Prelims: Analysis II

Continuity and Differentiability

Hilary Term 2023

Paul Balister

Last updated: January 5, 2023

A note on these notes

These notes are designed to accompany the University of Oxford Prelims *Analysis II* lecture course. They are adapted from, and owe much to, the notes of many previous lecturers, in particular H.A. Priestley, Z. Qian and R. Heath-Brown. They are quite heavily revised from the previous year's notes, so will undoubtedly contain typos and other mistakes. Please send any corrections or comments to `Paul.Balister@maths.ox.ac.uk`.

Lectures

To get the most out of the course you *must* attend the lectures. On the other hand, you should also read the relevant section of the notes *before* attending the lecture. The two complement each other, and having read the notes will make it easier to follow the lectures (even if you did not follow everything in the notes), and learn more from the lectures (even if you think you did follow everything in the notes). There will be more explanation in the lectures than there is in the notes. On the other hand I will not put everything on the board which is in the printed notes. In some places I have put in extra examples which I will not have time to demonstrate in the lectures. There is some extra material in the notes which I have put in for interest, but which I do not regard as central to the course and will probably not be covered in the lectures. This material will be marked as non-examinable.

Problem Sheets

The weekly problem sheets which accompany the lectures are an integral part of the course. You will only really understand the definitions and theorems in the course by doing the problems! I assume that week 1 tutorials are being devoted to the final sheets from the Michaelmas Term courses. I therefore suggest that the problem sheets for this course are tackled in tutorials in weeks 2–8, with the 8th sheet used a vacation work for a tutorial in the first week of Trinity Term. The problem sheets contain bonus questions 'for the enthusiasts' — these are usually harder questions and students are not expected to complete all, or even any, of them.

Contents

0 Summary of results from Analysis I

I will not cover this section in the lectures as it is material you should be familiar with from the *Introduction to University Mathematics* and *Analysis I* courses. I include it here as a summary and reminder of things you should know. Refer to previous course notes for more details.

Standard sets of numbers

$\mathbb{N}$: the set of natural numbers[1], $\{1, 2, 3, \dots\}$;

$\mathbb{Z}$: the set of integers, $\{\dots, -2, -1, 0, 1, 2, \dots\}$;

$\mathbb{Q}$: the set of rational numbers, $\{\frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0\}$;

$\mathbb{R}$: the set of all real numbers (the real line);

$\mathbb{C}$: the set of all complex numbers (the complex plane, $\{a + ib : a, b \in \mathbb{R}\}$).

We have

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Infinity (∞) and negative infinity ($-\infty$) are a convenient device for expressing certain notions concerning real numbers. *They are not themselves real numbers.*

Quantifiers

$\forall$: "for all" or "for every" or "whenever".

$\exists$: "there exist(s)" or "there is (are)".

Quantifiers matter. Treat them with care and respect. The order in which quantifiers are written down is important. For example[2]

$$\forall y \in \mathbb{R} \colon \exists x \in \mathbb{R} \colon x > y$$

is true as we can chose x depending on y, say $x = y + 1$, while

$$\exists x \in \mathbb{R} \colon \forall y \in \mathbb{R} \colon x > y$$

is false as we need the same x to work for all y. Statements such as

'There is an $x \in \mathbb{R}$ such that $x > y$ for all $y \in \mathbb{R}$.'

are therefore ambiguous. Good discipline is to put quantifiers at the front of a statement (even when written out in words), not at the back as an afterthought, and to read carefully from left to right.

---

[1] Some people prefer to start from zero. It makes little difference in this course.

[2] I prefer to use : to separate $\forall$ and $\exists$ from the statements they are quantifying, as opposed to 's.t.', commas or spacing. This is not standard, but I think it helps readability.

Arithmetic and ordering

The real numbers, with their usual arithmetic operations ($+$, $-$, $\times$, $\div$) and usual ordering ($<$, $\leq$, $>$, $\geq$), form an **ordered field**. See *Analysis I* notes for the formal details.

We define the **modulus** or **absolute value** by

$$|x| = \begin{cases} x, & x > 0; \\ 0, & x = 0; \\ -x, & x < 0. \end{cases}$$

Key facts about the modulus and inequalities:

(a) **Triangle Inequality**: for all $a, b \in \mathbb{R}$, $|a \pm b| \leq |a| + |b|$.

(b) **Reverse Triangle Inequality**: for all $a, b \in \mathbb{R}$, $|a \pm b| \geq ||a| - |b|| \geq |a| - |b|$.

(c) **Interval around a point**: for all $a, r, x \in \mathbb{R}$,
$|x - a| < r \iff a - r < x < a + r \iff x \in (a - r, a + r)$.

