - a_n| \\ &\le R\alpha^{m-1} + \cdots + R\alpha^n \\ &\le R\alpha^n(1 + \alpha + \alpha^2 + \cdots) \\ &= \frac{R\alpha^n}{1-\alpha}. \end{aligned}$$

Given $\varepsilon > 0$, choose $N \ge M$ so that $R\alpha^N/(1-\alpha) < \varepsilon$. Then for all $m, n \ge N$ we have $|a_m - a_n| < \varepsilon$, so $(a_n)$ is a Cauchy sequence and hence converges to a limit.

# Part III

# Series

# Chapter 4

# Summation

We now have a rigorous way to talk about convergence of sequences. We apply this to series, i.e. 'infinite sums'. If $(a_n)$ is a sequence then it forms the *terms* of the series

$$a_1 + a_2 + a_3 + \cdots$$

The key to attaching a mean to limit infinite sum $\sum_{j=1}^{\infty} a_j$ is to consider this as the limit of the sequence of *partial sums* $\sum_{j=1}^{n} a_j$.

**Definition 4.1.** If $(a_n)$ is a sequence, then we say that

$$\sum_{j=1}^{\infty} a_j = s$$

if the sequence $(s_n)$ of *partial sums*

$$s_n = \sum_{j=1}^{n} a_j$$

converges to $s$.

We could also start (as in our first example) at an index other than 1.

<u>Example: geometric sequences</u>

If $|r| < 1$ then

$$\sum_{j=0}^{\infty} r^j = \frac{1}{1-r}. \tag{4.1}$$

Indeed, the partial sums are

$$\sum_{j=0}^{n} r^j = \frac{1-r^{n+1}}{1-r} \tag{4.2}$$

provided that $r \neq 1$. You can prove this by induction of the neat trick of writing

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

and divide by $1-r$. To obtain (4.1) take $n \to \infty$ in (4.2)

Example: a telescoping sequence

We have

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

The partial sums are

$$\sum_{n=1}^{m} \frac{1}{n(n+1)} = \sum_{n=1}^{m} \frac{1}{n} - \frac{1}{n+1} = ... = \frac{m}{m+1} = \frac{1}{1+(1/m)};$$

so

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

## 4.1 General results about convergent series

We prove two simple results on the algebra of summation

**Lemma 4.2.** *If* $\sum_{j=1}^{\infty} a_j = t$ *and* $\sum_{j=1}^{\infty} b_j = s$ *then*

$$\sum_{j=1}^{\infty} (a_j + b_j) = t + s.$$

*If* $\sum_{j=1}^{\infty} a_j = s$ *and* $c \in \mathbb{R}$ *then*

$$\sum_{j=1}^{\infty} c a_j = cs.$$

*Proof.* We have

$$\sum_{j=1}^{n} a_j \to t \qquad \text{and} \qquad \sum_{j=1}^{n} b_j \to s,$$

so

$$\sum_{j=1}^{n} (a_j + b_j) = \left(\sum_{j=1}^{n} a_j\right) + \left(\sum_{j=1}^{n} b_j\right) \to t + s$$

by algebra of limits. The other part is easy. □

If a sequence does converge then its terms must tend to zero; but this is not enough for convergence.

**Lemma 4.3.** *If* $a_n \not\to 0$ *then* $\sum_{n=1}^{\infty} a_n$ *does not converge.*

We will actually prove that if $\sum_{n=1}^{\infty} a_n$ does converge then $a_n \to 0$ as $n \to \infty$. But note that this is *not* a condition that ensures convergence. We will shortly see, for example, that $\sum \frac{1}{n}$ does not converge, even though its terms tend to zero. This is the reason for stating the result as we have done as a 'non-convergence' test.

*Proof.* We show that if $\sum_{n=1}^{\infty} a_n$ converges then $a_n \to 0$ as $n \to \infty$

Suppose that $\sum_{n=1}^{\infty} a_n = s$. This means that

$$\sum_{j=1}^{n} a_j \to s.$$

But we also have (shift rule)

$$\sum_{j=1}^{n-1} a_j \to s,$$

so

$$a_n = \left(\sum_{j=1}^{n} a_j\right) - \left(\sum_{j=1}^{n-1} a_j\right) \to s - s = 0. \qquad \square$$

To show that $a_n \to 0$ does not imply that $\sum a_n$ converges we give an example: $\sum_{j=1}^{\infty} \frac{1}{j}$ ('the harmonic series') does not converge.

Suppose that it does, i.e. that $s_n := \sum_{j=1}^{n} \frac{1}{n} \to l$ as $n \to \infty$. Now note that we have

$$\begin{aligned} 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots &> \frac{1}{2} + \left[\frac{1}{2} + \frac{1}{2}\right] + \left[\frac{1}{4} + \frac{1}{4}\right] + \left[\frac{1}{6} + \frac{1}{6}\right] \cdots \\ &= \frac{1}{2} + \left[1 + \frac{1}{2} + \frac{1}{3} + \cdots\right] \end{aligned}$$

which would imply that $l \geq \frac{1}{2} + l$, which is impossible. [To do this properly, observe that we are in fact showing that $s_{2n} > \frac{1}{2} + s_n$, and now we can use the algebra of limits to deduce that $l \geq \frac{1}{2} + l$.]

To see why $a_n \to 0$ might not be enough, we show that convergence actually implies much more. First, we observe that

$$\sum_{j=1}^{\infty} a_j \text{ converges} \quad \Leftrightarrow \quad \sum_{j=k}^{\infty} a_j \text{ converges for some } k \in \mathbb{N}.$$

This follows since we can write

$$\sum_{j=1}^{n} a_j = \left( \sum_{j=1}^{k-1} a_j \right) + \sum_{j=k}^{n} a_j. \tag{4.3}$$

If either sum in (4.3) converges then we can take $n \to \infty$ and obtain

$$\sum_{j=1}^{\infty} a_j = \left( \sum_{j=1}^{k-1} a_j \right) + \sum_{j=k}^{\infty} a_j. \tag{4.4}$$

**Lemma 4.4** (Vanishing 'tails'). *If $\sum a_n$ converges then*

$$\sum_{j=n}^{\infty} a_j \to 0 \qquad \textit{as} \qquad n \to \infty.$$

*Proof.* Rearrange (4.4) as

$$\sum_{j=k}^{\infty} a_j = \sum_{j=1}^{\infty} a_j - \left( \sum_{j=1}^{k-1} a_j \right);$$

take $k \to \infty$, using the fact that $\lim_{k\to\infty} \sum_{j=1}^{k-1} a_j = \sum_{j=1}^{\infty} a_j$. □

## 4.2 Series with positive terms

If $a_j \geq 0$ for every $j$ then

$$\sum_{j=1}^{n} a_j$$

is an increasing sequence, so it converges if and only if it is bounded above. We often write

$$\sum_{j=1}^{\infty} a_j < \infty$$

to mean that a series with positive terms converges. If a sequence of positive terms is not bounded above then its partial sums become arbitrarily large and we say that the series diverges; in this case we might write $\sum_{j=1}^{\infty} a_j = \infty$ (but this is just a 'shorthand'). We have just proved the following.

**Lemma 4.5.** *If $a_n \geq 0$ and the partial sums $\sum_{j=1}^{n} a_j$ are bounded above then $\sum_{j=1}^{\infty} a_j$ converges. If the partial sums are not bounded above then the series diverges.*

Example: the harmonic series

We now show again that
$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

To see that the sum is not bounded above, we divide the sum into blocks:
$$1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{>2/4=1/2} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{>4/8=1/2} + \cdots;$$
each of the $2^{n-2}$ terms in the $n$th block is at least $1/2^{n-1}$, so the $n$th block adds up to at least $1/2$. So $s_{2^n} \geq n/2$, and the partial sums are unbounded.

Example: the Basel problem

We show that
$$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$
If we note that $n^2 \geq n(n-1)$, then
$$\sum_{n=1}^{m} \frac{1}{n^2} \leq 1 + \sum_{n=1}^{m-1} \frac{1}{n(n+1)},$$
and we showed before that $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} < \infty$. [What is the value of this sum? Finding this value is known as the 'Basel Problem', which was originally solved by Euler, who showed that it is $\pi^2/6$. We will see an 'elementary' proof of this a little later.]

The following simple lemma is the key result that allows us to deduce convergence/divergence of one series from another.

**Lemma 4.6** (Comparison Test). *Suppose that $0 \leq a_n \leq b_n$ for every $n$. Then*

- *if $\sum_{j=1}^{\infty} b_j$ converges then $\sum_{j=1}^{\infty} a_j$ converges;*
- *if $\sum_{j=1}^{\infty} a_j$ diverges then $\sum_{j=1}^{\infty} b_j$ diverges.*

(We already used essentially the argument in the proof to show that $\sum 1/n^2 < \infty$ converges.)

*Proof.* Write $\alpha_n = \sum_{j=1}^{n} a_j$ and $\beta_n = \sum_{j=1}^{n} b_j$. Since $a_n \leq b_n$ for every $n$,
$$\alpha_n \leq \beta_n$$

for every $n$. Both $(\alpha_n)$ and $(\beta_n)$ are increasing sequences. If $\sum b_j$ converges then $\alpha_n \leq \beta_n \leq \sum b_j$ for every $n$, so $(\alpha_n)$ is an increasing sequence that is bounded above and so converges. Conversely, if $\sum a_j$ does not converge then the partial sums $\alpha_n$ are not bounded above, so nor are the partial sums $\beta_n$, and so the series $\sum b_j$ must diverge. $\square$

One example: $\sum \frac{1}{n^p}$ converges if $p \geq 2$, since $\frac{1}{n^p} \leq \frac{1}{n^2}$ and $\sum \frac{1}{n^2}$ converges; it diverges if $p \leq 1$ since then $\frac{1}{n^p} \geq \frac{1}{n}$ and $\sum \frac{1}{n}$ diverges. What happens for $1 < p < 2$? For one approach based on grouping as in the proof above that $\sum 1/n$ diverges, see Examples 3. For another approach using integration see later.

**Lemma 4.7** (Comparison Test (improved))**.** *Suppose that there exists $N \in \mathbb{N}$ such that $0 \leq a_n \leq b_n$ for every $n \geq N$. Then*

- *if $\sum_{j=1}^{\infty} b_j$ converges then $\sum_{j=1}^{\infty} a_j$ converges;*
- *if $\sum_{j=1}^{\infty} a_j$ diverges then $\sum_{j=1}^{\infty} b_j$ diverges.*

*Proof.* We proved earlier that $\sum_{j=1}^{\infty} a_j$ converges if and only if $\sum_{j=N}^{\infty} a_j$ converges. So we can just apply the comparison test to the sums starting at $j = N$. $\square$

We now look at two examples. First, we take

$$a_n = \frac{n+1}{n^3 - n - 1}$$

and claim that $\sum_{n=2}^{\infty} a_n$ converges. For large $n$, $a_n \sim \frac{1}{n^2}$; so we expect $\sum a_n < \infty$. For $n \geq 2$ we have $n + 1 < \frac{1}{2}n^3$, so

$$a_n = \frac{n+1}{n^3 - n - 1} < \frac{2n}{n^3 - n - 1} < \frac{2n}{n^3 - (n^3/2)} = \frac{4}{n^2}.$$

Since $\sum \frac{4}{n^2} < \infty$, it follows that $\sum a_n$ converges.

Now take

$$a_n = \frac{n-2}{n^2 + 2n + 4}.$$

For large $n$, $a_n \sim \frac{1}{n}$, so we would expect $\sum a_n$ to diverge. Note that $n^2 + 2n + 4 > n^2 + 2n^2 + 4n^2 = 7n^2$, and that $n - 2 \geq n/2$ for all $n \geq 4$; so we have

$$a_n = \frac{n-2}{n^2 + 2n + 4} > \frac{n/2}{7n^2} = \frac{1}{14n}.$$

Since $\sum \frac{1}{14n}$ diverges, so does $\sum a_n$.

For another version of the comparison test that is easily to apply in these two examples, see Examples 3.

## 4.3 The number e

We define

$$\mathrm{e} = \sum_{k=0}^{\infty} \frac{1}{k!}.$$

This sum converges by the comparison test: for all $k \geq 2$ we have

$$\frac{1}{k!} \leq 2^{-(k-1)}.$$

In fact, the sum converges very fast: for[1] any $n \geq 1$ we have

$$\begin{aligned}
\mathrm{e} - \sum_{k=0}^{n} \frac{1}{k!} &= \sum_{k=n+1}^{\infty} \frac{1}{k!} \\
&= \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \cdots\right) \\
&< \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \cdots\right) \leq \frac{1}{n!},
\end{aligned}$$

since for any $m$, $\sum_{k=1}^{m}(\frac{1}{n+1})^k < \sum_{k=1}^{\infty}(\frac{1}{n+1})^k = \frac{1}{n} \leq 1$. So for any $n \geq 1$ we have

$$0 < \mathrm{e} - \sum_{k=0}^{n} \frac{1}{k!} < \frac{1}{n!}. \tag{4.5}$$

We will use this to show that e is irrational.

### 4.3.1 Irrationality of e

**Theorem 4.8.** *The number* e *is irrational.*

*Proof.* Suppose that e is rational, so that $\mathrm{e} = p/q$ for some integers $p$ and $q$, where $p$ and $q$ have no common factors. Choosing $n = q$ in (4.5) we obtain

$$0 < \frac{p}{q} - \sum_{k=0}^{q} \frac{1}{k!} < \frac{1}{q!}.$$

If we multiply this inequality by $q!$ then we obtain

$$0 < p[(q-1)!] - \sum_{k=0}^{q} \frac{q!}{k!} < 1.$$

1 If we take $n = 0$ then the second line of the inequality is not true, since $\frac{1}{1!} = 1 = \frac{1}{0!}\frac{1}{1}$.

Now, note that all the terms in the expression

$$p[(q-1)!] - \sum_{k=0}^{q} \frac{q!}{k!}$$

are integers: since $q \geq k$, we have $q!/k! = q(q-1)\cdots(k+1)$. So this term is the sum of integers (most of them negative), and so is an integer itself. This gives a contradiction, since there are no integers lying strictly between 0 and 1. $\square$

### 4.3.2 $(1+\frac{1}{n})^n \to \mathrm{e}$

**Proposition 4.9.**

$$\left(1+\frac{1}{n}\right)^n \to \mathrm{e}.$$

*Proof.* Set $x_n = (1+(1/n))^n$; we have already shown that $x_n \to l$ for some $l \in \mathbb{R}$ in Section 3.1.

If we expand $x_n$ using the binomial theorem we have

$$\begin{aligned} x_n &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k (1)^{n-k} \\ &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{1}{n^k} \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left(\frac{n(n-1)\cdots(n-k+1)}{n^k}\right) \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}\right) \\ &\leq \sum_{k=0}^{n} \frac{1}{k!}. \end{aligned}$$

It follows that $x_n \leq \mathrm{e}$ for all $n$, and since we already know that $x_n$ converges (or using the comparison test) we have $l \leq \mathrm{e}$.

Now fix $m \in \mathbb{N}$, and consider

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^{m+n} &= \sum_{k=0}^{n+m} \binom{n+m}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n+m} \frac{1}{k!} \frac{(n+m)!}{(n+m-k)!} \frac{1}{n^k} \\
&\geq \sum_{k=0}^{m} \frac{1}{k!} \frac{(n+m)!}{(n+m-k)!} \frac{1}{n^k}.
\end{aligned}$$

So we have

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^{m+n} &\geq \sum_{k=0}^{m} \frac{1}{k!} \frac{(n+m)(n+m-1)\cdots(n+m-k+1)}{n^k} \\
&\geq \sum_{k=0}^{m} \frac{1}{k!}.
\end{aligned}$$

Since $m$ is fixed, we can let $n \to \infty$ and use the fact that limits preserve inequalities to deduce that

$$\begin{aligned}
\sum_{k=0}^{m} \frac{1}{k!} \leq \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^{n+m} &= \lim_{n\to\infty} \left[\left(1+\frac{1}{n}\right)^{n} \left(1+\frac{1}{n}\right)^{m}\right] \\
&= \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^{n} \lim_{n\to\infty} \left(1+\frac{1}{n}\right)^{m} \\
&= l \times 1 = l
\end{aligned}$$

(using the product rule for limits for $\lim(1+\frac{1}{n})^m$ and on the whole expression). So $l \geq \sum_{k=0}^{m} \frac{1}{k!}$ for all $m$, which implies, letting $m \to \infty$, that $l \geq \mathrm{e}$. □

## *4.4 The Basel problem

We use an elementary argument to find the value of $\sum 1/n^2$, which we showed to converge in Section 4.2. It uses elementary properties of trigonometric functions (which we have not yet introduced rigorously). [The proof is from T.J. Ransford (1982) An elementary proof of $\sum_1^\infty \frac{1}{n^2} = \frac{\pi^2}{6}$. *Eureka* **42**, Summer 1982.]

We start with the identity

$$\cos nx + \mathrm{i}\sin nx = \mathrm{e}^{\mathrm{i}nx} = (\mathrm{e}^{\mathrm{i}x})^n = (\cos x + \mathrm{i}\sin x)^n = \sum_{j=0}^{n} \binom{n}{j} \mathrm{i}^j \cos^{n-j} x \sin^j x,$$

which implies, by taking the imaginary part of both sides, that

$$\sin(2m+1)x = \sum_{j=0}^{m} \binom{2m+1}{2j+1} \cos^{2m-2j} x(-1)^j \sin^{2j+1} x.$$

[You could prove this by induction rather than using complex numbers.]

If we let $x = k\pi/(2m+1)$ with $k = 1, \ldots, m$ then $\sin(2m+1)x = \sin k\pi = 0$ and $\sin x \neq 0$, so we can divide this equation by $\sin^{2m+1} x$ to obtain

$$\sum_{j=1}^{m} (-1)^j \binom{2m+1}{2j} \cot^{2(m-j)} \left(\frac{k\pi}{2m+1}\right) = 0, \qquad k = 1, \ldots, m.$$

This means that the $m$ roots of the equation

$$\sum_{j=0}^{m} (-1)^j \binom{2m+1}{2j+1} t^{m-j} = 0$$

are $\cot^2(k\pi/(2m+1))$. The sum of these roots are given by the coefficient of $t^{m-1}$ if the leading term has coefficient 1, so

$$\sum_{k=1}^{m} \cot^2 \left(\frac{k\pi}{2m+1}\right) = \binom{2m+1}{3} \Big/ \binom{2m+1}{1} = \frac{1}{3} m(2m-1). \tag{4.6}$$

Now add $m$ to both sides and use the identity $1 + \cot^2 \theta = \text{cosec}^2\theta$ to give

$$\sum_{k=1}^{m} \text{cosec}^2 \left(\frac{k\pi}{2m+1}\right) = \frac{1}{3} m(2m+2). \tag{4.7}$$

Finally, we know that for $0 < \theta < \pi/2$ we have

$$0 < \sin\theta < \theta < \tan\theta \qquad \Rightarrow \qquad \cot\theta < \frac{1}{\theta} < \text{cosec}\theta.$$

If we use this inequality in both (4.6) and (4.7) we obtain

$$\frac{1}{3} m(2m-1) < \sum_{k=1}^{m} \left(\frac{2m+1}{k\pi}\right)^2 < \frac{1}{3} m(2m+2),$$

or equivalently

$$\frac{\pi^2}{6} \left(1 - \frac{1}{2m+1}\right) \left(1 - \frac{2}{2m+1}\right) < \sum_{k=1}^{m} \frac{1}{k^2} < \frac{\pi^2}{6} \left(1 - \frac{1}{2m+1}\right) \left(1 + \frac{1}{2m+1}\right).$$

If we let $m \to \infty$ then we obtain

$$\sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{6}.$$

## 4.5 Absolute convergence

We say that $\sum a_n$ *converges absolutely* if $\sum |a_n| < \infty$.

Note that there are sequences that converge that do not converge absolutely; for example, while $\sum \frac{1}{n}$ does not converge, the series

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

does converge.

If we group the terms together in pairs, we have

$$\frac{1}{2n-1} - \frac{1}{2n} = \frac{2n-(2n-1)}{(2n-1)(2n)} = \frac{1}{(2n-1)(2n)} < \frac{1}{n^2},$$

and we know that $\sum \frac{1}{n^2}$ converges. So $s_{2n} = \sum_{k=1}^{2n} (-1)^{k+1} \frac{1}{k}$ is an increasing sequence that is bounded above and converges to some limit $l$; since the terms also converge to zero, $s_{2n+1} = s_{2n} + a_{2n+1} \to l$ as well, so (by Q2 on Examples 2) $s_n \to l$ as $n \to \infty$. [We will work out the value of this sum later.]

**Lemma 4.10** (Absolute convergence implies convergence)**.** *Suppose that $\sum |a_n|$ converges. Then $\sum a_n$ converges.*

There is a much 'neater' proof using Cauchy sequences, which you can find on Examples 3. But we give a proof using a 'trick' that we will also use later in Section 4.8.

*Proof.* We define two new sequences with positive terms:

$$b_n = \begin{cases} a_n & a_n > 0, \\ 0 & a_n \leq 0; \end{cases} \qquad c_n = \begin{cases} 0 & a_n \geq 0, \\ -a_n & a_n < 0. \end{cases}$$

Note that $b_n \leq |a_n|$, $c_n \leq |a_n|$, and $a_n = b_n - c_n$.

Since $\sum |a_n|$ converges, it follows from the Comparison Test (version 1) that both $\sum b_n$ and $\sum c_n$ converge. Since $a_n = b_n - c_n$, it now follows from the sum rule for series that $\sum a_n$ converges. $\square$

As series that converges but does not converge absolutely (like the alternating version of the harmonic series) is called *conditionally convergent*.

## 4.6 Tests for convergence

**Lemma 4.11** (Comparison Test, version 2)**.** *Suppose that* $|a_n| \leq b_n$ *for every* $n$*, and* $\sum b_n < \infty$*. Then* $\sum a_n$ *converges.*

Combine version 1 (Lemma 4.7) with Lemma 4.10.

**Lemma 4.12** (Ratio test)**.** *If* $a_n \neq 0$ *and*

$$\left|\frac{a_{n+1}}{a_n}\right| \to r \qquad \textit{as} \qquad n \to \infty$$

*then if*

- *if* $r < 1$ *the series* $\sum a_n$ *converges absolutely;*
- *if* $r > 1$ *the series* $\sum a_n$ *does not converge;*
- *if* $r = 1$ *it could be either.*

E.g. $a_n = 1/n$ and $a_n = 1/n^2$ both give $r = 1$. The first does not converge but the second does.

*Proof.* (i) Assume that

$$\left|\frac{a_{n+1}}{a_n}\right| \to r < 1.$$

Choose $\rho$ with $r < \rho < 1$; there exists $N$ such that

$$\left|\frac{a_{n+1}}{a_n}\right| \leq \rho \qquad \text{for all} \qquad n \geq N. \tag{4.8}$$

So then for all $n \geq N$ we have

$$|a_n| \leq \rho^{n-N}|a_N|,$$

and it follows that $\sum a_n$ converges absolutely by the comparison test.

(ii) If $r > 1$ then, fixing $\rho$ with $1 < \rho < r$ there exists $N$ such that

$$\left|\frac{a_{n+1}}{a_n}\right| \geq \rho \qquad \text{for all} \qquad n \geq N. \tag{4.9}$$

Then $|a_n| \geq \rho^{n-N}|a_N| \geq |a_N|$ for all $n \geq N$; so $a_n \not\to 0$, which means that $\sum a_n$ does not converge. $\square$

Note that the proof only really uses the upper/lower bounds on $|a_{n+1}|/|a_n|$ in (4.8) and (4.9); so either of these conditions would be sufficient, and we could write a 'large $n$' forms of the ratio test. [If there exists $N \in \mathbb{N}$ and $\rho$ with $0 < \rho < 1$ such that $|a_{n+1}|/|a_n| \leq \rho$ for all $n \geq N$ then $\sum a_n$ converges absolutely.]

E.g. Take $a_n = \frac{x^n}{n!}$. Then

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{x^{n+1}}{(n+1)!}\frac{n!}{x^n}\right| = \frac{|x|}{n+1} \to 0$$

as $n \to \infty$. So $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converges absolutely for any $x \in \mathbb{R}$. [This series *defines* $\exp(x) = \mathrm{e}^x$.]

E.g. Take $a_n = \frac{n!(2n)!}{(3n)!}$. Then

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!(2n+2)!}{(3n+3)!}\frac{(3n)!}{n!(2n)!} = \frac{(n+1)(2n+1)(2n+2)}{(3n+3)(3n+2)(3n+1)} \to \frac{4}{27} < 1,$$

so $\sum \frac{n!(2n)!}{(3n)!} < \infty$.

**Lemma 4.13** (Cauchy's root test)**.** *Suppose that*

$$|a_n|^{1/n} \to r \qquad \textit{as} \qquad n \to \infty.$$

*Then*

- *if $r < 1$ the series $\sum a_n$ converges absolutely;*
- *if $r > 1$ the series $\sum a_n$ does not converge; and*
- *if $r = 1$ it could be either.*

The proof is very similar to that of the Ratio Test, see Examples 3. Again, the two examples $a_n = 1/n$ and $a_n = 1/n^2$ show that if $r = 1$ we cannot decide on the convergence.

If we took $a_n = x^n/n!$ then we would have $|a_n|^{1/n} = |x|/(n!)^{1/n}$; since $(n!)^{1/n} \to \infty$ as $n \to \infty$ (see Examples 3), it follows that this $n$th root $\to 0$ as $n \to \infty$, so the series converges absolutely (which we already showed using the ratio test). [This raises the question how fast $n!$ grows; we will give an estimate below in Section 4.7.2.]

## 4.7 Comparing series and integrals

Although you will not cover integration rigorously until Analysis II, arguments that bound certain series by integrals are extremely useful; and since you do actually know how to integrate (even if not exactly why it 'works') it would be artificial not to cover this approach now.

We say that a function is decreasing if $f(x) \le f(y)$ whenever $x \ge y$.

**Lemma 4.14** (Integral test)**.** *Suppose that $f\colon [1,\infty) \to [0,\infty)$ is a non-negative decreasing function. Then*

- *if $\int_1^n f(x)\,\mathrm{d}x$ is bounded then $\sum_{n=1}^\infty f(n)$ converges;*
- *if $\int_1^n f(x)\,\mathrm{d}x$ is unbounded then $\sum_{n=1}^\infty f(n)$ diverges.*

*Proof.* (i) Since $f$ is decreasing we have

$$f(n) \le f(x) \qquad \text{for all} \quad x \in [n-1, n];$$

so, integrating both sides between $n-1$ and $n$,

$$f(n) = \int_{n-1}^n f(n)\,\mathrm{d}x \le \int_{n-1}^n f(x)\,\mathrm{d}x.$$

It follows that

$$\sum_{k=2}^n f(k) \le \sum_{k=2}^n \int_{k-1}^k f(x)\,\mathrm{d}x = \int_1^n f(x)\,\mathrm{d}x,$$

see Figure 4.1. If the right-hand side is bounded the sum converges by Lemma 4.5.

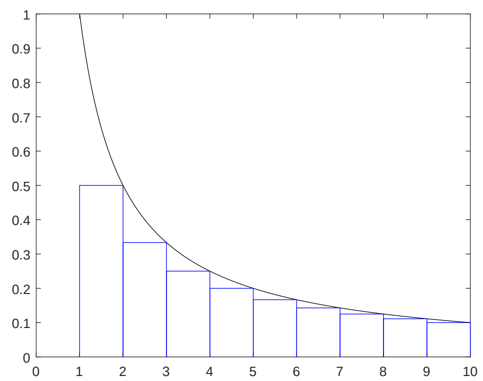


Figure 4.1: Bounding the sum from above using an integral.

(ii) Conversely, for a lower bound we use the fact that

$$f(n) \ge f(x) \qquad \text{for all} \quad x \in [n, n+1];$$

integrating both sides between $n$ and $n+1$ we obtain

$$f(n) = \int_n^{n+1} f(x)\,\mathrm{d}x \geq \int_n^{n+1} f(x)\,\mathrm{d}x.$$

Therefore

$$\sum_{k=1}^{n} f(k) \geq \sum_{k=1}^{n} \int_k^{k+1} f(x)\,\mathrm{d}x = \int_1^{n+1} f(x)\,\mathrm{d}x,$$

see Figure4.2. If the right-hand side is unbounded, so is the left-hand side and the sum diverges (see Lemma 4.5).

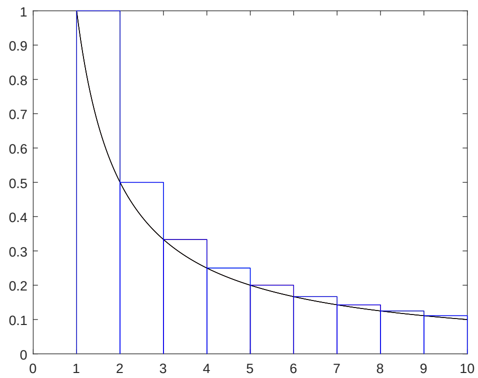


Figure 4.2: Bounding the sum from below using an integral.

□

Examples:

<u>$\sum n^{-\alpha}$ converges if $\alpha > 1$ and diverges if $\alpha \leq 1$.</u>

Assume that $\alpha \neq 1$ (we will treat this case soon). The function $x \mapsto x^{-\alpha}$ is non-negative and decreasing. We have

$$\int_1^n x^{-\alpha}\,\mathrm{d}x = \left[\frac{1}{1-\alpha} x^{1-\alpha}\right]_1^n = \frac{1}{\alpha-1}(1 - n^{-(\alpha-1)}).$$

This is bounded by $1/(\alpha-1)$ if $\alpha > 1$, but tends to $\infty$ as $n \to \infty$ if $\alpha < 1$.

### 4.7.1 Euler's constant

How fast does $\sum_{k=1}^{n} \frac{1}{n}$ grow?

We could use the integral test to give a third proof that this sum is infinite, since[2]

$$\sum_{k=1}^{n} \frac{1}{k} \geq \int_1^{n+1} \frac{1}{x}\,\mathrm{d}x = \log(n+1). \tag{4.10}$$

Let us define[3]

$$\gamma_n := \sum_{k=1}^{n} \frac{1}{k} - \int_1^{n} \frac{1}{x}\,\mathrm{d}x = \sum_{k=1}^{n} \frac{1}{k} - \log n.$$

We will show that $\gamma_n$ converges to some $\gamma \in \mathbb{R}$ as $n \to \infty$, by showing that $(\gamma_n)$ is decreasing and bounded below.

In fact, we have

$$\gamma_{n+1} - \gamma_n = \frac{1}{n+1} - \int_n^{n+1} \frac{1}{x}\,\mathrm{d}x < 0,$$

so $\gamma_n$ is decreasing. Use (4.10) in the form $\int_1^n \frac{1}{x}\,\mathrm{d}x \leq \sum_{k=1}^{n-1} \frac{1}{k}$ we have

$$\gamma_n \geq \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n-1} \frac{1}{k} = \frac{1}{n} \geq 0,$$

so $\gamma_n$ is bounded below.

It follows that $\gamma_n$ converges to some limit, which we call $\gamma$, as $n \to \infty$. It is not known whether $\gamma \approx 0.5772\ldots$ is rational or irrational.

The fact that $\gamma_n$ converges can be used to evaluate $\sum (-1)^{n+1}\frac{1}{n}$, see Question 13 on Examples 3.

### 4.7.2 Bounds on the factorial using comparison with integrals

To bound $n!$ above and below we use logarithms and comparison with integrals.

**Lemma 4.15** (Bounds on $n!$). *For all $n \geq 1$ we have*

$$n^n \mathrm{e}^{-n+1} \leq n! \leq n^{n+1} \mathrm{e}^{-n+1}.$$

*Proof.* We start by turning $n!$ into a sum by taking logs: we have

$$\log n! = \log 2 + \log 3 + \cdots + \log(n-1) + \log n$$

2 Forget ln! The only logarithm worth anything is the 'natural logarithm', the functional inverse of $x \mapsto \mathrm{e}^x$. This is the one true log; from henceforth, log = ln, and there is no log but log (to the base e).

3 The same argument works with the definition from lectures, replacing $\log n$ with $\log(n+1)$, which perhaps fits in more neatly with the lower bound in (4.10). The resulting value of $\gamma$ is the same: $\lim_{n\to\infty}[\log n - \log(n+1)] = \lim_{n\to\infty} \int_n^{n+1} \frac{1}{x}\,\mathrm{d}x$, which is zero since $0 < \int_n^{n+1} \frac{1}{x}\,\mathrm{d}x < \frac{1}{n}$.

since $\log 1 = 0$. Then since we have

$$\log k \leq \log x \qquad x \in [k, k+1]$$

we have

$$\begin{aligned}\log n! &= (\log 2 + \cdots + \log(n-1)) + \log n \\ &\leq \left(\int_1^n \log x \, \mathrm{d}x\right) + \log n \\ &= [x \log x - x]_1^n + \log n \\ &= (n+1)\log n - n + 1.\end{aligned}$$

See Figure 4.3. Taking exponentials of both sides we obtain

$$n! \leq n^{n+1} \mathrm{e}^{-n+1}.$$

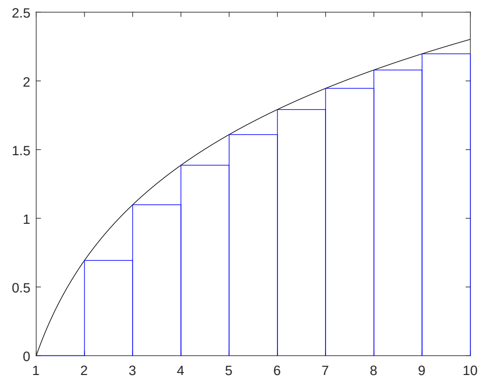


Figure 4.3: Bounding $\log(n!)$ from above using an integral.