The complex numbers $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$ also forms a field with the usual operations $+$, $-$, $\times$ and $\div$, but do *not* have an ordering[3]. We can still define the modulus or absolute value as the *real* number $|a + ib| = \sqrt{a^2 + b^2}$. This satisfies the triangle inequality and reverse triangle inequality. The set $\{x \in \mathbb{C} : |x - a| < r\}$ is now a disc of radius r about the point a in the complex plane.

Boundedness properties and the Completeness Axiom

A subset S of $\mathbb{R}$ is **bounded above** if there exists an **upper bound**, that is a $b \in \mathbb{R}$ such that for all $x \in S$, $x \leq b$. Similarly S is **bounded below** if there exists a **lower bound** $a \in \mathbb{R}$, so that for all $x \in S$, $a \leq x$. A set S is **bounded** if it is bounded above and below. This happens if and only if there exists $M \in \mathbb{R}$ such that for all $x \in S$, $|x| \leq M$. Note that the bounds don't need to be in the set S.

The notion of boundedness (but not upper or lower bounds) also applies to sets of complex numbers: $S \subseteq \mathbb{C}$ is **bounded** if $\exists M \in \mathbb{R}\colon \forall z \in S\colon |z| \leq M$.

We assume that $\mathbb{R}$ satisfies the following.

Completeness Axiom. *A non-empty subset S of $\mathbb{R}$ which is bounded above has a least upper bound.*

The least upper bound, also called the **supremum**, of S is denoted $\sup S$ and can easily seen to be unique when it exists[4]. In symbols, $s = \sup S$ satisfies

[3] At least not an arithmetically useful one. Of course you could define any ordering you like on $\mathbb{C}$, but it would not play well with $+$ and $\times$.

[4] Sometimes it is convenient to extend the definition of sup so that $\sup S = +\infty$ for sets that are not bounded above and $\sup \emptyset = -\infty$. Similarly $\inf S = -\infty$ for sets that are not bounded below and $\inf \emptyset = +\infty$. Properties (a) and (b) then still hold with the obvious ordering conventions on $\mathbb{R} \cup \{\pm\infty\}$, and the standard mathematical interpretation of vacuous statements as being true.

(a) $\forall x \in S\colon x \le s$ (s is an upper bound)

(b) $\forall b \in \mathbb{R}\colon ((\forall x \in S\colon x \le b) \implies s \le b)$ (and it is the least one)

Combining (a) with the contrapositive of (b) we get the following.

Approximation property. If $c < \sup S$ then there exists an $x \in S$ with $c < x \le \sup S$.

The Completeness Axiom can equivalently be formulated as the assertion that non-empty subset of $\mathbb{R}$ which is bounded below has a greatest lower bound, or **infimum**, $\inf S$. Reversing all the inequalities in the properties above for sup gives corresponding properties for inf.

The Completeness Axiom underpins the deeper results in *Analysis I* and the same is true in *Analysis II*.

Intervals

A subset $I \subseteq \mathbb{R}$ is called an **interval** if whenever I contains two points, it also contains all points between them. In symbols:

$$\forall x, y, z \in \mathbb{R}\colon ((x, z \in I \text{ and } x \le y \le z) \implies y \in I). \qquad \textbf{(Interval property)}$$

One can prove using the completeness axiom (exercise, see problem sheet 1, question 1) that every interval is of one of the following forms:

$$\begin{aligned} \emptyset &:= \{\} & (-\infty, \infty) &:= \mathbb{R} \\ (a, b) &:= \{x \in \mathbb{R} : a < x < b\} & (-\infty, b) &:= \{x \in \mathbb{R} : x < b\} \\ (a, b] &:= \{x \in \mathbb{R} : a < x \le b\} & (-\infty, b] &:= \{x \in \mathbb{R} : x \le b\} \\ [a, b) &:= \{x \in \mathbb{R} : a \le x < b\} & (a, \infty) &:= \{x \in \mathbb{R} : x > a\} \\ [a, b] &:= \{x \in \mathbb{R} : a \le x \le b\} & [a, \infty) &:= \{x \in \mathbb{R} : x \ge a\} \end{aligned}$$

An interval is called **non-trivial** if it has infinitely many points, i.e., it is not empty ($\emptyset$) and not a singleton set ($[a, a] = \{a\}$). Intervals on the left in the above table are all bounded, the ones on the right are unbounded. Intervals of types $\emptyset$, (a, b), $(-\infty, b)$, (a, ∞) and $\mathbb{R}$ are called **open**. Intervals of types $\emptyset$, $[a, b]$, $(-\infty, b]$, $[a, \infty)$ and $\mathbb{R}$ are called **closed** — we will see why later.

Limits of sequences

A **sequence** of real (respectively complex, integer, ...) numbers is a function $a\colon \mathbb{N} \to \mathbb{R}$ (respectively $\mathbb{N} \to \mathbb{C}$, $\mathbb{N} \to \mathbb{Z}$, ...) which assigns to each natural number n a real (respectively complex, integer, ...) number $a(n)$, which in this context is more usually denoted a_n. We denote[5] a sequence as $(a_1, a_2, a_3, \dots)$, or $(a_n)_{n\in\mathbb{N}}$, or $(a_n)_{n=1}^{\infty}$ or more usually we just abbreviate it as (a_n).

[5] This is consistent with the notation for an ordered pair/n-tuple/vector (a_1, a_2) or $(a_1, \dots, a_n)$, which can be thought of as a function $a\colon \{1, \dots, n\} \to \mathbb{R}$ giving a real value for each 'coordinate'. Not to be confused with the *set* $\{a_1, \dots, a_n\}$ where the order does not matter and repetitions ignored.

Terms such as boundedness, supremum, etc. when applied to sequences refer to the set $\{a_n : n \in \mathbb{N}\}$ of values taken by the sequence.

The key definition in *Analysis I* is that of a limit of a sequence: a sequence (a_n) of real or complex numbers **tends to** (or **converges to**) **the limit** $\ell \in \mathbb{R}$ or $\mathbb{C}$ if[6]

$$\forall \varepsilon > 0\colon \exists N \in \mathbb{N}\colon \forall n > N\colon |a_n - \ell| < \varepsilon. \tag{1}$$

We then write $a_n \to \ell$ as $n \to \infty$ or $\lim_{n\to\infty} a_n = \ell$. We say (a_n) **converges** if there exists $\ell \in \mathbb{R}$ or $\mathbb{C}$ such that $a_n \to \ell$, otherwise we say (a_n) **diverges**.

Important fact. The limit of a convergent sequence is unique, if it exists.

Useful fact. A convergent sequence is always bounded.

Sometimes, in the definition (1) of a limit, it is neater to work with the condition $n \geq N$, or require only that $N \in \mathbb{R}$. This makes no difference as one can adjust N by 1 or replace $N \in \mathbb{R}$ with[7] $\lfloor N \rfloor$ respectively (as $n > N$ if and only if $n > \lfloor N \rfloor$).

We say a sequence (a_n) of *real* numbers **tends to infinity** if

$$\forall M \in \mathbb{R}\colon \exists N \in \mathbb{N}\colon \forall n > N\colon a_n > M.$$

We then use the notation $a_n \to \infty$ or $\lim_{n\to\infty} a_n = \infty$. A similar definition exists for $a_n \to -\infty$.

Warning. If $a_n \to \pm\infty$ we do *not* say that a_n *converges*. Also note that if (a_n) does not converge, it does *not* imply that $a_n \to \pm\infty$ (e.g., it might oscillate). Finally, $a_n \to \pm\infty$ only makes sense for *real* sequences[8] as the definition uses ordering, which is not defined on $\mathbb{C}$.

Important fact. Whether or not a sequence converges, and what its limit is, does not depend on the first few terms. Thus we only need the sequence to be defined from some point onwards — we don't have to start with a_1.

Algebra of Limits (AOL) (Real or Complex sequences). If $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then $|a_n| \to |a|$; $a_n \pm b_n \to a \pm b$; $a_n b_n \to ab$; and, provided $b \neq 0$, $a_n/b_n \to a/b$. Also constant sequences converge: if all $c_n = c$ then $c_n \to c$.

Important fact. $b_n \to b \neq 0$ implies that from some point onwards $b_n \neq 0$ (which is needed in the proof that $a_n/b_n \to a/b$).

AOL can be extended by induction to any *fixed* number of arithmetic operations. For example $a_n \to a$ implies $a_n^k \to a^k$ for a fixed $k \in \mathbb{N}$. However, it does *not* apply when we are taking an unbounded number of operations. For example $\sum_{k=1}^n \frac{1}{n} = 1 \not\to \sum_{k=1}^\infty 0 = 0$.

AOL results apply when the limits are in $\mathbb{R}$ or $\mathbb{C}$. Generalisations to real sequences which tend to $\pm\infty$ need care (and separate proofs even when they work, see later).

[6] And to fully make sense of this definition requires prior attendance of the *Analysis I* course!