For a lower bound, we have

$$\begin{aligned}\log n! &= \log 2 + \log 3 + \cdots + \log(n-1) + \log n \\ &\geq \int_1^n \log x \, \mathrm{d}x \\ &= n \log n - n + 1,\end{aligned}$$

which gives

$$n! \geq n^n \mathrm{e}^{-n+1}.$$

So we have

$$n^n \mathrm{e}^{-n+1} \leq n! \leq n^{n+1} \mathrm{e}^{-n+1},$$

as claimed. □

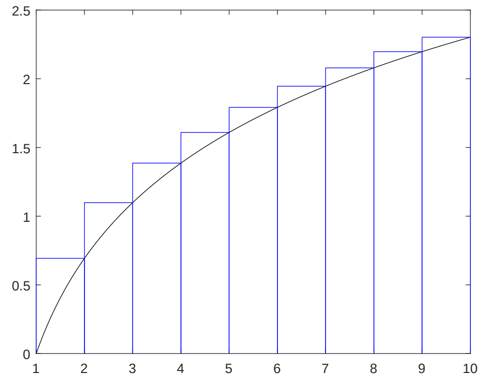


Figure 4.4: Bounding $\log(n!)$ from below using an integral.

We can use these bounds to evaluate limits: e.g.

$$\frac{n!2^n}{n^n} \leq \frac{n^{n+1}\mathrm{e}^{-n+1}2^n}{n^n} = \frac{1}{\mathrm{e}}n\left(\frac{2}{\mathrm{e}}\right)^n;$$

since $2/\mathrm{e} < 1$ it follows that $n!2^n/n^n$ tends to zero (see Proposition 2.18 (iii), using the fact that if $x_n \to \infty$ then $1/x_n \to 0$). Another example,

$$\frac{n!4^n}{n^n} \geq \frac{n^n\mathrm{e}^{-n+1}4^n}{n^n} = \frac{1}{e}\left(\frac{4}{\mathrm{e}}\right)^n;$$

since $4/\mathrm{e} > 1$ it follows that $n!4^n/n^n \to \infty$ as $n \to \infty$.

### 4.7.3 Alternating series test

We showed earlier that $\sum(-1)^{n+1}\frac{1}{n}$ converges. We now use a similar argument to prove the 'Alternating Series Test'.

**Lemma 4.16.** *Suppose that $a_n \geq 0$ with $a_{n+1} \leq a_n$ and $a_n \to 0$. Then*

$$\sum_{n=1}^{\infty}(-1)^{n+1}a_n$$

*converges.*

In this series we add odd terms and subtract even terms: the sum is $a_1 - a_2 + a_3 - a_4 + \cdots$

*Proof.* Let

$$s_k = \sum_{n=1}^{k}(-1)^{n+1}a_n.$$

We have $s_{2k} \leq s_{2k+1}$, since $s_{2k+1} = s_{2k} + a_{2k+1}$.

Notice that

$$s_{2k+2} - s_{2k} = \sum_{n=1}^{2k+2} (-1)^{n+1} a_n - \sum_{n=1}^{2k} (-1)^{n+1} a_n = a_{2k+1} - a_{2k+2} \geq 0;$$

so $s_{2k}$ is an increasing sequence. In particular $s_{2k} \geq s_2$ for every $k$.

Similarly

$$s_{2k+1} - s_{2k-1} = \sum_{n=1}^{2k+1} (-1)^{n+1} a_n - \sum_{n=1}^{2k-1} (-1)^{n+1} a_n = -a_{2k} + a_{2k+1} \leq 0;$$

so $s_{2k+1}$ is an increasing sequence. In particular, $s_{2k+1} \leq s_1$ for every $k$.

It follows that $s_{2k}$ is an increasing sequence, with $s_{2k} \leq s_{2k+1} \leq s_1$ for all $k$, i.e. it is bounded above. So $s_{2k} \to l$ for some $l \in \mathbb{R}$.

Now by the sum rule, since

$$s_{2k+1} = s_{2k} + a_{2k+1}$$

and $a_{2k+1} \to 0$ as $k \to \infty$, it follows that $s_{2k+1} \to l$ also.

It now follows by the result of Q2 on Examples 2 that $s_k \to l$. □

This result clearly applies to $\sum (-1)^{n+1} \frac{1}{n}$. Question 13 on Examples 2 outlines the proof that the value of this sum is $\log 2$.

## 4.8 Rearrangements

We now investigate the behaviour of rearrangements of series. The idea is that we keep the same terms, but change their order. To do this properly, we let $\sigma \colon \mathbb{N} \to \mathbb{N}$ be a bijection (one-to-one and onto). Then we want to compare

$$\sum_{n=1}^{\infty} a_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} a_{\sigma(n)}.$$

**Lemma 4.17.** *If $a_n \geq 0$ and $\sum a_n < \infty$ then every rearrangement has the same limit.*

*Proof.* Let $b_n$ be a rearrangement, so that $b_n = a_{\sigma(n)}$.

For each $N$, let $M_N = \max\{\sigma(n) : n = 1, \ldots, N\}$: the first $N$ terms of $(b_n)$ all occur within the first $M_N$ terms of $a_n$. So

$$\sum_{n=1}^{N} b_n \leq \sum_{n=1}^{M_N} a_n \leq \sum a_n.$$

So the partial sums of $\sum b_n$ are increasing and bounded above by $\sum a_n$: so $\sum b_n$ converges and $\sum b_n \leq \sum a_n$.

If we let $L_N = \max\{\sigma^{-1}(r) : r = 1, \ldots, N\}$ then the first $N$ terms of $(a_n)$ are included in the first $L_N$ terms of $(b_n)$. So $\sum_{j=1}^{N} a_j \leq \sum_{j=1}^{L_N} b_j$. Arguing as before we see that $\sum a_n \leq \sum b_n$. □

From this we can readily deduce a much more powerful result.

**Corollary 4.18.** *If $\sum a_n$ is absolutely convergent then every rearrangement has the same limit.*

*Proof.* We use the same trick that we did in the proof that absolutely convergent sequences converge (Lemma 4.10): we define two new sequences with positive terms,

$$b_n = \begin{cases} a_n & a_n > 0, \\ 0 & a_n \leq 0; \end{cases} \qquad c_n = \begin{cases} 0 & a_n \geq 0, \\ -a_n & a_n < 0. \end{cases}$$

Note that $b_n \leq |a_n|$, $c_n \leq |a_n|$, and $a_n = b_n - c_n$. We showed

By Lemma 4.17 we have

$$\sum_{n=1}^{\infty} b_{\sigma(n)} = \sum_{n=1}^{\infty} b_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} c_{\sigma(n)} = \sum_{n=1}^{\infty} c_n.$$

Since $a_{\sigma(n)} = b_{\sigma(n)} - c_{\sigma(n)}$ it follows that

$$\sum a_{\sigma(n)} = \sum b_{\sigma(n)} - \sum c_{\sigma(n)} = \sum b_n - \sum c_n = \sum a_n.$$ □

If $\sum a_n$ converges but $\sum |a_n|$ does not converge then recall that we say that $\sum a_n$ is *conditionally convergent.*

Suppose that $\sum a_n$ is conditionally convergent: it follows by the sum rule that in this case both $b_n$ and $c_n$ (as define above) must diverge to $+\infty$ (if one converges, then by the sum rule the other must converge too; and then it would follow, since $|a_n| = b_n + c_n$, then $\sum |a_n|$ would converge, i.e. $\sum a_n$ would converge absolutely).

**Theorem 4.19** (Riemann Rearrangement Theorem)**.** *If $\sum a_n$ is conditionally convergent, then it can be rearranged to converge to any real number $l$.*

*Proof.* Let $(p_n)$ and $(q_n)$ be the subsequences of $(a_n)$ containing all the positive and negative terms, respectively. Note that since $\sum a_n$ converges, $a_n \to 0$, which implies that $p_n \to 0$ and $q_n \to 0$ also.

Suppose that $l > 0$.

We make a new sequence that includes all the terms from $(p_n)$ and $(q_n)$, so all the terms from $(a_n)$, and make sure that it converges to $l$.

We start by including the smallest number of terms from $(p_n)$ to ensure that

$$S_1 := \sum_{j=1}^{N_1} p_j > l.$$

This means that

$$S_1 = \sum_{j=1}^{N_1-1} p_j + p_{N_1} \leq l + p_{N_1} \qquad \Rightarrow \qquad l < S_1 \leq l + p_{N_1}.$$

We now include the minimum number of negative terms from $(q_n)$ to make the sum less than $l$ once again: we choose the smallest $M_1$ such that

$$T_1 := S_1 + \sum_{i=1}^{M_1} q_i < l \qquad \Rightarrow \qquad l + q_{M_1} \leq T_1 < l.$$

We continue in this way, choosing sums that are (after the required number) greater than or less than $l$. The sequence we obtain like this,

$$p_1, p_2, \ldots, p_{N_1}, q_1, q_2, \ldots, q_{M_1}, p_{N_1+1}, \ldots, p_{N_2}, q_{M_1+1}, \ldots, q_{M_2}, p_{N_2+1}, \ldots$$

is a rearrangement of the original series.

We know that

$$l < S_i \leq l + p_{N_i} \qquad \text{and} \qquad l + q_{M_i} \leq T_i < l.$$

Since $N_i \to \infty$ and $M_i \to \infty$, it follows that $S_i \to l$ and $T_i \to l$. To see that the rearranged sum tends to $l$, note that the sum of the series between $S_i$ and $T_i$ is decreasing, so

$$l + q_{M_i} \leq \sum \leq l + p_{N_i}$$

and between $T_i$ and $S_{i+1}$ is increasing, so

$$l + q_{M_i} \leq \sum \leq l + p_{N_{i+1}}.$$

□

# Part IV

# Continuous functions

# Chapter 5

# Continuity

## 5.1 Functions of one real variable

We consider functions defined on a subset $E$ of $\mathbb{R}$, usually an interval (we will have more to say about intervals later), for example $f\colon E \to \mathbb{R}$. It is useful to remember that such a function is a rule that assigns a real number to every element of $E$; so, strictly speaking, '$x^2$' is not a function, but '$x \mapsto x^2$' is (this may seem a little pedantic, but is worth bearing in mind).

For a function $f\colon E \to \mathbb{R}$, we say that $f$ is defined on $E$, or that $E$ is the *domain* of $f$. The *range* or *image* of $f$ is

$$f(E) = \{y : \ y = f(x) \text{ for some } x \in E\}.$$

Here are some examples that we will return to later:

- $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $f(x) = 1/x$. [Cannot define $f$ on all of $\mathbb{R}$.]
- $f\colon \mathbb{R} \to \mathbb{R}$ defined by
$$f(x) = \begin{cases} x \sin(1/x) & x \neq 0, \\ 0 & x = 0. \end{cases}$$
[The factor of $x$ damps the 'wild oscillations' coming from $\sin(1/x)$ near $x = 0$.] The range of this function is $[\alpha^{-1} \sin \alpha, 1)$, where $\alpha$ is the smallest positive root of $\alpha = \tan \alpha$ (neither end of the range is entirely obvious); see Figure 5.1.
- $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \lfloor x \rfloor$ is the largest integer $\leq x$ (the function 'jumps' at every integer).

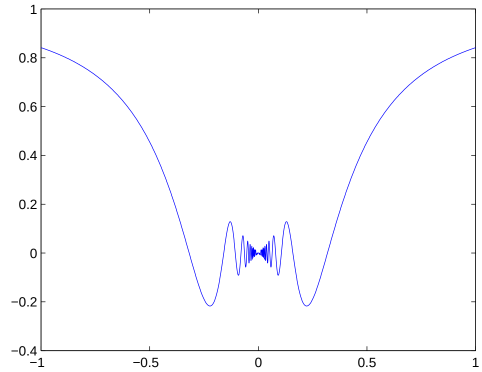


Figure 5.1: The graph of $x \sin(1/x)$ between $x = -1$ and $x = +1$.

- Thomae's function $f\colon (0,1) \to \mathbb{R}$ defined by setting

$$f(x) = \begin{cases} 1/q & x = p/q \text{ in lowest form, } p \geq 0 \\ 0 & x \notin \mathbb{Q}. \end{cases} \tag{5.1}$$

  Range is $\{0\} \cup \{1/q : \ q \in \mathbb{N}\}$. See Figure 5.2.

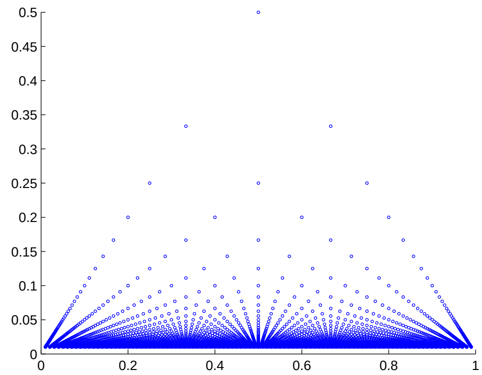


Figure 5.2: The graph of the function defined in (5.1).

- Functions can be defined 'piecewise', e.g. $E = [0, \infty)$,

$$f(x) = \begin{cases} \sin x & 0 \leq x \leq \pi/2; \\ 2x/\pi & \pi/2 < x \leq \pi; \\ 2 & \pi < x \leq 4; \\ 2 - 3(x-4)^2 & x > 4. \end{cases} \tag{5.2}$$

  See Figure 5.3. This is often a good way to construct counterexamples, as is specifying a function by a sketch.

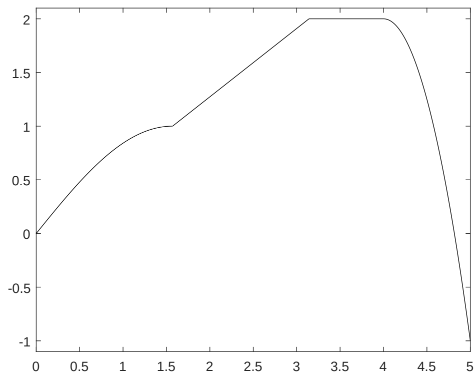


Figure 5.3: The graph of the function defined piecewise in (5.2).

## 5.2 Continuity

We say that $f$ is continuous at $c$ if $f(x)$ can be made arbitrarily close to $f(c)$ by taking $x$ sufficiently close to $c$, see Figure 5.4.

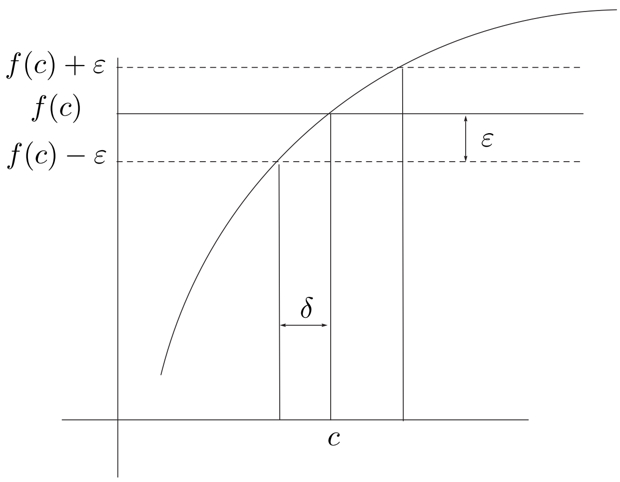


Figure 5.4: Illustration of the $\varepsilon$–$\delta$ definition of continuity

**Definition 5.1** ($\varepsilon$–$\delta$ definition of continuity)**.** A function $f\colon E \to \mathbb{R}$ is *continuous* at $c \in E$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$x \in E \quad \text{and} \quad |x-c| < \delta \qquad \Rightarrow \qquad |f(x) - f(c)| < \varepsilon. \tag{5.3}$$

If $f$ is continuous at every point of $E$ we say that $f$ is continuous on $E$. If $E = \mathbb{R}$ and $f$ is continuous at every point of $\mathbb{R}$ we often just say that $f$ is continuous.

Given a point $c$, for any 'challenge' $\varepsilon$ we have to find a $\delta$ that works in (10.2). In some cases it might be interesting to try to find the 'largest' $\delta$ (there might not always be a largest one, see below), but in general we just need some $\delta$ that works.

Some further remarks on the definition:

- If $\delta$ works in (10.2) then so does any $\delta' < \delta$.
- If $E$ is an open interval, $E = (a, b)$ then for any $c \in (a, b)$, if $\delta$ is sufficiently small then $|x - c| < \delta$ implies that $x \in E$ (take $\delta \leq \min(c - a, b - c)$).
- If $E = [a, b]$ then at $c = a$ the condition is in fact (for small $\delta$)

  $$0 \leq x - a < \delta$$

  (and a similar one-sided inequality if $c = b$).