[7] The **floor** function $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$ is x rounded down and the **ceiling** function $\lceil x \rceil = \min\{n \in \mathbb{Z} : n \geq x\}$ is x rounded up to the next integer. These are well-defined: see *Analysis I*.

[8] Although $|a_n| \to \infty$ makes perfect sense for complex sequences as then $|a_n|$ is real.

Limits preserve weak inequalities. If $a_n \to a$ and $b_n \to b$ and $a_n \le b_n$ then $a \le b$. (Also applies with a and/or b replaced by $\pm\infty$ with the obvious ordering conventions.)

Warning. This only applies to *weak* inequalities: $a_n < b_n$ does *not* imply $\lim a_n < \lim b_n$. Also, as inequalities are involved, this applies to *real* sequences only, as does the following.

A real sequence (a_n) is **increasing**[9] (respectively **strictly increasing**, **decreasing**, **strictly decreasing**) if $m < n$ implies $a_m \le a_n$ (respectively $a_m < a_n$, $a_m \ge a_n$, $a_m > a_n$).

Monotone limits. If (a_n) is an increasing sequence of real numbers that is bounded above, then it converges and $\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}$. If (a_n) is a decreasing sequence of real numbers that is bounded below, then it converges and $\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}$.

If (a_n) is increasing and *not* bounded above, then[10] $a_n \to +\infty$. Similarly, if (a_n) is decreasing and *not* bounded below then $a_n \to -\infty$.

Sandwiching (or **Squeeze theorem**). If $a_n \le b_n \le c_n$ and $a_n \to \ell$, $c_n \to \ell$, then $b_n \to \ell$.

To prove a version of sandwiching which also works for complex sequences we note the following.

Observation (for real or complex sequences). $a_n \to 0$ iff[11] $|a_n| \to 0$.
Indeed, it is enough to note that $||a_n| - 0| = |a_n - 0|$ in the definition (1) of a limit.

Sandwiching, alternative form. If $a_n \to \ell$, $|b_n - a_n| \le r_n$ and $r_n \to 0$, then $b_n \to \ell$.

Proof. $0 \le |b_n - a_n| \le r_n$ (for real r_n) and $r_n \to 0$ implies $|b_n - a_n| \to 0$ by sandwiching. Now $|b_n - a_n| \to 0 \Longrightarrow b_n - a_n \to 0 \Longrightarrow b_n = a_n + (b_n - a_n) \to \ell + 0 = \ell$ by AOL. □

Subsequences

A **subsequence** of a sequence (a_n) is a sequence $(a_{s_n}) = (a_{s_1}, a_{s_2}, \dots)$ where (s_n) is a strictly increasing sequence of natural numbers (i.e., $s_1 < s_2 < \cdots$).

Limits of subsequences. If $a_n \to \ell$ as $n \to \infty$ then, for any subsequence (a_{s_n}) of (a_n), $a_{s_n} \to \ell$ as $n \to \infty$.

This result is often used in the form of the contrapositive: to show a sequence does *not* converge it is enough to exhibit two subsequences that converge to different limits, or find one that does not converge at all (e.g., because it tends to $\pm\infty$).

The following is one of the main theorems of *Analysis I* — we will be needing it!

[9] Sometimes the term **non-decreasing** is used in place of **increasing** to emphasise that it not necessarily strictly increasing. Similarly **non-increasing** is the same as (not necessarily strictly) decreasing.
[10] This is one good reason to extend the definitions of inf and sup as in footnote [4] on page 2.
[11] If and only if.

Theorem 0.1 (Bolzano–Weierstrass Theorem)**.** *A bounded sequence of real or complex numbers has a convergent subsequence.*

Cauchy sequences

A real or complex sequence (a_n) is a **Cauchy sequence** if

$$\forall \varepsilon > 0\colon \exists N \in \mathbb{N}\colon \forall n, m > N\colon |a_n - a_m| < \varepsilon.$$

Theorem 0.2 (Cauchy Convergence Criterion *or* General Principle of Convergence)**.** *A sequence of real or complex numbers converges if and only if it is a Cauchy sequence.*

The Cauchy convergence criterion is extremely useful when we want to show something converges, but don't know (or can't easily describe) what the limit should be.

Series

Given a sequence (a_k) the **series** $\sum a_k$ or $\sum_{k=1}^{\infty} a_k$ is defined to be the limit (if it exists) of the sequence (s_n) of **partial sums**[12] $s_n := \sum_{k=1}^{n} a_k$ as $n \to \infty$.

Useful fact. If $\sum a_k$ converges, then $a_k \to 0$.