The following definition arises naturally in the context of the above.

**Definition 5.2.** We call $x \in E$ with $|x-c| < \delta$, i.e. $E \cap (c-\delta, c+\delta)$, the *$\delta$-neighbourhood* of $c$ in $E$. Figure 5.5 illustrates some $\delta$-neighbourhoods.

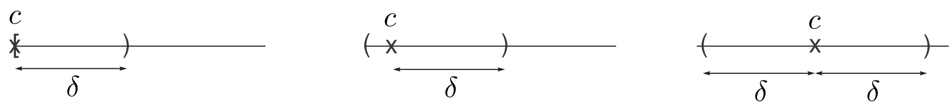


Figure 5.5: Different $\delta$-neighbourhoods of $c$ for the same value of $\delta$. In the left-hand picture $c$ is the left-hand endpoint of a closed interval $E$. In the middle picture $c$ is close to the left-hand endpoint of an open interval $E$. If $\delta$ was smaller the interval $(c-\delta, c+\delta)$ would be entirely contained in $E$, as in the right-hand picture.

**Example 5.3.** The function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = \ell$, $\ell$ a constant, is continuous at every $c \in \mathbb{R}$.

Take $c \in \mathbb{R}$; then for any $x \in \mathbb{R}$ we have

$$|f(x) - f(c)| = |\ell - \ell| = 0.$$

Whatever $\varepsilon > 0$ we take, we have $|f(x) - f(c)| = 0 < \varepsilon$, so we can take any $\delta > 0$ in (10.2) [so in this case there is not a 'largest' $\delta$]. Since $f$ is continuous at every $c \in \mathbb{R}$ we can say that '$f$ is continuous'.

This example is perhaps so simple that it fails to illustrate the definition adequately. The next most complicated example would probably be $x \mapsto x$, but making this a little more general gives a more useful example.

**Example 5.4.** For any $\alpha \in \mathbb{R}$ the function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = \alpha x$ is continuous.

If $\alpha = 0$ the $f(x) = 0$ for all $x$, i.e. $f$ is constant and so continuous by the previous example. So we can assume that $\alpha \neq 0$.

Take $c, x \in \mathbb{R}$ then

$$|f(x) - f(c)| = |\alpha x - \alpha c| = |\alpha||x - c|.$$

Given $\varepsilon > 0$, if we take $\delta = \varepsilon/|\alpha|$ and then whenever $|x - c| < \delta$ we have

$$|f(x) - f(c)| = |\alpha||x - c| < |\alpha|\frac{\varepsilon}{|\alpha|} = \varepsilon.$$

In both these examples, given any $\varepsilon$ we can take the same $\delta$ for every choice of $c$. This is not usually true.

**Example 5.5.** The function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is continuous.

Take some $c \in \mathbb{R}$; then for any $x \in \mathbb{R}$ we have

$$|f(x) - f(c)| = |x^2 - c^2| = |(x + c)(x - c)| = |x + c||x - c|.$$

If $|x - c| < 1$ then

$$|x + c| = |(x - c) + 2c| \leq |x - c| + 2|c| < 1 + 2|c|,$$

and so if $|x - c| < 1$ we have

$$|x^2 - c^2| < (1 + 2|c|)|x - c|. \tag{5.4}$$

Given any $\varepsilon > 0$, we now choose $\delta := \min(1, \varepsilon/(1 + 2|c|)$; then if $|x - c| < \delta$ we have $|x - c| < 1$ so that (5.4) is valid; and then

$$|x^2 - c^2| < (1 + 2|c|)\delta = \varepsilon,$$

as required.

**Example 5.6.** The function $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $f(x) = 1/x$ is continuous at every $c \neq 0$.

Take $c \in \mathbb{R} \setminus \{0\}$. First, observe that for every $x \neq 0$ we have

$$\left|\frac{1}{x} - \frac{1}{c}\right| = \left|\frac{c - x}{xc}\right| = \frac{|x - c|}{|xc|}.$$

Now, notice that since $c \neq 0$, if $|x - c| < \delta_1 := |c|/2$ then we have

$$|x| = |x - c + c| \geq ||c| - |x - c|| > |c| - |c|/2 = |c|/2,$$

using the reverse triangle inequality. So for $|x - c| < \delta_1$

$$\left|\frac{1}{x} - \frac{1}{c}\right| < \frac{|x - c|}{|c|^2/2}.$$

Now, given any $\varepsilon > 0$, if we also have $|x - c| < \delta_2 := |c|^2\varepsilon/2$ then

$$\frac{|x - c|}{|c|^2/2} < \varepsilon.$$

So if we take $\delta = \min(\delta_1, \delta_2)$ it follows that

$$x \neq 0, \ |x - c| < \delta \qquad \Rightarrow \qquad \left|\frac{1}{x} - \frac{1}{c}\right| < \varepsilon$$

and so $f$ is continuous at $c$ as claimed. Note that $\delta$ now has to be much smaller when $c$ is small. [When $c$ is large you can actually use a much larger value of $\delta$ than our argument here shows.]

Here is a first general result about continuous functions. In the particular case $\alpha = 0$ this is about 'preservation of sign'.

**Lemma 5.7** (Preservation of inequalities)**.** *If $f\colon E \to \mathbb{R}$ is continuous at $c \in E$ and $f(c) > \alpha$ then there exists a $\delta > 0$ such that $f(x) > \alpha$ for all $x \in E$ with $|x - c| < \delta$. Similarly, if $f(c) < \alpha$ then there exists a $\delta > 0$ such that $f(x) < \alpha$ for all $x \in E$ with $|x - c| < \delta$.*

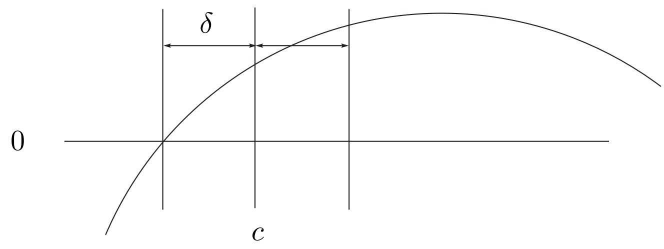


Figure 5.6: If $f$ is continuous at $c$ and $f(c) > 0$ then $f > 0$ in a $\delta$-neighbourhood of $c$.

*Proof.* Suppose that $f(c) > \alpha$. Apply the definition of continuity with $\varepsilon = f(c) - \alpha$. Then there exists a $\delta > 0$ such that if $x \in E$ and $|x - c| < \delta$ then

$$|f(x) - f(c)| < f(c) - \alpha$$

and so

$$f(x) = f(c) + (f(x) - f(c)) \geq f(c) - |f(x) - f(c)| > \alpha.$$

The proof for $f(c) < 0$ is similar, but with $\varepsilon = \alpha - f(c)$. $\square$

## 5.3 Algebra of continuous functions

**Proposition 5.8.** *If $f, g\colon E \to \mathbb{R}$ are both continuous at $c \in E$ then*

*(i) $f + g$ is continuous at $c$;*

*(ii) $fg$ is continuous at $c$.*

[The function $f + g\colon E \to \mathbb{R}$ is the function defined by $(f + g)(x) = f(x) + g(x)$ and the function $fg\colon E \to \mathbb{R}$ is the function defined by $(fg)(x) = f(x)g(x)$. Note that $fg$ *does not* mean the composition of $f$ with $g$, which we write (see below) as $f \circ g$.]

*Proof.* (i) Since

$$\begin{aligned} |(f+g)(x) - (f+g)(c)| &= |f(x) - f(c) + g(x) - g(c)| \\ &\leq |f(x) - f(c)| + |g(x) - g(c)| \end{aligned}$$

we find $\delta_1$ and $\delta_2$ such that $|x - c| < \delta_1$ implies that $|f(x) - f(c)| < \varepsilon/2$ and $|x - c| < \delta_2$ implies that $|g(x) - g(c)| < \varepsilon/2$. Take $\delta = \min(\delta_1, \delta_2)$. Then if $|x - c| < \delta$ we have

$$|(f+g)(x) - (f+g)(c)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

(ii) Add and subtract $f(c)g(x)$ and use the triangle inequality.

$$\begin{aligned} |(fg)(x) - (fg)(c)| &= |f(x)g(x) - f(c)g(c)| \\ &= |f(x)g(x) - f(c)g(x) + f(c)g(x) - f(c)g(c)| \\ &\leq |f(x)g(x) - f(c)g(x)| + |f(c)g(x) - f(c)g(c)| \\ &\leq |g(x)||f(x) - f(c)| + |f(c)||g(x) - g(c)|. \end{aligned}$$

We treat the second term first, since $|f(c)|$ does not vary ($c$ is fixed). We can use the continuity of $g$ at $c$ to find a $\delta_1 > 0$ such that

$$x \in E,\ |x - c| < \delta_1 \qquad \Rightarrow \qquad |g(x) - g(c)| < \frac{\varepsilon}{2(1 + |f(c)|)},$$

and then certainly (even if $|f(c)| = 0$)

$$x \in E,\ |x - c| < \delta_1 \quad \Rightarrow \quad |f(c)||g(x) - g(c)| < \frac{|f(c)|}{1 + |f(c)|}\frac{\varepsilon}{2} < \frac{\varepsilon}{2}. \tag{5.5}$$

To deal with the first term we use the continuity of $g$ at $c$ to control $|g(x)|$, and then the argument is essentially the same as the second term. Since $g$ is continuous at $c$, we can take $\varepsilon = 1$ in the definition of continuity and find a $\delta_2 > 0$ such that

$$x \in E,\ |x - c| < \delta_3 \quad \Rightarrow \quad |g(x) - g(c)| < 1 \quad \Rightarrow \quad |g(x)| < |g(c)| + 1. \tag{5.6}$$

Now we use the fact that $f$ is continuous at $c$ to choose a $\delta_2$ such that

$$x \in E,\ |x - c| < \delta_2 \quad \Rightarrow \quad |f(x) - f(c)| < \frac{\varepsilon}{2(|g(c)| + 1)}. \tag{5.7}$$

Then combining (5.6) and (5.7) it follows that

$$\begin{aligned} &x \in E,\ |x - c| < \min(\delta_2, \delta_3) \\ &\qquad \Rightarrow \quad |g(x)||f(x) - f(c)| < (|g(c)| + 1)\,\frac{\varepsilon}{2(|g(c)| + 1)} = \frac{\varepsilon}{2}. \end{aligned} \tag{5.8}$$

Now set $\delta = \min(\delta_1, \delta_2, \delta_3)$, and then combing (5.8) and (5.5) we have shown that

$$x \in E,\ |x - c| < \delta \qquad \Rightarrow \qquad |(fg)(x) - (fg)(c)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required. □

Some new continuous functions, all proved using induction and sum/product properties of continuity:

- $x \mapsto x^n$ is continuous on $\mathbb{R}$ for any $n \in \mathbb{N}$;
- all polynomials
$$P(x) = \sum_{k=0}^{n} a_k x^k$$
are continuous on $\mathbb{R}$ (for any $n$ and choice of $a_k$);
- $x^{-n}$ is continuous on $\mathbb{R} \setminus \{0\}$ for any $n \in \mathbb{N}$;
- the expression
$$\sum_{k=-n}^{n} a_k x^k$$
is continuous on $\mathbb{R} \setminus \{0\}$ (for any $n$ and choice of $a_k$).

What about a function like $\exp(x)$ defined by a power series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}\ ?$$

This function *is* continuous, but we cannot prove this yet (you will see a proof in Analysis II). Sine and cosine (which can also be defined by power series) are also continuous; we will prove this shortly using their more geometric definition.

**Proposition 5.9** (Composition of continuous functions)**.** *Suppose that $f\colon E \to \mathbb{R}$ is continuous at $c$, $f(E) \subset D$, and $g\colon D \to \mathbb{R}$ is continuous at $f(c)$; then the composition of $g$ with $f$, $g \circ f\colon E \to \mathbb{R}$, defined by setting*

$$(g \circ f)(x) = g(f(x)) \qquad x \in E,$$

*is continuous at $c$.*

*Proof.* Take $\varepsilon > 0$. Using continuity of $g$ at $f(c)$, there exists a $\delta_1 > 0$ such that

$$y \in D,\ |y - f(c)| < \delta_1 \qquad \Rightarrow \qquad |g(y) - g(f(c))| < \varepsilon. \tag{5.9}$$

Using continuity of $f$ at $c$, we take $\varepsilon = \delta_1$; there exists a $\delta_2 > 0$ such that

$$|x - c| < \delta_2 \quad \Rightarrow \quad |f(x) - f(c)| < \delta_1.$$

Since $f(x) \in f(E) \subseteq D$ it now follows that

$$|x - c| < \delta_2 \ \Rightarrow\ f(x) \in D,\ |f(x) - f(c)| < \delta_1 \ \Rightarrow\ |g(f(x)) - g(f(c))| < \varepsilon$$

by (5.9); therefore $g \circ f$ is continuous at $c$, as required. $\square$

**Corollary 5.10.** *If $f\colon E \to \mathbb{R}$ is continuous at $c$ and $f(c) \neq 0$ then $1/f$ is well-defined in some $\delta$-neighbourhood of $c$ in $E$ and is continuous at $c$.*

*Proof.* We have already shown in Lemma 5.7 that if $f(c) \neq 0$ then $f \neq 0$ in some $\delta$-neighbourhood $N$ of $c$ ($N = \{x \in E : |x-c| < \delta\}$ for some $\delta > 0$). Thus $f\colon N \to \mathbb{R}\setminus\{0\}$, and we compose this with $h\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $h(x) = 1/x$, which is certainly continuous at $f(c) \neq 0$. Thus $h \circ f\colon N \to \mathbb{R}$ is continuous, and $h \circ f(x) = 1/f(x)$. $\square$

**Corollary 5.11.** *If $f, g\colon E \to \mathbb{R}$ are continuous at $c \in E$ and $g(c) \neq 0$ then $f/g$ is well-defined in some $\delta$-neighbourhood of $c$ in $E$ and continuous at $c$.*

All rational functions

$$\frac{P(x)}{Q(x)},$$

where $P(x)$ and $Q(x)$ are polynomials, are continuous on $\mathbb{R} \setminus \{x : Q(x) = 0\}$.

**Lemma 5.12.** *Suppose that $f, g\colon E \to \mathbb{R}$ are continuous at $c \in E$. Then*

*(i) the function $|f| : E \to \mathbb{R}$ given by $|f|(x) = |f(x)|$ is continuous at $c$;*

*(ii) the function $\max(f, g) : E \to \mathbb{R}$ given by $\max(f, g)(x) = \max(f(x), g(x))$ is continuous at $c$.*

For a proof see Examples Sheet 4.

# Chapter 6

# Discontinuity and Sequential continuity

We are going to show that we can define continuity by using limits of sequences.

**Lemma 6.1** (Continuity and sequences 1)**.** *Suppose that $f\colon E \to \mathbb{R}$ is continuous at $c \in E$, and that $(x_n) \in E$ is a sequence such that $x_n \to c$ as $n \to \infty$. Then $f(x_n) \to f(c)$ as $n \to \infty$.*

*Proof.* Fix $\varepsilon > 0$. We need to show that there exists an $N \in \mathbb{N}$ such that

$$n \geq N \qquad \Rightarrow \qquad |f(x_n) - f(c)| < \varepsilon.$$

First, observe that since $f$ is continuous at $c$, given any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$x \in E \qquad \text{and} \qquad |x - c| < \delta \qquad \Rightarrow \qquad |f(x) - f(c)| < \varepsilon. \tag{6.1}$$

Now, since $x_n \to c$ there exists an $N \in \mathbb{N}$ such that

$$n \geq N \qquad \Rightarrow \qquad |x_n - c| < \delta. \tag{6.2}$$

Therefore for $n \geq N$ it follows from (6.2) combined with (6.1) that

$$n \geq N \quad \Rightarrow \quad |x_n - c| < \delta \quad \Rightarrow \quad |f(x_n) - f(c)| < \varepsilon,$$

and so $f(x_n) \to f(c)$ as $n \to \infty$ as claimed. $\square$

We make a definition of 'sequentially continuous' based on this result, so that Lemma 6.1 can be summarised as 'continuous implies sequentially continuous'.

**Definition 6.2.** A function $f\colon E \to \mathbb{R}$ is *sequentially continuous* at $c \in E$ if whenever $(x_n) \in E$ and $x_n \to c$ as $n \to \infty$ then $f(x_n) \to f(c)$ as $n \to \infty$.

In order to show that sequentially continuous implies continuous, we will first consider how to formulate exactly when a function fails to be continuous at a point $c$ (we say 'discontinuous at $c$'). We will slowly negate the definition of continuity. We want to end up with a 'positive' statement, i.e. one without the words "it is not true that"

1. It is not true that: for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $x \in E$ and $|x-c| < \delta$ implies that $|f(x) - f(c)| < \varepsilon$.

2. There exists an $\varepsilon > 0$ such that it is not true that: there exists a $\delta > 0$ such that $x \in E$ and $|x-c| < \delta$ implies that $|f(x) - f(c)| < \varepsilon$.

3. There exists an $\varepsilon > 0$ such that for every $\delta > 0$ it is not true that: for every $x \in E$ with $|x-c| < \delta$ we have $|f(x) - f(c)| < \varepsilon$.

4. There exists an $\varepsilon > 0$ such that for every $\delta > 0$ there is an $x \in E$ with $|x-c| < \delta$ but $|f(x) - f(c)| \geq \varepsilon$.

[Rule of thumb: $\forall \leftrightarrow \exists$.]

The following lemma is almost a definition; but it comes from the above negation process, so is actually a consequence of our definition of 'continuous at $c$' and 'discontinuous = not continuous'.

**Lemma 6.3.** *A function $f\colon E \to \mathbb{R}$ is discontinuous at $c \in E$ if there exists an $\varepsilon > 0$ such that for every $\delta > 0$ there exists an $x \in E$ such that*

$$|x-c| < \delta \qquad \textit{but} \qquad |f(x) - f(c)| \geq \varepsilon.$$

We can use this to show that continuity and sequential continuity are equivalent (we prove that sequential continuity implies continuity).

**Lemma 6.4** (Continuity and sequences 2)**.** *Suppose that for any sequence $(x_n) \in E$ with $x_n \to c$ as $n \to \infty$ we have $f(x_n) \to f(c)$ as $n \to \infty$; then $f$ is continuous at $c$.*

*Proof.* The neatest way to prove this is a contrapositive argument: we show that if $f$ is not continuous at $c$ then it is not sequentially continuous at $c$.

If $f$ is not continuous at $c$, then, if for each $n \in \mathbb{N}$ we take $\delta = 1/n$ in Lemma 6.3 we can find an $x_n \in E$ with

$$|x_n - c| < 1/n \qquad \text{and} \qquad |f(x_n) - f(c)| \geq \varepsilon.$$

In this way we have constructed a sequence $(x_n) \in E$ such that $x_n \to c$ but $f(x_n) \not\to f(c)$; so $f$ is not sequentially continuous at $c$. $\square$

We have therefore shown that *a function is continuous at $c$ if and only if it is sequentially continuous at $c$.*

We now have two ways to show that a function is discontinuous at a point: we can find a sequence $(x_n) \in E$ such that $x_n \to c$ but $f(x_n) \not\to f(c)$, or we can use Lemma 6.3 and find an $\varepsilon > 0$ such that for every $\delta > 0$ there is an $x \in E$ with $|x - c| < \delta$ and $|f(x) - f(c)| \geq \varepsilon$.

Some examples using sequences:

**Example 6.5.** The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

is not continuous at $x = 0$.

Take $x_n = 1/n$, then $x_n \to 0$, but $f(x_n) \to 1 \neq 0 = f(0)$.

**Example 6.6.** The function $f(x) = \lfloor x \rfloor$ is discontinuous at every point $c \in \mathbb{Z}$.

Take $x_n = c - [1/(n+1)]$; then $x_n \to c$ but $f(x_n) \to c - 1 \neq c = f(c)$.

**Example 6.7.** Suppose that $g, h\colon \mathbb{R} \to \mathbb{R}$ are both continuous and let

$$f(x) = \begin{cases} g(x) & x \in \mathbb{Q} \\ h(x) & x \notin \mathbb{Q} \end{cases}.$$

Then $f$ is continuous at $c$ if and only if $g(c) = h(c)$.

Suppose that $g(c) \neq h(c)$ and consider the case $c \in \mathbb{Q}$. Then there exists a sequence $x_n \notin \mathbb{Q}$ such that $x_n \to c$; then $\lim_{n\to\infty} f(x_n) = h(x_n) \to h(c) \neq g(c) = f(c)$. A similar argument works if $c \notin \mathbb{Q}$ (taking $(x_n) \in \mathbb{Q}$ with $x_n \to c$).

If $g(c) = h(c)$ then given $\varepsilon > 0$ since $g$ is continuous at $c$ there exists $\delta_1 > 0$ such that

$$|x - c| < \delta_1 \qquad \Rightarrow \qquad |g(x) - g(c)| < \varepsilon$$

and since $h$ is continuous at $c$ there exists $\delta_2 > 0$ such that

$$|x - c| < \delta_2 \qquad \Rightarrow \qquad |h(x) - h(c)| < \varepsilon.$$

Then, with $\delta = \min(\delta_1, \delta_2)$, if $|x - c| < \delta$ we have either $x \in \mathbb{Q}$ and then

$$|x - c| < \delta \leq \delta_1 \qquad \Rightarrow \qquad |f(x) - f(c)| = |g(x) - g(c)| < \varepsilon$$

or $x \notin \mathbb{Q}$ and then

$$|x - c| < \delta \leq \delta_2 \qquad \Rightarrow \qquad |f(x) - f(c)| = |h(x) - h(c)| < \varepsilon;$$

in both cases $|x - c| < \delta$ ensures that $|f(x) - f(c)| < \varepsilon$ so $f$ is continuous at $c$.

Some examples using the $\varepsilon$–$\delta$ approach from Lemma 6.3.

**Example 6.8.** The function $f \colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

is not continuous at $x = 0$.

Let $\varepsilon = 1$. Then for any $\delta > 0$, the point $x = \delta/2$ satisfies

$$|x - 0| = |x| = \delta/2 < \delta \qquad \text{and} \qquad |f(x) - f(0)| = |f(x)| = 1 \geq 1,$$

and so $f$ is not continuous at $x = 0$.

**Example 6.9.** The function $f \colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is discontinuous everywhere.

Take $c \in \mathbb{Q}$ and let $\varepsilon = 1$. Then for any $\delta > 0$ there is an irrational $x$ with $|x - c| < \delta$, and so for which $|f(x) - f(c)| = 1$. Similarly if $c \notin \mathbb{Q}$, again take $\varepsilon = 1/2$ and find a rational $x$ with $|x - c| < \delta$ so that $|f(x) - f(c)| = 1$.

**Example 6.10.** The function $f \colon (0, 1) \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1/q & x \in \mathbb{Q} \text{ with } x = p/q \text{ in its lowest terms,} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

(see Figure 5.2) is discontinuous at every rational and continuous at every irrational.

If $c = p/q \in \mathbb{Q}$ then $f(c) = 1/q$. Now take $\varepsilon = 1/q$. Given any $\delta > 0$ there exists an irrational $x$ with $|x - c| < \delta$, but for which $f(x) = 0$ and $|f(x) - f(c)| = 1/q \geq \varepsilon$.

Now take $c \notin \mathbb{Q}$ and some $\varepsilon > 0$. If $x = p/q \in \mathbb{Q}$ then

$$|f(x) - f(c)| = 1/q;$$

this is $\geq \varepsilon$ only if $x$ has denominator $q \leq 1/\varepsilon$.

Choose some $q' \in \mathbb{N}$ such that $1/q' \leq \varepsilon$. Then within $(0,1)$ there are only finitely many $x$ for which $f(x) \geq 1/q'$ (certainly no more than $q'(q'+1)/2$). So the set

$$A = \{x \in (0,1) : \ |f(x) - f(c)| \geq \varepsilon\}$$

is finite. If $\varepsilon > 1/2$ then this set will be empty; this means that whatever our choice of $\delta$, $x \in (0,1)$ and $|x - c| < \delta$ will imply that $|f(x) - f(c)| < \varepsilon$. Otherwise the set is non-empty, and so there is an element $x^* \in A$ that is closest (but not equal) to $c$: in this case we take $\delta = |x^* - c|$. (The set $\{|x - c| : \ x \in A\}$ is a non-empty collection of strictly positive numbers, so has a least element $\delta$.) Now if $|x - c| < \delta$ then either $x$ is irrational, in which case $f(x) = f(c)$, or $x$ is rational with denominator $q > q' \geq 1/\varepsilon$, whence $1/q < \varepsilon$.

# Chapter 7

# Open sets, closed sets, and intervals

## 7.1 Open and closed sets

We are used to saying that the interval $(a, b)$ is 'open' and that the interval $[a, b]$ is 'closed'; the terminology is not arbitrary. In general, a set is open if every point in the set lies 'inside' the set itself.

**Definition 7.1.** A subset $A$ of $\mathbb{R}$ is *open* if for every $x \in A$ there exists an $\varepsilon > 0$ such that

$$(x - \varepsilon, x + \varepsilon) \subset A.$$

Clearly $\mathbb{R}$ is open, and so is $\varnothing$, since it satisfies the definition (there's no $x \in \varnothing$ so there's nothing to check). Here is what seems like a very trivial lemma, with perhaps just a little more content: open intervals are open. But there is some point to this, since (so it seems) we just decided to call intervals of the form $(a, b)$ 'open', and now we have defined what seems like a more generally applicable property of 'being open'.

**Lemma 7.2.** *The interval $(a, b)$ is open.*

*Proof.* Take $x \in (a, b)$, and let $\varepsilon = \min(x - a, b - x)$. Note that $\varepsilon > 0$ and that $a \leq x - \varepsilon \leq x + \varepsilon \leq b$, so $(x - \varepsilon, x + \varepsilon) \subset (a, b)$ and $(a, b)$ is open as claimed. □

The interval $I = (a, b]$ is *not* open: the point $b$ lies on the 'edge' of the interval - there is no $\varepsilon > 0$ such that $(b - \varepsilon, b + \varepsilon) \subset I$, since $I$ contains no points larger than $b$.

**Lemma 7.3.** *Suppose that $A$ and $B$ are open subsets of $\mathbb{R}$. Then $A \cup B$ and $A \cap B$ are open.*

*Proof.* Suppose that $x \in A \cup B$; if $x \in A$ then for some $\varepsilon > 0$ we have $(x - \varepsilon, x + \varepsilon) \subset A \subset A \cup B$, and similarly if $x \in B$. So $A \cup B$ is open.

If $A \cap B$ is empty then it is open. Otherwise, if $x \in A \cap B$ then $x \in A$, so for some $\varepsilon_1 > 0$ we have $(x - \varepsilon_1, x + \varepsilon_1) \in A$, and $x \in B$ and so for some $\varepsilon_2 > 0$ we have $(x - \varepsilon_2, x + \varepsilon_2) \in B$. If we take $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ then $(x - \varepsilon, x + \varepsilon) \in A \cap B$ and $A \cap B$ is open as claimed. $\square$

It follows by induction that any finite union of open sets is open, and any finite intersection of open sets is open. In fact any countable union of open sets is open (see Examples 4) but the result for intersections cannot be extended to countable collections, since

$$\bigcap_{n=1}^{\infty} \left( a - \frac{1}{n}, a + \frac{1}{n} \right) = \{a\},$$

and any single point is not open.

A set is closed if its complement is open.

**Definition 7.4.** A subset $A$ of $\mathbb{R}$ is *closed* if $\mathbb{R} \setminus A$ is open.

Note that this means that $\mathbb{R}$ and $\emptyset$ are closed (so they are both open and closed). The definition allows for another 'very trivial' lemma.

**Lemma 7.5.** *The interval* $[a, b]$ *is closed.*

*Proof.* We have $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$. This is the union of two open sets, and so open. $\square$

However, there is a more useful (and equivalent) definition - that '$A$ contains all its limit points'. This is very familiar in the context of closed intervals: if $x_n \in [a, b]$ and $x_n \to x$ then $x \in [a, b]$.

**Lemma 7.6.** *A subset $A$ of $\mathbb{R}$ is closed if and only if $(a_n) \in A$ with $a_n \to a$ implies that $a \in A$.*

*Proof.* Suppose that $A$ is closed, i.e. $\mathbb{R} \setminus A$ is open. We want to show that if $(a_n) \in A$ and $a_n \to a$ then $a \in A$; suppose not, i.e. that $a \in \mathbb{R} \setminus A$. Since $\mathbb{R} \setminus A$ is open, there exists $\varepsilon > 0$ such that $(a - \varepsilon, a + \varepsilon) \cap A = \varnothing$. But, as $a_n \to a$, there exists $n_0$ such that $|a_n - a| < \varepsilon$ for all $n \geq n_0$, contradicting the fact that $a_n \in A$.

Now suppose that whenever $(a_n) \in A$ and $a_n \to a$ we have $a \in A$; we want to show that $\mathbb{R} \setminus A$ is open, i.e. if we take $x \notin A$ then there exists $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \cap A = \varnothing$. If this is not true, then for each $n \in \mathbb{N}$ we can find $x_n \in A$ such that $x_n \in (x - 1/n, x + 1/n)$; then $|x_n - x| < 1/n$ and so $x_n \to x$. It follows that $x \in A$, contradicting the assumption $x \notin A$. $\square$

## 7.2 Continuity via open sets

We can, in fact, define continuity in terms of open sets. In the following definition, the function $f$ need not be invertible; rather we use $f^{-1}(U)$ to denote the *preimage* of $U$, i.e. all those points that $f$ maps into $U$:

$$f^{-1}(U) = \{x \in \mathbb{R} : \ f(x) \in U\}.$$

**Lemma 7.7.** *A function $f \colon \mathbb{R} \to \mathbb{R}$ is continuous if and only if $f^{-1}(U)$ for every open set $U \subset \mathbb{R}$.*

*Proof.* Suppose that $f$ is continuous an $U$ is open. Take $c \in f^{-1}(U)$; we need to show that there exists a $\delta > 0$ such that $(c - \delta, c + \delta) \subset f^{-1}(U)$.

Since $c \in f^{-1}(U)$ we know that $f(c) \in U$. Since $U$ is open, we know that there is some $\varepsilon > 0$ such that $(f(c) - \varepsilon, f(c) + \varepsilon) \in U$. Now we use the continuity of $f$ at $c$: there exists a $\delta > 0$ such that if $|x - c| < \delta$ then $|f(x) - f(c)| < \varepsilon$. This implies that $f(x) \in U$, and so $x \in f^{-1}(U)$. It follows that $f^{-1}(U) \supset (c - \delta, c + \delta)$, so $f^{-1}(U)$ is indeed open.

Now suppose that $f^{-1}(U)$ is open whenever $U$ is open. We want to show that $f$ is continuous. So we pick $c \in \mathbb{R}$ and consider the open set $(f(c) - \varepsilon, f(c) + \varepsilon)$. We know that

$$f^{-1}(f(c) - \varepsilon, f(c) + \varepsilon)$$

is open and contains $c$. It follows that there exists a $\delta > 0$ such that

$$(c - \delta, c + \delta) \subset f^{-1}(f(c) - \varepsilon, f(c) + \varepsilon),$$

from which it follows that if $x \in (c - \delta, c + \delta)$, i.e. if $|x - c| < \delta$, then

$$f(x) \in (f(c) - \varepsilon, f(c) + \varepsilon), \qquad \text{i.e.} \quad |f(x) - f(c)| < \varepsilon,$$

so $f$ is continuous at $c$. $\square$

This definition of continuity is very widely applicable, as you will see in *Norms, Metrics, & Topologies* next year.

# Chapter 8

# Definition and continuity of trigonometric functions

We now define sine, cosine, and tangent geometrically, and show that they are continuous. Later we will show that these geometric definitions agree with the power series definitions. The proofs in this section are non-examinable.

We let $(\cos x, \sin x)$ be the Cartesian coordinates of the point on a circle of radius one centred at the origin, where $x$ is the length of the arc as indicated in Figure 8.1. We define $\tan x = \sin x / \cos x$.

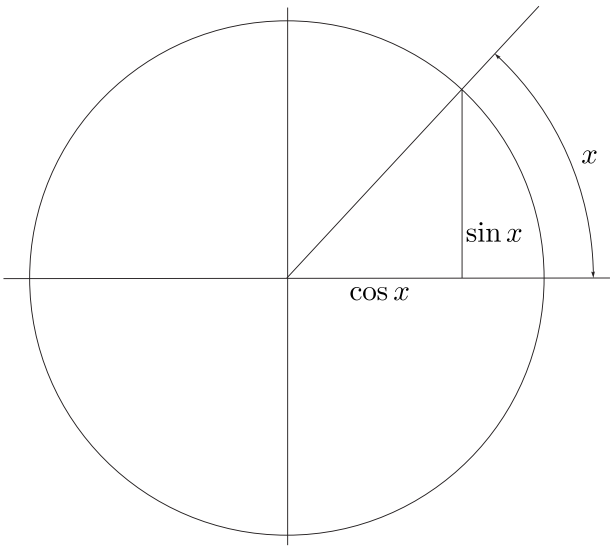


Figure 8.1: Geometric definition of sine and cosine.