A series $\sum a_k$ is **absolutely convergent** if $\sum |a_k|$ converges. Note that as $\sum_{k=1}^{n} |a_k|$ is increasing in n, it is either bounded, so converges ($\sum_{k=1}^{\infty} |a_k| = \ell < \infty$), or tends to infinity ($\sum_{k=1}^{\infty} |a_k| = \infty$).

Important fact. Absolute convergence implies convergence (for both real and complex series).

Infinite triangle inequality. If $\sum |a_k|$ converges then $|\sum_{k=1}^{\infty} a_k| \leq \sum_{k=1}^{\infty} |a_k|$.

[This applies to complex as well as real series. The proof starts with induction on n to deduce the finite triangle inequality $|\sum_{k=1}^{n} a_k| \leq \sum_{k=1}^{n} |a_k|$. We then take limits, using 'absolute convergence implies convergence' so that $|\sum_{k=1}^{\infty} a_k|$ is well defined, and then 'limits preserve weak inequalities' (see *Analysis I*, problem sheet 5, question 3).]

Tests for convergence of series

Comparison Test. If $0 \leq a_n \leq b_n$ and $\sum b_n$ converges, then $\sum a_n$ converges.

Often this is applied to $|a_n|$ (a_n can now be complex) to first show $\sum |a_n|$ converges.

[12] When working both with individual terms of the series and with the partial sums of the series it is sensible to use different dummy variables: here we use k for the first and n for the second.

Comparison Test$^+$. If $|a_n| \le b_n$ and $\sum b_n$ converges, then $\sum a_n$ converges (absolutely).

Ratio Test. Assume $a_n \neq 0$ are real or complex and $|a_{n+1}/a_n| \to \ell$ as $n \to \infty$. If $0 \le \ell < 1$ then $\sum a_n$ converges; if $\ell > 1$ or $\ell = \infty$ then $\sum a_n$ diverges.

[No information is obtained when $\ell = 1$. The Ratio Test is quite weak as $|a_{n+1}/a_n|$ converging is a rather strong assumption. But when it works it is often easy to apply.]

Alternating Series Test. If (a_n) is a real (non-negative) decreasing sequence and $a_n \to 0$ then $\sum(-1)^n a_n$ converges.

[Strictly speaking, non-negativity follows automatically from the other conditions. This is the only general test listed here that can show convergence for sequences that don't converge absolutely.]

Integral Test. If $f\colon [1,\infty) \to \mathbb{R}$ is a non-negative decreasing function, then $\sum_{k=1}^{\infty} f(k)$ converges iff $\int_1^n f(x)\,\mathrm{d}x$ converges as $n \to \infty$.

[A powerful test for series with slowly decreasing positive terms.]

Power series

A **power series** is a series[13] of the form $\sum_{n=0}^{\infty} a_n x^n$ where we consider x a (real or complex) parameter that can be varied. It can be used to define a function $f(x) := \sum_{n=0}^{\infty} a_n x^n$ whenever this series converges.

The **radius of convergence (ROC)** of the power series $\sum a_n x^n$ is defined by[14]

$$R := \begin{cases} \sup\left\{|x| : \sum a_n x^n \text{ converges}\right\}, & \text{if this set is bounded;} \\ +\infty, & \text{otherwise.} \end{cases}$$

[Sometimes the definition is given in terms of $\sum |a_n x^n|$ converging — it makes no difference. It also makes no difference if x is allowed to be complex, or is restricted to real values — you get the same value of R as a consequence of the following theorem.]

Theorem 0.3. *If $\sum a_n x^n$ is a power series with ROC R and $x \in \mathbb{C}$, then*

(a) *if $|x| < R$ then $\sum a_n x^n$ converges, (and in fact it converges absolutely),*

(b) *if $|x| > R$ then $\sum a_n x^n$ diverges, (and in fact the terms $a_n x^n$ are unbounded, so do not even tend to zero).*

Proof. (a) As $|x| < R$ there exists an y with $|x| < |y| < R$ and $\sum a_n y^n$ converging (by the approximation property of sup, or by the unboundedness of the set of convergence when $R = \infty$). But then $a_n y^n \to 0$, so in particular $(a_n y^n)$ is bounded, say $|a_n y^n| \le M$. Now $\sum |a_n x^n| \le \sum M(|x|/|y|)^n$ converges by comparison with a geometric series. Absolute convergence now implies convergence of $\sum a_n x^n$.

[13] We usually start at $n = 0$ here as we want to include a constant term.

[14] Another good reason to adopt the extension of the definition on sup in footnote [4] on page 2.

(b) As $|x| > R$ there exists a y with $R < |y| < |x|$. Assume $\sum a_n x^n$ converges. Then $a_n x^n \to 0$, so $(a_n x^n)$ is bounded, say $|a_n x^n| \le M$ (which from now on is all we shall assume). As in the proof of (a) this implies that $\sum a_n y^n$ converges (absolutely), contradicting the definition of R. □

1 Functions and limits

Functions

Analysis II is a course about **functions**. Given two sets X and Y (which will usually be subsets of $\mathbb{R}$ in this course), a **function** $f\colon X \to Y$ assigns to each element x of the set X an element $f(x)$ of the set Y. Sometimes we also write $x \mapsto f(x)$. We call X the **domain** of f, or $\operatorname{dom}(f)$; and Y the **codomain**[15] of f, or $\operatorname{codom}(f)$. The **image** of f is $f(X) := \{f(x) : x \in X\}$, i.e., the set of values that are actually achieved by f. This is a subset, possibly a proper subset, of the codomain Y.

There is no expectation here that the mapping $x \mapsto f(x)$ has to be specified by a single formula, or even a formula at all. Specification of a function 'by cases' or by complicated rules will be common in this course — the modulus function is one example of this. Thus we shall allow our examples to include functions like the following:

1. $f\colon (0,1] \to \mathbb{R}$ defined by $f(x) := \begin{cases} \frac{1}{q}, & \text{if } x \text{ is rational, } x = \frac{p}{q} \text{ in lowest terms;} \\ 0, & \text{otherwise.} \end{cases}$
2. $f\colon \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ defined by $\tan(f(x)) = x$. (This defines arctan.)
3. $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} (\frac{x}{2})^{2n}$. (This is the Bessel function $J_0(x)$.)
4. For $x \in (-\infty, 2.512)$, define $a_0 = x$ and inductively $a_{n+1} = e^{a_n/2} - 1$ for $n \ge 0$. Then set $f(x) := \lim_{n\to\infty} 2^n a_n$. (A bizarre function that satisfies $f(x) = 2f(e^{x/2} - 1)$.)

We want to encompass the familiar functions of everyday mathematics: polynomials; exponential functions; trigonometric functions; hyperbolic functions — all of which can be defined on the whole of $\mathbb{R}$. We shall also encounter associated inverse functions, logarithms, arcsin, etc. You will know from *Analysis I* that many of these functions can be *defined* using power series. One of our objectives in *Analysis II* will be to develop properties of functions defined by power series (continuity, differentiability, useful inequalities and limits, ...). But until our general theory of functions has been developed far enough to cover this material *we shall make use of the standard properties we need of standard functions in our examples.*

[15] Some authors use the term **range** in place of codomain, but others use range to mean the image. I will therefore avoid using this term.

The material in this section is unashamedly technical, but necessary if we are to build firm foundations for the study of real-valued functions defined on subsets of $\mathbb{R}$, many of them having graphs neither you nor any computer software can hope to sketch effectively.

Limit points

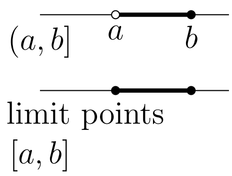
We want to define what is meant by the limit of a function. Intuitively f has a limit ℓ at the point p if the values of $f(x)$ are close to ℓ when x is close to (but not equal to) p. But for the definition of limit to be meaningful it is necessary that f is defined at 'enough' points close to p. So we are interested only in points p that x can get close to, where x is in the domain of f. This leads us to the definition of a *limit point.*

Definition. Let $E \subseteq \mathbb{R}$. A point $p \in \mathbb{R}$ is called a **limit point** (or **cluster point** or **accumulation point**) of E if E contains points $\neq p$ arbitrarily close to p. Formally:

$$\forall \varepsilon > 0 \colon \exists x \in E \colon 0 < |x - p| < \varepsilon.$$

Here p may be in E, but need not be. Note that the condition $0 < |x - p|$ is important in the case that $p \in E$ as we want points close to p that are *not equal* to p.

Example 1.1. Let $E = (a, b]$ where $a < b$. Then p is a limit point of E if and only if $p \in [a, b]$. To prove this, there are four cases to consider. If $p < a$ take $\varepsilon := |p - a|$ and get a contradiction. If $p > b$ take $\varepsilon := |p - b|$, similarly. If $p \in [a, b)$, given $\varepsilon > 0$ choose $x = p + \frac{1}{2}\min\{\varepsilon, |p - b|\}$. If $p = b$ (or indeed any $p \in (a, b]$) we can take $x = p - \frac{1}{2}\min\{\varepsilon, |p - a|\}$.
The same conclusion holds when $E = (a, b)$, $E = [a, b)$ or $E = [a, b]$.