This shows that sin is odd, i.e. $\sin(-x) = -\sin x$; that cos is even, i.e. $\cos(-x) = \cos(x)$, that

$$\cos(x) = \sin(\pi/2 - x), \qquad \sin(x) = \cos(\pi/2 - x), \tag{8.1}$$

that

$$\sin(x + \pi) = -\sin x, \qquad \cos(x + \pi) = -\cos x, \tag{8.2}$$

and that sin and cos have period $2\pi$, i.e.

$$\sin(x + 2\pi) = \sin(x), \qquad \cos(x + 2\pi) = \cos(x).$$

From Pythagoras's Theorem (see Figure 8.2) we have

$$\sin^2 x + \cos^2 x = 1$$

and in particular $|\sin x|, |\cos x| \le 1$.

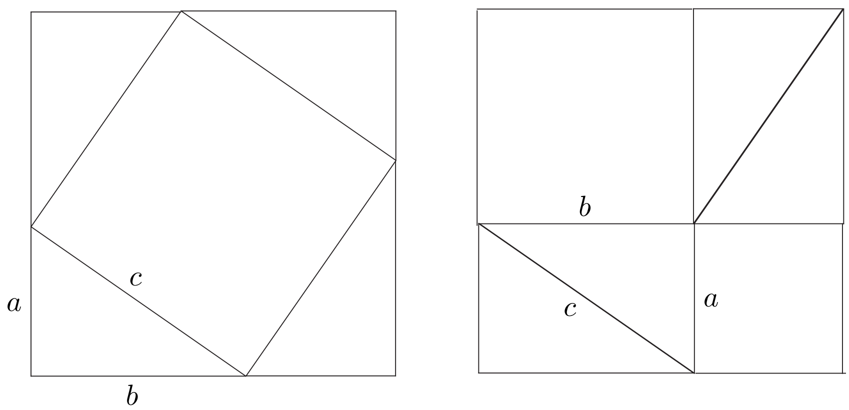


Figure 8.2: Graphical proof of Pythagoras's Theorem. Two squares of sides $a + b$ containing four copies of the same right-angled triangle. On the left, the complement of the four triangles is one square of area $c^2$, on the right it is two squares of areas $a^2$ and $b^2$.

**Lemma 8.1.** *For $x \in (0, \pi/2)$ we have*

$$0 < \sin x < x < \tan x.$$

*And for $x \in (-\pi/2, \pi/2)$*

$$|\sin x| \le |x| \le |\tan x|.$$

Since $|\sin x| \le 1$ the inequality $|\sin x| \le |x|$ is clearly valid for $x \in (-\pi, \pi)$ also.

*Proof.* Referring to Figure 8.3, the area of the sector OBA is $x/2$ $(x/2\pi \times \pi)$, while the area of triangle OBA is $(\sin x)/2$. So $\sin x < x$.

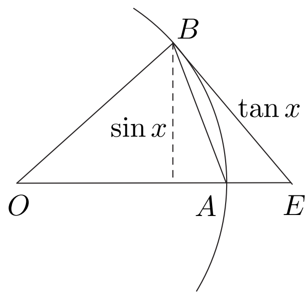


Figure 8.3: Figure for proof of Lemma 8.1.

The area of the triangle OBE is $(\tan x)/2$ and this contains the sector OBA with area $x/2$, so $x < \tan x$.

Since $\sin x$, $\tan x$, and $x$ are all odd, it follows that for $x \in (-\pi/2, \pi/2)$ we have $|\sin x| \leq |x| \leq |\tan x|$ (equality at $x = 0$). $\square$

We can also give a graphical proof of the sine and cosine summation formulae. First recall that the area of the triangle in Figure 8.4 is $\frac{1}{2}ab\sin\theta$.

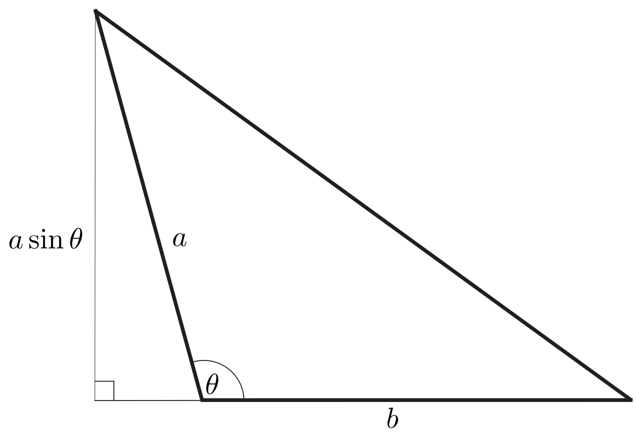


Figure 8.4: The area of the bold triangle is $\frac{1}{2}ab\sin\theta$.

**Lemma 8.2.** *For* $0 < \alpha, \beta < \pi/2$,

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta, \quad \sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta.$$

*Proof.* The area of the large triangle in Figure 8.5 is

$$\frac{1}{2}cd\sin(\alpha+\beta)$$

and the areas of the two sub-triangles are

$$\frac{1}{2}(c\sin\alpha)(d\cos\beta) \qquad \text{and} \qquad \frac{1}{2}(c\cos\alpha)(d\sin\beta).$$

Equating the whole area to the sum of the two parts yields the sine summation formula provided that $0 < \alpha, \beta$ and $\alpha + \beta < \pi$.

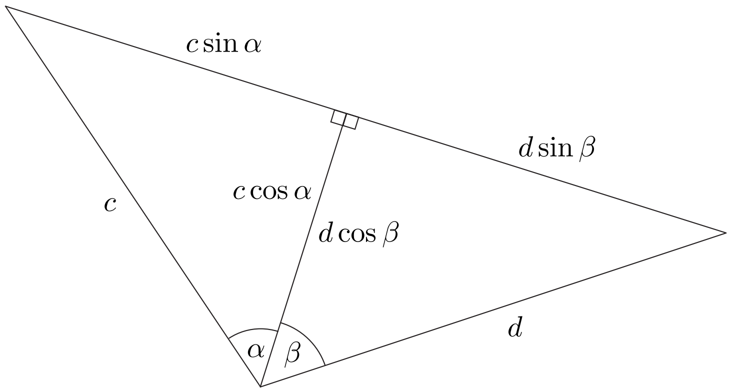


Figure 8.5: Graphical proof of the sine summation formula.

The difference formula can be proved similarly using Figure 8.6, see Examples Sheet 2. $\square$

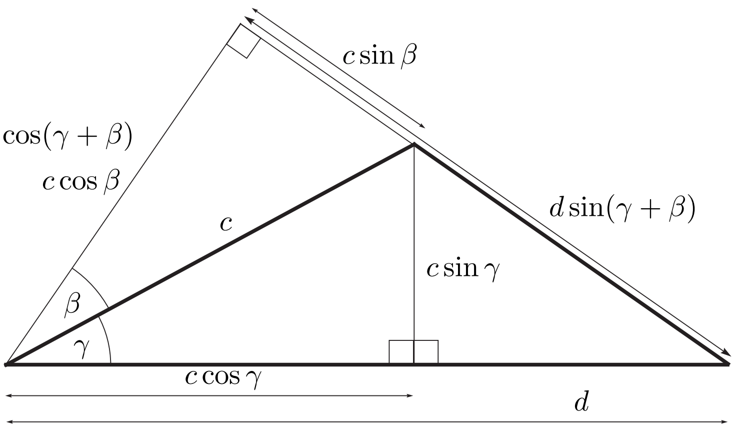


Figure 8.6: Graphical proof of the sine difference formula.

**Corollary 8.3.** *For* $0 < \alpha, \beta < \pi/2$

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \quad \cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta.$$

*Proof.* Use (8.1) along with the fact that sine is odd and cosine is even (see Examples Sheet 2 again). □

By using the elementary properties of sine and cosine it is possible to extend these equalities to all $\alpha, \beta \in \mathbb{R}$. We only state this result.

**Corollary 8.4.** *The formulae in Lemma 8.2 and Corollary 8.3 hold for all* $\alpha, \beta \in \mathbb{R}$.

We can now prove the continuity of the trigonometric functions. For a very simple proof, note that it follows from the sine summation/difference formulae that

$$\sin(\alpha+\beta) - \sin(\alpha-\beta) = 2\cos\alpha\sin\beta. \tag{8.3}$$

**Theorem 8.5.** *The function* $f\colon \mathbb{R} \to [-1,1]$ *given by* $f(x) = \sin x$ *is continuous.*

*Proof.* Fix $c \in \mathbb{R}$. Choose $\alpha = (x+c)/2$ and $\beta = (x-c)/2$ in (8.3), which gives

$$\sin x - \sin c = 2\cos((x+c)/2)\sin((x-c)/2).$$

Therefore

$$|\sin x - \sin c| = 2|\cos((x+c)/2)||\sin((x-c)/2)| \le 2 \times 1 \times \left|\frac{x-c}{2}\right|.$$

Continuity of $f$ at $c$ now follows by taking $\delta = \varepsilon$. □

Since $\cos x = \sin(\pi/2 - x)$, cosine is the composition of two continuous functions, and is therefore continuous itself. The function $\tan x := \sin x/\cos x$ is the quotient of two continuous functions and is therefore continuous away from zeros of cos (which are $\pm(2k+1)\pi/2$, $k \in \mathbb{N} \cup \{0\}$).

**Example 8.6.** The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x\sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is continuous on $\mathbb{R}$.

We know that

- $x$ is continuous on $\mathbb{R}$;
- $\sin x$ is continuous on $\mathbb{R}$;
- $1/x$ is continuous at any $c \neq 0$.

It follows using the algebra of continuous functions (product and composition) that $f$ is continuous at every $c \neq 0$. At $c = 0$ observe that

$$|f(x) - f(0)| = |f(x)| = |x \sin(1/x)| \leq |x|.$$

Given $\varepsilon > 0$, we can choose $\delta = \varepsilon$ and then

$$|x - 0| < \delta \quad \Rightarrow \quad |x| < \varepsilon \quad \Rightarrow \quad |f(x)| = |f(x) - f(0)| < \varepsilon,$$

so $f$ is also continuous at zero.

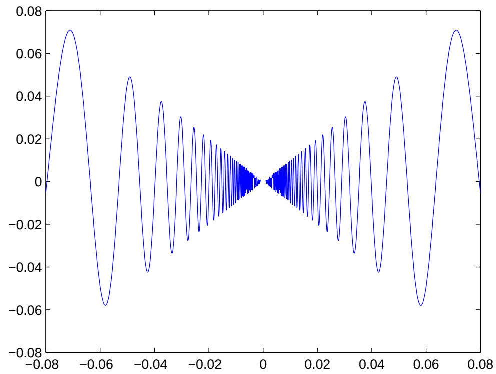


Figure 8.7: The graph of $x \sin(1/x)$ between $x = -0.08$ and $x = +0.08$.

# Chapter 9

# Continuous functions on closed bounded intervals

## 9.1 The Intermediate Value Theorem

Continuous functions on closed bounded intervals (i.e. $[a, b]$ for some $a < b$) have some nice properties. We first prove the Intermediate Value Theorem. The proof is very like the proof that there is a real number $\alpha > 0$ such that $\alpha^2 = 2$, and like that proof uses the Completeness of the Real Numbers.

**Theorem 9.1** (Intermediate Value Theorem)**.** *Suppose that $f$ is continuous on $[a, b]$ and that $f(a) < f(b)$. Then for any $g$ with $f(a) < g < f(b)$ there exists $c \in (a, b)$ such that $f(c) = g$. A similar statement holds if $f(a) > f(b)$ and $f(a) > g > f(b)$.*

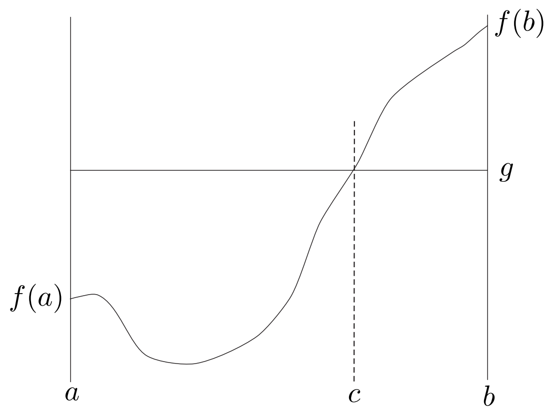


Figure 9.1: The Intermediate Value Theorem

The theorem requires continuity on the whole interval. For example,

$$f(x) = \begin{cases} -1 & x \in [-1, 0) \\ 1 & x \in [0, 1] \end{cases}$$

is continuous at every $x \in [-1, 1]$ apart from at $x = 0$. But lack of continuity at this single point invalidates the theorem.

*Proof.* Consider the set

$$S = \{x \in [a, b] : \ f(x) < g\}.$$

Note that since $a \in S$ this set is non-empty, and since $s \leq b$ for every $s \in S$ it is bounded above. Therefore, by the Least Upper Bound Axiom, we can define

$$c = \sup S$$

and clearly $c \in [a, b]$.

Either (i) $f(c) < g$; (ii) $f(c) = g$; or (iii) $f(c) > g$. We show that (i) and (iii) are impossible, and so $f(c) = g$.

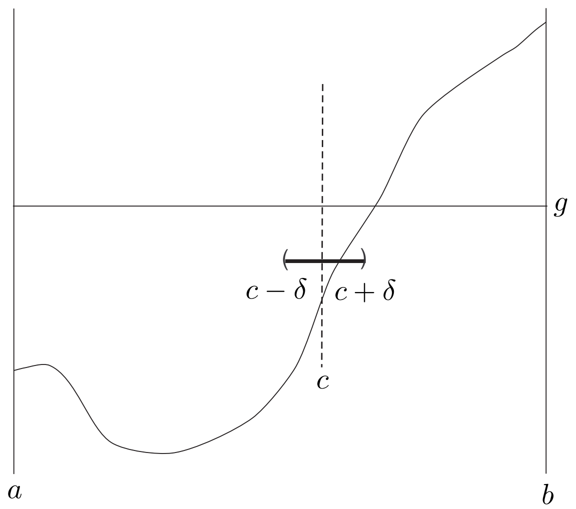


Figure 9.2: The argument that $f(c) < g$ is impossible. By continuity $f(x) < g$ on $[c, c + \delta)$; in particular $f(c + \delta/2) < g$ and so $c + \delta/2 \in S$, contradicting the fact that $c$ is an upper bound for $S$.

In case (i) we know that $c \neq b$ (since $f(b) > g$). Using the continuity of $f$ at $c$ there exists a $\delta > 0$ such that

$$x \in (c - \delta, c + \delta) \cap [a, b] \qquad \Rightarrow \qquad f(x) < g.$$

In particular, we can find $\delta > 0$ such that

$$x \in [c, c+\delta) \qquad \Rightarrow \qquad f(x) < g;$$

so $[c, c+\delta) \subset S$ and $c$ is not an upper bound for $S$. See Figure 9.2.

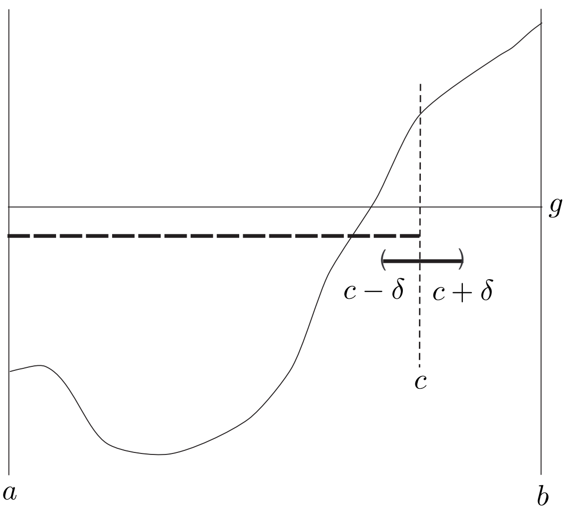


Figure 9.3: The argument that $f(c) > g$ is impossible. By continuity $f(x) > g$ on $(c-\delta, c]$, so no point in this interval can belong to $S$. The set $S$ lies within the dashed line, and we can exclude the interval $(c-\delta, c]$. This implies that $c-\delta$ is an upper bound for $S$, contradicting the fact that $c$ is the *least* upper bound for $S$.

In case (iii) we know that $c \neq a$ (because $g > f(a)$), and using the continuity of $f$ at $c$, there exists a $\delta > 0$ such that

$$x \in (c-\delta, c+\delta) \cap [a, b] \qquad \Rightarrow \qquad f(x) > g;$$

in particular we can find $\delta > 0$ such that

$$x \in (c-\delta, c] \qquad \Rightarrow \qquad f(x) > g.$$

So $(c-\delta, c]$ is disjoint from $S$, so $c - \delta$ is an upper bound for $S$, contradicting the fact that $c$ is the least upper bound. See Figure 9.3.

The only possibility left is that $f(c) = g$. $\square$

The IVT has many interesting consequences/applications.