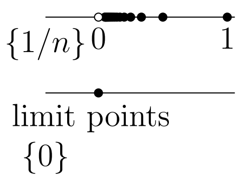


Definition. A set E is called **closed** if it contains all its limit points. A set E is called **open** if it is the complement of a closed set, or equivalently:

$$\forall p \in E \colon \exists \delta > 0 \colon (p - \delta, p + \delta) \subseteq E.$$

Exercise. Check the 'equivalently' condition is indeed equivalent, and that the definitions of open and closed are consistent with the terminology used for intervals on page 3.

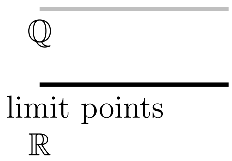
Example 1.2. Let $E = \{\frac{1}{n} : n \in \mathbb{N}\}$. Here p is a limit point of E if and only if $p = 0$. Indeed, if $p < 0$ or $p > 1$ take $\varepsilon = |p|$ or $|p - 1|$ respectively. If $\frac{1}{n+1} < p < \frac{1}{n}$ take $\varepsilon = \min\{p - \frac{1}{n+1}, \frac{1}{n} - p\}$. If $p = \frac{1}{n}$ then take $\varepsilon = \frac{1}{n(n+1)}$ and note that there is no *other* point of E within distance ε of p. If $p = 0$ then for any $\varepsilon > 0$ pick $n > \frac{1}{\varepsilon}$ so that $|\frac{1}{n} - 0| < \varepsilon$.



Definition. An **isolated** point of a set E is a point of E that is *not* a limit point of E, i.e., it is a point $p \in E$ such that for some $\delta > 0$, $(p - \delta, p + \delta) \cap E = \{p\}$.

For example, all the points of E in Example 1.2 are isolated points.

Example 1.3. Let $E = \mathbb{Q}$. Then for *any* point $p \in \mathbb{R}$, p is a limit point of E as there are rationals in any set of the form $(p, p + \varepsilon)$. Note that none of the points in $\mathbb{Q}$ are isolated. A similar argument also applies to the set $E = \mathbb{R} \setminus \mathbb{Q}$ of irrational numbers.



The notion of limit point is important well beyond the present course, in which we shall encounter only simple instances of it. Much more exotic examples exist. The structure of the real line is rich, with $\mathbb{R}$ having many subsets which are very complicated. Such complexities are important in topology and measure theory for example. The following gives a simple criterion for limit points.

Proposition 1.4 (Limit points via sequences)**.** *A point $p \in \mathbb{R}$ is a limit point of $E \subseteq \mathbb{R}$ if and only if there exists a sequence (p_n) of points with $p_n \in E$, $p_n \neq p$, such that $\lim_{n\to\infty} p_n = p$.*

Proof. If p is a limit point of E then for any $n \in \mathbb{N}$ choose $\varepsilon := \frac{1}{n}$. Then there exists $p_n \in E$ such that $0 < |p_n - p| < \frac{1}{n}$. Now $p_n \to p$ as $n \to \infty$ (by sandwiching), and $p_n \in E$ and $p_n \neq p$ (by assumption).

Conversely, if such a sequence (p_n) exists, given $\varepsilon > 0$, $\exists N \in \mathbb{N} \colon \forall n \geq N \colon |p_n - p| < \varepsilon$. So in particular $p_N \in E$ and $0 < |p_N - p| < \varepsilon$ as $p_N \neq p$. $\square$

Corollary 1.5 (Closed sets are closed under limits)**.** *If $E \subseteq \mathbb{R}$ is closed and $p_n \in E$ with $p_n \to p \in \mathbb{R}$ as $n \to \infty$, then $p \in E$.*

Proof. Either $p = p_n \in E$ for some n, so $p \in E$; or $p \neq p_n$ for all n in which case p is a limit point of E by Proposition 1.4, and hence in E as E is closed. $\square$

Proposition 1.4 together with Example 1.3 gives the following useful consequences.

- Given $x \in \mathbb{R}$, there exists a sequence (r_n) of rational numbers such that $r_n \to x$.
- Given $x \in \mathbb{R}$, there exists a sequence (q_n) of irrational numbers such that $q_n \to x$.