### 9.1.1 Existence of roots of equations

We can sometimes use the IVT to prove that an equation has a root: suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous, and that there are points $x, y \in [a,b]$ such that

$$f(x) > 0 \qquad \text{and} \qquad f(y) < 0.$$

Then the IVT applied to the interval $[x,y]$ (or $[y,x]$ if $y < x$) implies that there is a point $c$ between $x$ and $y$ such that $f(c) = 0$. To find the root, one can use 'interval halving', see Examples Sheet 3.

**Example 9.2.** The equation $x^2 = 2$ has a root in the interval $(0,2)$.

Consider the function $f(x) = x^2 - 2$. Then $f(0) = -2$ and $f(2) = 2$, so there exists a $c \in (0,2)$ such that $f(c) = c^2 - 2 = 0$.

**Example 9.3.** The equation $x = \mathrm{e}^{-x}$ has a root in the interval $(0,1)$.

For now we will assume standard properties of the exponential function, including the fact that it is continuous. Consider the function $f(x) = x - \mathrm{e}^{-x}$, which is continuous on $[0,1]$ (and in fact on $\mathbb{R}$). Then $f(0) = -1$ and $f(1) = 1 - 1/\mathrm{e} > 0$. By the IVT, there is a point $c \in (0,1)$ such that $f(c) = 0$, i.e. for which $c = \mathrm{e}^{-c}$, see Figure 9.4.

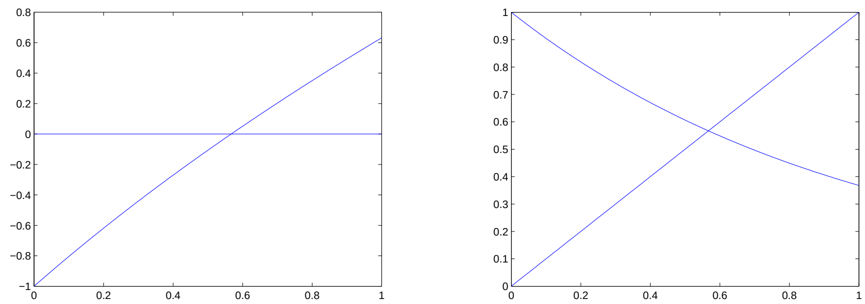


Figure 9.4: There is a point $c \in (0,1)$ such that $c = \mathrm{e}^{-c}$. On the left: if $f(x) = x - \mathrm{e}^{-x}$ then $f(0) < 0$ and $f(1) > 0$, so the IVT implies that there is a $c \in (0,1)$ such that $f(c) = 0$. On the right: graphs of $y = x$ and $y = \mathrm{e}^{-x}$ intersecting at $c \simeq 0.5671$ (found using the MATLAB `fzero` function).

**Proposition 9.4.** *Any odd degree polynomial has at least one real root.*

*Proof.* Consider

$$P(x) = \sum_{j=0}^{2n+1} a_j x^j$$

where $a_{2n+1} \neq 0$. We can assume without loss of generality that $a_{2n+1} > 0$, since $P(x) = 0$ if and only if $-P(x) = 0$.

We have already seen that $P$ is continuous on $\mathbb{R}$. We show that there exists an $x^* > 0$ such that $P(x^*) > 0$ and an $x_* < 0$ such that $P(x_*) < 0$. The existence of a root in the interval $(x_*, x^*)$ then follows from the IVT applied to $P\colon [x_*, x^*] \to \mathbb{R}$.

Now let $A = \sum_{j=0}^{2n} |a_j|$. Note that

$$P(x) \geq a_{2n+1}x^{2n+1} - \sum_{j=0}^{2n} |a_j||x|^j.$$

For $x > 1$ this implies that

$$P(x) \geq a_{2n+1}x^{2n+1} - \left(\sum_{j=0}^{2n} |a_j|\right) x^{2n} = a_{2n+1}x^{2n+1} - Ax^{2n}.$$

Thus $P(x) > 0$ if $x > A/a_{2n+1}$ so we can take $x^* = 2A/a_{2n+1}$. Similarly $P(x) < 0$ if $x < -A/a_{2n+1}$ so we can take $x_* = -2A/a_{2n+1}$. □

### 9.1.2 A fixed point theorem

Another application is a simple 'fixed point theorem'.

**Lemma 9.5.** *Any continuous function $f\colon [a, b] \to [a, b]$ has a fixed point, i.e. there exists an $x^* \in [a, b]$ such that $f(x^*) = x^*$.*

*Proof.* Consider the function $g(x) = f(x) - x$. Then $0 \leq g(a)$ and $g(b) \leq 0$.

If $g(a) = 0$ then $f(a) = a$, and if $g(b) = 0$ then $f(b) = b$. Otherwise $g(a) > 0$ and $g(b) < 0$, in which case by the IVT there exists a $c \in (a, b)$ such that $g(c) = 0$, i.e. $f(c) = c$. □

The same result does not hold if $[a, b]$ is replaced by $(a, b)$: consider, for example, $f\colon (0, 1) \to (0, 1)$ given by $f(x) = x/2$. Then $f(x) \neq x$ for all $x \in (0, 2)$. If we considered $f$ as a function from $[0, 1]$ into itself then we would have $f(0) = 0$.

## 9.2 Continuous functions map intervals to intervals

A more abstract result requires a preliminary definition.

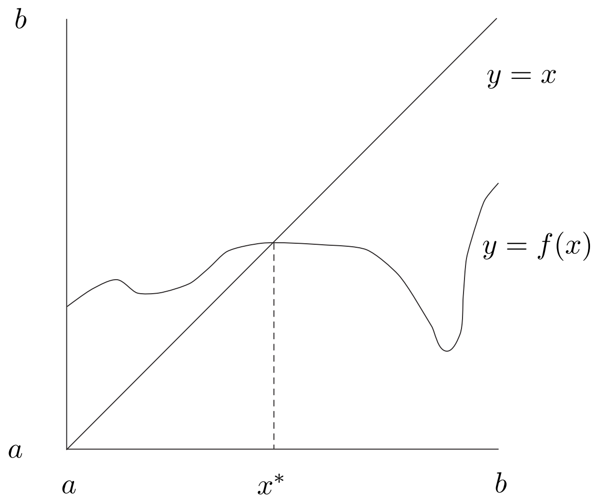


Figure 9.5: A continuous function $f\colon [a,b] \to [a,b]$ must have a fixed point, i.e. there exists an $x^* \in [a,b]$ such that $f(x^*) = x^*$.

**Definition 9.6.** A non-empty subset $I$ of $\mathbb{R}$ is an *interval* if whenever $x, y \in I$ and $x < y$ then $[x, y] \in I$.

Note that the definition allows the 'degenerate' case of a single point $\{a\}$. We specify non-empty since otherwise the empty set would satisfy the definition.

**Lemma 9.7.** *Any interval is a set of the following form: a single point $\{a\}$, a (non-trivial) bounded interval*

$$(a,b), \qquad (a,b], \qquad [a,b), \qquad [a,b]; \tag{9.1}$$

*a semi-infinite interval*

$$(-\infty,b), \qquad (-\infty,b], \qquad (a,\infty), \qquad [a,\infty); \tag{9.2}$$

*or the whole line $\mathbb{R}$.*

*Proof.* If $I$ is an interval then by definition it is non-empty. Now let $a = \inf I$ and $b = \sup I$; then $a = -\infty$ or $a \in \mathbb{R}$, and $b \in \mathbb{R}$ or $b = +\infty$. We have $a \le x \le b$ for every $x \in I$.

If $a = b$ then $I = \{a\}$.

Otherwise we start by showing that $(a, b) \subset I$. Indeed, if $a < x$ then there exists $a' \in I$ such that $a' < x$: if $a = -\infty$ then $I$ is not bounded below (so there exists $a' \in I$

with $a' < x$), while if $a \in \mathbb{R}$ then we can take $\varepsilon = (x-a)/2$ and find $a' \in I$ such that $a' < a + \varepsilon < (x+a)/2 < x$. Similarly, if $x < b$ then there exists $b' \in I$ with $b' > x$.

Therefore for any $x \in (a,b)$, for which $a < x < b$, there exists $a' \in I$ with $a' < x$ and $b' \in I$ with $b' > x$. Since $I$ is an interval, $[a', b'] \subset I$ and so in particular $x \in I$.

It follows that $(a,b) \subset I$ as claimed; if $a = -\infty$ and $b = \infty$ then $I = \mathbb{R}$.

If both $a$ and $b$ are finite then we are in case (9.1); whether the endpoints are included depends on whether $\inf I \in I$ and $\sup I \in I$; for example, if $a = \inf I \in I$ but $b = \sup I \notin I$ then $I = [a, b)$.

If only one of $a$ and $b$ are finite then we are in case (9.2); for example, if $a = -\infty$ and $b \in \mathbb{R}$ then the interval is either $(-\infty, b)$ or $(-\infty, b]$, depending on whether $b = \sup I \in I$ or not. $\square$

**Corollary 9.8.** *Let $I$ be an interval and suppose that $f\colon I \to \mathbb{R}$ is continuous. Then $f(I)$ is an interval.*

Note that our definition of an 'interval' allows $f(I)$ to be a single point (and this is clearly possible as any constant function shows).

*Proof.* Observe that if $a, b \in f(I)$ with $a < b$ then there exist $x, y \in I$ $(x \neq y)$ such that

$$f(x) = a \qquad \text{and} \qquad f(y) = b.$$

Suppose that $x < y$ and $a < b$. Then for any $g \in (a,b)$ there exists a point $c \in (x,y)$ such that $f(c) = g$. In particular, $g \in f(I)$. Since this holds for any $g \in (a,b)$, it follows that $[a,b] \subset f(I)$. The other three cases can be treated similarly. $\square$

You are asked to show on Examples Sheet 4 that if $I$ is an open interval then $f(I)$ can be any kind of interval. We will see shortly that if $I$ is a closed bounded interval then $f(I)$ must also be a closed bounded interval.

## 9.3 The Extreme Value Theorem

A function $f\colon E \to \mathbb{R}$ is bounded above on $E$ if there exists an $M \in \mathbb{R}$ such that

$$f(x) \leq M \qquad \text{for all} \qquad x \in E;$$

and bounded below on $E$ if there exists an $m \in \mathbb{R}$ such that

$$f(x) \geq m \qquad \text{for all} \qquad x \in E.$$

It is bounded on $E$ if it is bounded above and below on $E$.

If $f$ is bounded above on $E$ (and $E$ is non-empty) then the set $f(E) = \{f(x) : x \in E\}$ is non-empty and bounded above, so has a least upper bound ('supremum'), which we denote by

$$\sup\{f(x) : x \in E\} = \sup_{x \in E} f(x)$$

(if we are not being precise we might just write '$\sup f$' or $\sup_E f$).

The proof of the Extreme Value Theorem uses the following ingredients:

(i) The Bolzano–Weierstrass Theorem (any bounded sequence of real numbers has a convergent subsequence);

(ii) if $(x_n) \in [a, b]$ and $x_n \to y$ then $y \in [a, b]$ (since $a \leq x_n \leq b \; \forall \; n$);

(iii) if $f : E \to \mathbb{R}$ is continuous at $c \in E$, and $(x_n) \subset E$ with $x_n \to c$, then $f(x_n) \to f(c)$ (Lemma 6.1).

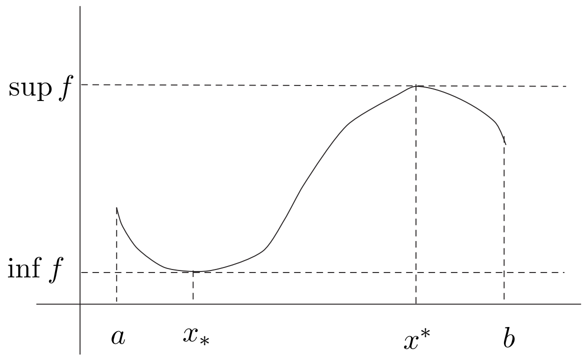


Figure 9.6: Illustration of the Extreme Value Theorem. Note that the range of $f$ is the closed interval $[\inf f, \sup f]$, as proved in Corollary 9.10.

The fact that the function is defined on an interval that is *closed* and *bounded* is crucial in the proof.

**Theorem 9.9** (The Extreme Value Theorem)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous then $f$ is bounded and attains its bounds. In other words, $f$ is bounded above and below on $[a, b]$, and there exist $x^*, x_* \in [a, b]$ such that*

$$f(x_*) = \inf_{x \in [a,b]} f(x) \qquad \textit{and} \qquad f(x^*) = \sup_{x \in [a,b]} f(x).$$

For a proof that does not use the Bolzano–Weierstrass Theorem see Examples 4.

*Proof.* We give the proof for the supremum. Suppose that $f$ is not bounded above on $[a, b]$. Then for each $n \in \mathbb{N}$ there exists an $x_n \in [a, b]$ such that

$$f(x_n) \geq n.$$

Since $(x_n)$ is a bounded sequence, the Bolzano–Weierstrass Theorem guarantees that it has a convergent subsequence, $x_{n_j} \to y$. Since $a \leq x_{n_j} \leq b$ for every $j$, it follows that $a \leq y \leq b$, i.e. $y \in [a, b]$.

Since $f$ is continuous at $y$, it follows (from Lemma 6.1) that $f(x_{n_j}) \to f(y)$. But by assumption $f(x_{n_j}) \to \infty$, while $f(y) < \infty$. This yields a contradiction, and so $f$ must be bounded on $[a, b]$.

Since $f$ is bounded on $[a, b]$, it follows that the set

$$S = \{f(x) : \ x \in [a, b]\}$$

is bounded. It is clearly non-empty, so $M = \sup S$ exists. We know that for any $\varepsilon > 0$, there exists an element $s \in S$ such that $s > M - \varepsilon$, since $M$ is the least upper bound. In particular, if for each $n \in \mathbb{N}$ we choose $\varepsilon = 1/n$ we can find an element $x_n \in [a, b]$ such that

$$M - 1/n < f(x_n) \leq M.$$

Once again we use the Bolzano–Weierstrass Theorem to find a subsequence of the $x_n$, $x_{n_j}$, that converges to some $x^* \in [a, b]$ as $j \to \infty$. Since $f$ is continuous on $[a, b]$ it is continuous at $x^*$, and so it follows (from Lemma 6.1 again) that

$$f(x^*) = \lim_{j \to \infty} f(x_{n_j}).$$

Since

$$M - 1/n_j < f(x_{n_j}) \leq M$$

it follows from the sandwich rule for sequences that $f(x^*) = M$.

For the lower bound consider the function $x \mapsto -f(x)$. $\square$

Note that we need all the hypotheses:

- the function $x \mapsto 1/x$ on $(0, 1)$ is not bounded (interval is not closed);
- the function $x \mapsto x$ on $[0, \infty)$ is not bounded (interval is not bounded);
- the function $f \colon [0, 1] \to \mathbb{R}$ given by

  $$f(x) = \begin{cases} 1/x & x \neq 0 \\ 0 & x = 0 \end{cases}$$

  is not bounded ($f$ is not continuous);

- the function $f\colon [0,1) \to \mathbb{R}$ defined by $f(x) = x$ is bounded, $0 \le f(x) \le 1$, but the supremum is not attained (interval is not closed);
- the function $f\colon [0,3] \cap \mathbb{Q} \to \mathbb{R}$ defined by $f(x) = \sin x$ is bounded, $0 \le f(x) \le 1$, but the supremum is not attained, since $\pi/2 \notin [0,3] \cap \mathbb{Q}$ (the function is not defined on an interval).

**Corollary 9.10.** *If $f\colon [a,b] \to \mathbb{R}$ is continuous then $f([a,b]) = [\inf f, \sup f]$.*

*Proof.* We have already shown that a continuous function maps intervals to intervals. The previous proof shows that

$$\sup f([a,b]) = f(x^*) \in f([a,b]) \qquad \inf f([a,b]) = f(x_*) \in f([a,b]),$$

and so $f([a,b]) = [f(x_*), f(x^*)]$ as claimed. $\square$

We can now address the question of the existence of zeros of polynomials of even degree.

**Theorem 9.11.** *Let $P : \mathbb{R} \to \mathbb{R}$ be given by*

$$P(x) = \sum_{k=0}^{2n} a_k x^k, \qquad a_{2n} > 0,$$

*i.e. an even degree polynomial with positive leading coefficient. Then $P$ has a minimum at some point $x = x_*$, and (i) if $P(x_*) > 0$ there are no real roots; (ii) if $P(x_*) = 0$ there is at least one real root; and (iii) if $P(x_*) < 0$ then there are at least two real roots.*

*Proof.* We want to apply the EVT, but we cannot do it immediately since $P$ is not defined on a closed bounded interval.

First, note that certainly the minimum of $P(x)$ (if it exists) is no larger than $P(0)$. Following the argument in the proof of Proposition 9.4, for all $x$ with $|x| > 1$ we have

$$\begin{aligned} P(x) &\ge a_{2n}|x|^{2n} - \sum_{k=0}^{2n-1} |a_k||x|^k \\ &\ge a_{2n}|x|^{2n} - \left(\sum_{k=0}^{2n-1} |a_k|\right) |x|^{2n-1}; \end{aligned}$$

so[1] there exists an $R > 0$ such that $P(x) \ge \max(1, P(0) + 1)$ for all $|x| \ge R$.

1 We have $P(x) \ge \alpha|x|^{2n} - \beta|x|^{2n-1}$ for some $\alpha > 0$, $\beta \ge 0$. So for any choice of $M \ge 0$, in order to ensure that $P(x) \ge M$ it is enough to have

$$\alpha|x|^{2n} - \beta|x|^{2n-1} \ge M \qquad \Leftrightarrow \qquad \alpha \ge \frac{\beta}{|x|} + \frac{M}{|x|^{2n}},$$

which we can always ensure by taking $|x| > R := \max(2\beta/\alpha, (2M/\alpha)^{1/2n})$. (In the proof here we also need to ensure that $R > 1$ so that the inequalities we have used for $P(x)$ are valid.)

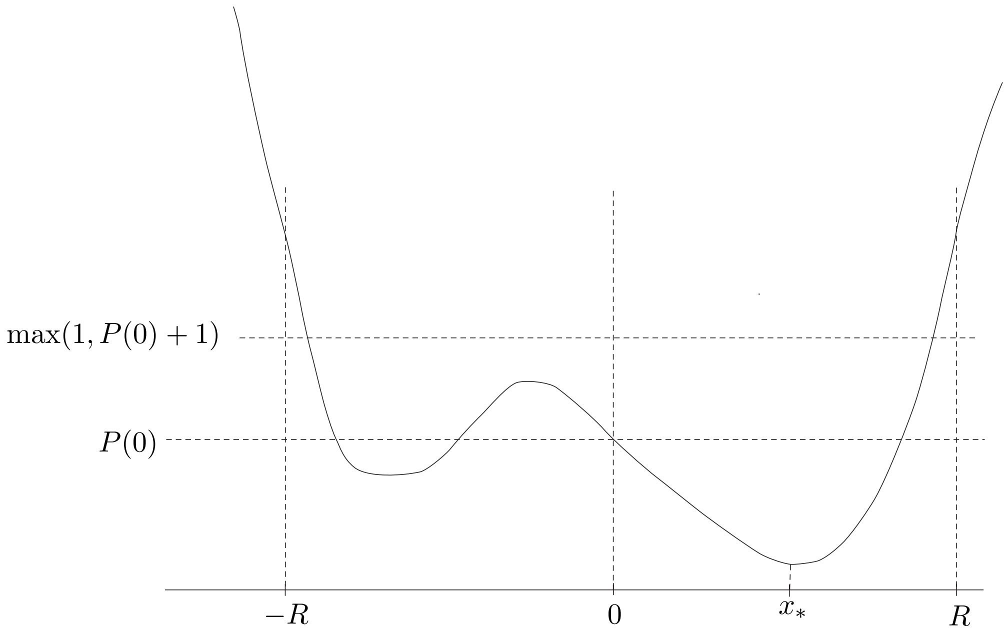


Figure 9.7: Application of the EVT to find the minimum of an even degree polynomial. $P(x) \geq \max(1, P(0) + 1)$ for $|x| \geq R$ (this need not be the optimum choice of $\varrho$). We can apply the EVT on $[-R, R]$ to prove the existence of $x_*$.

We can now apply the EVT on the interval $[-R, R]$ to prove the existence of a point $x_* \in [-R, R]$ such that

$$P(x) \geq P(x_*) \qquad \text{for all} \quad x \in [-R, R].$$

Then for $|x| \geq R$ we have

$$P(x) \geq P(0) \geq P(x_*),$$

since $0 \in [-R, R]$, see Figure 9.7. Therefore $P(x) \geq P(x_*)$ for all $x \in \mathbb{R}$ and we have found the minimum of $P$.

Now parts (i) and (ii) are immediate, and since

$$P(-R) \geq 1 > 0, \qquad P(x_*) < 0, \qquad P(R) \geq 1 > 0$$

part (iii) follows using the IVT on the intervals $[-R, x_*]$ and $[x_*, R]$. $\square$

# Chapter 10

# Uniform continuity

In general, in the definition of continuity

> A function $f\colon E \to \mathbb{R}$ is *continuous* at $c \in E$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that
>
> $$x \in E \quad \text{and} \quad |x - c| < \delta \qquad \Rightarrow \qquad |f(x) - f(c)| < \varepsilon. \tag{10.1}$$

the choice of $\delta$ depends on both $\varepsilon$ and on the point $c$.

For example, when we showed in Example 5.5 that $x \mapsto 1/x$ is continuous at a point $c \neq 0$, we had to take $\delta = \min(\frac{1}{2}|c|, \frac{1}{2}|c|^2\varepsilon)$. This reflects the fact that $1/x$ changes very quickly when $c$ is small. For more on this see Q22 on Examples 4.

If $\delta$ only depends on $\varepsilon$ and not on $c$ for all $c \in E$, we say that $f$ is *uniformly continuous* on $E$.

**Definition 10.1.** A function $f\colon E \to \mathbb{R}$ is *uniformly continuous* on $E$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$x \in E \quad \text{and} \quad |x - y| < \delta \qquad \Rightarrow \qquad |f(x) - f(y)| < \varepsilon \tag{10.2}$$

for every $x, y \in E$.

Note that if a function is uniformly continuous on $E$, then it must be continuous on $E$.

For example, we showed in Example 5.4 that for the function $x \mapsto \alpha x$ we can take $\delta = \varepsilon/\alpha$, and this is independent of $x$, so this function is uniformly continuous on $\mathbb{R}$.

We can show that a continuous function on a closed and bounded interval is also uniformly continuous.

**Theorem 10.2.** *If $f\colon [a,b] \to \mathbb{R}$ is continuous, then $f$ is uniformly continuous on $[a,b]$.*

*Proof.* Suppose that $f$ is not uniformly continuous. Then there exists an $\varepsilon > 0$ such that for every $\delta > 0$ we can find $x, y \in E$ such that $|x - y| < \delta$ and $|f(x) - f(y)| \geq \varepsilon$.

For each $n$ we take $\delta = 1/n$, and find $x_n, y_n \in E$ such that

$$|x_n - y_n| < \frac{1}{n} \qquad \text{and} \qquad |f(x_n) - f(y_n)| \geq \varepsilon. \tag{10.3}$$

Since $(y_n)$ is a bounded sequence, we can find a subsequence $(y_{n_j})$ that converges to some $y \in [a,b]$ as $j \to \infty$.

Since

$$|x_{n_j} - y| \leq |x_{n_j} - y_{n_j}| + |y_{n_j} - y| < \frac{1}{n_j} + |y_{n_j} - y|$$

it follows that $x_{n_j} \to y$ as well.

By assumption we know that $f$ is continuous at $y$: so for our particular $\varepsilon$ there exists $\delta > 0$ such that

$$|x - y| < \delta \qquad \Rightarrow \qquad |f(x) - f(y)| < \varepsilon/2.$$

However, for $j$ sufficiently large we have $|x_{n_j} - y| < \delta$ and $|y_{n_j} - y| < \delta$, and so

$$|f(x_n) - f(y_n)| \leq |f(x_n) - f(y)| + |f(y) - f(y_n)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

contradicting (10.3). $\square$

# Chapter 11

# Monotonic functions and continuous inverses

Suppose that $I$ is an interval. We have already shown that if $f\colon I \to \mathbb{R}$ is continuous then $f(I)$ is an interval, $J$. In this chapter we will show that if $f\colon I \to J$ is continuous and strictly monotonic (strictly increasing/decreasing) then $f^{-1}\colon J \to I$ is continuous.

Recall that a function $f\colon A \to B$ is *injective/one-to-one* if

$$f(x) = f(y) \qquad \Rightarrow \qquad x = y.$$

Equivalently, if $x \neq y$ then $f(x) \neq f(y)$. A function $f\colon A \to B$ is *surjective/onto* if $f(A) = B$, i.e.

$$\text{for every } \alpha \in B \quad \text{there exists} \quad x \in A \quad \text{such that} \quad f(x) = \alpha.$$

A function $f\colon A \to B$ is *bijective* and termed a *bijection* if it is both injective and surjective (one-to-one and onto).

When $f\colon A \to B$ is a bijection we can define $f^{-1}\colon B \to A$ by setting

$$f^{-1}(\alpha) = x \qquad \text{when} \qquad f(x) = \alpha.$$

Note that $f\colon A \to f(A)$ (the range of $A$) is always surjective, so if $f$ is injective we can define $f^{-1}\colon f(A) \to A$.

The key notion for obtaining an inverse is in fact monotonicity.

**Definition 11.1.** Let $E \subset \mathbb{R}$. A function $f\colon E \to \mathbb{R}$ is said to be *increasing* (on $E$) if

$$x \geq y \qquad \Rightarrow \qquad f(x) \geq f(y)$$

and *strictly increasing* (on $E$) if

$$x > y \qquad \Rightarrow \qquad f(x) > f(y).$$

A function $f\colon E \to \mathbb{R}$ is said to be *decreasing* (on $E$) if

$$x \geq y \qquad \Rightarrow \qquad f(x) \leq f(y).$$

and *strictly decreasing* (on $E$) if

$$x > y \qquad \Rightarrow \qquad f(x) < f(y).$$

(Throughout the definition we take $x, y \in E$.)

A function that is strictly increasing or strictly decreasing is called *strictly monotonic*, and *monotonic* if it is decreasing or increasing.

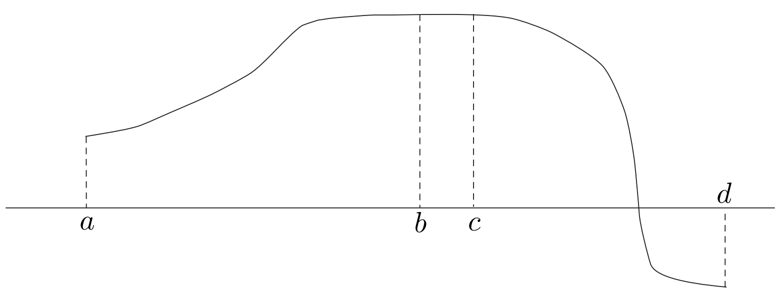


Figure 11.1: The function whose graph is shown here is constant on $[b, c]$. It is strictly increasing on $[a, b]$; increasing on $[a, c]$; decreasing on $[b, d]$; and strictly decreasing on $[c, d]$. It is not monotonic on $[a, d]$

Some care is needed over this terminology - note that any constant function is both 'increasing' and 'decreasing' in terms of this definition. In particular, what is meant by a function being 'increasing' varies from author to author, some taking this to mean strictly increasing, others 'non-decreasing' as here.

**Theorem 11.2** (Inverse Function Theorem for continuous functions)**.** *Let $I$ be an interval and suppose that $f\colon I \to \mathbb{R}$ is continuous and strictly monotonic. Then $J = f(I)$ is an interval and $f^{-1}\colon J \to I$ is continuous and strictly monotonic (in the same sense as $f$).*

*Proof.* We have already shown that $J = f(I)$ is an interval in Corollary 9.8.

For the rest of the result we treat the case where $f$ is strictly increasing. If $f$ is strictly decreasing then we can consider $-f$ instead.

Since $f$ is strictly increasing, it is injective: if $x \neq y$ then $x > y$ or $y > x$, in which case $f(x) > f(y)$ or $f(x) < f(y)$, i.e. $f(x) \neq f(y)$. Since $J = f(I)$ the function $f\colon I \to J$ is surjective. So $f\colon I \to J$ is bijective, and we can define its inverse $f^{-1}\colon J \to I$.

Suppose that $f^{-1}$ is not strictly increasing. Then there exist $\alpha, \beta \in J$ such that $\alpha > \beta$ but $f^{-1}(\alpha) \leq f^{-1}(\beta)$. Set

$$x = f^{-1}(\alpha) \qquad \text{and} \qquad y = f^{-1}(\beta).$$

Then

$$x \leq y \qquad \text{and} \qquad f(x) = \alpha > \beta = f(y),$$

which contradicts the fact that $f$ is strictly increasing.

We now have a strictly increasing function $f^{-1}\colon J \to I$. Using the fact that $f^{-1}$ is surjective (onto) we now deduce that $f^{-1}$ must be continuous.

Take $c \in J$ and choose $\varepsilon > 0$. We need to show that there exists a $\delta > 0$ such that

$$y \in J,\ |y - c| < \delta \qquad \Rightarrow \qquad |f^{-1}(y) - f^{-1}(c)| < \varepsilon.$$

Suppose that $z = f^{-1}(c)$ is not an endpoint of $I$, so that $(z - \varepsilon, z + \varepsilon) \subset I$ for $\varepsilon$ sufficiently small. Since $f$ is strictly increasing,

$$f(z + \varepsilon) > f(z) = c, \quad \Rightarrow \quad f(z + \varepsilon) = c + \delta_1, \quad \delta_1 > 0.$$

Similarly

$$f(z - \varepsilon) < f(z) = c \quad \Rightarrow \quad f(z - \varepsilon) = c - \delta_2, \quad \delta_2 > 0.$$

Now set $\delta = \min(\delta_1, \delta_2)$. Then, since $f^{-1}$ is strictly increasing,

$$y \in (c - \delta, c + \delta) \ \Rightarrow \ \begin{cases} f^{-1}(y) < f^{-1}(c + \delta_1) = f^{-1}(c) + \varepsilon \\ f^{-1}(y) > f^{-1}(c - \delta_2) = f^{-1}(c) - \varepsilon, \end{cases}$$

i.e.

$$|f^{-1}(y) - f^{-1}(c)| < \varepsilon.$$

See Figure 11.2.

If $z = f^{-1}(c)$ is an endpoint of $I$ the argument is similar, but easier since we only have to look 'on one side' of $c$: we treat the case when $z$ is the left-hand endpoint of $I$. Note that it follows immediately that

$$f^{-1}(c) \leq f^{-1}(y) \qquad \text{for all} \quad y \in J.$$

In this case, if $\varepsilon > 0$ is sufficiently small then $[z, z + \varepsilon) \subset I$. Since $f$ is strictly increasing,

$$f(z + \varepsilon) > f(z) = c, \quad \Rightarrow \quad f(z + \varepsilon) = c + \delta, \quad \delta > 0.$$

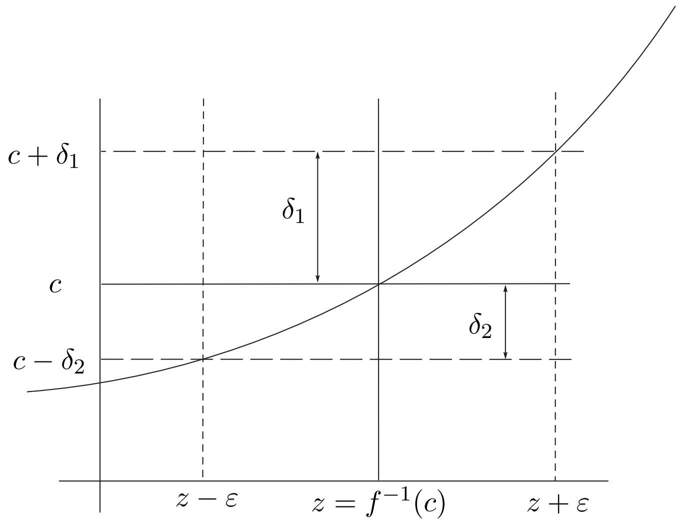


Figure 11.2: The argument to show that $f^{-1}$ is continuous at $c$. With $z = f^{-1}(c)$, given $\varepsilon > 0$, find $\delta_1, \delta_2 > 0$ such that $f(z - \epsilon) = c - \delta_2$ and $f(z + \epsilon) = c + \delta_1$.

Since $f^{-1}$ is strictly increasing,

$$y \in [c, c + \delta) \;\Rightarrow\; f^{-1}(y) < f^{-1}(c + \delta) = f^{-1}(c) + \varepsilon,$$

i.e.

$$f^{-1}(c) \leq f^{-1}(y) < f^{-1}(c) + \varepsilon.$$

See Figure 11.2. □

There are some more general results hidden in this proof:

(i) If $E \subset \mathbb{R}$ and $f\colon E \to \mathbb{R}$ is strictly monotonic, then it has an inverse $f^{-1}\colon f(E) \to E$ that is strictly monotonic in the same sense.

(ii) Let $I$ and $J$ be intervals. If $f\colon I \to J$ is monotonic and surjective then it is continuous.

The final question on Examples 4 uses Theorem 11.2 to deduce that $x \mapsto x^{1/n}$ is a continuous function on $[0, \infty)$ for every $n \in \mathbb{N}$.