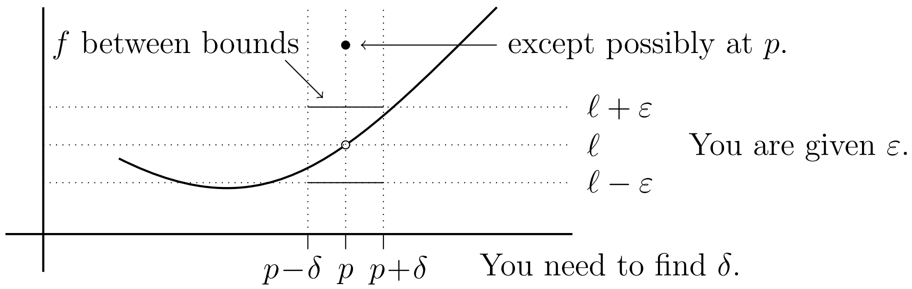
Limits of functions

Now we come to the most important definition in this course.

Definition. Let $E \subseteq \mathbb{R}$ and $f \colon E \to \mathbb{R}$ be a real-valued function. Let p be a limit point of E and let $\ell \in \mathbb{R}$. We say that f **tends to** (or **converges to**) ℓ **as** x **tends to** p if

$$\forall \varepsilon > 0 \colon \exists \delta > 0 \colon \forall x \in E \colon (0 < |x - p| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon). \tag{2}$$

In words: given any $\varepsilon > 0$ we can find a $\delta > 0$ such that $f(x)$ will be within distance ε of ℓ for any $x \in E$, $x \neq p$, that is within distance δ of p.



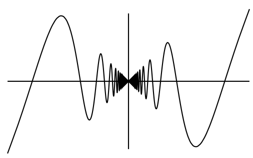
We also write this as $\lim_{x\to p} f(x) = \ell$ or $f(x) \to \ell$ as $x \to p$. If one needs to emphasise the domain E we can write this more formally as

$$\lim_{\substack{x\to p\\ x\in E}} f(x) = \ell.$$

We say $f(x)$ **converges** as $x \to p$ if p is a limit point of E and $\lim_{x\to p} f(x) = \ell$ for some $\ell \in \mathbb{R}$. Otherwise we say $f(x)$ **diverges** as $x \to p$.

Note that, in the definition, δ may, and almost always will, depend on ε.

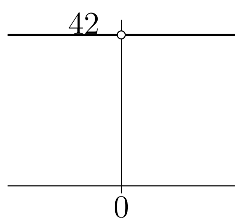
Important note. In the limit definition it may or may not happen that f is defined *at* p. And when $f(p)$ is defined, its value has no influence on whether or not $\lim_{x\to p} f(x)$ exists. Moreover, when the limit ℓ does exist and $f(p)$ is defined, there is no reason to assume that $f(p)$ will equal ℓ.



Example 1.6. Let $\alpha > 0$. Consider the function $f(x) = |x|^\alpha \sin\frac{1}{x}$ on the domain $E := \mathbb{R} \setminus \{0\}$. We claim that $f(x) \to 0$ as $x \to 0$. Since $|\sin\theta| \leq 1$ for any $\theta \in \mathbb{R}$, we have $\left||x|^\alpha \sin\frac{1}{x}\right| \leq |x|^\alpha$ for any $x \neq 0$. For any $\varepsilon > 0$, choose $\delta := \varepsilon^{1/\alpha} > 0$. Then for $0 < |x-0| < \delta$

$$\left||x|^\alpha \sin\tfrac{1}{x} - 0\right| \leq |x|^\alpha < \delta^\alpha = \varepsilon.$$

According to the definition, $|x|^\alpha \sin\frac{1}{x} \to 0$ as $x \to 0$.

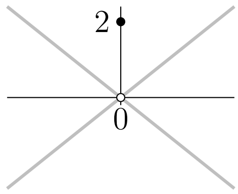


Example 1.7. Let f be defined on $E = \mathbb{R} \setminus \{0\}$ by $f(x) = 42$. Then 0 is a limit point of E. Let $\ell = 42$. Then for $x \neq 0$ we have $|f(x) - \ell| = 0$. So, for any $\varepsilon > 0$, we can take $\delta = 1$, say, to get that

$$0 < |x-0| < \delta \implies |f(x) - \ell| = 0 < \varepsilon.$$

So $f(x) \to 42$ as $x \to 0$.

Example 1.8. Let f be defined on $\mathbb{R}$ by



$$f(x) := \begin{cases} x, & \text{if } x \in \mathbb{Q},\ x \neq 0; \\ 2, & \text{if } x = 0; \\ -x, & \text{otherwise.} \end{cases}$$

We claim that $f(x) \to 0$ as $x \to 0$. To prove this, simply note that $|f(x) - 0| = |x| < \varepsilon$ if $0 < |x - 0| < \delta := \varepsilon$. (Here, following the definition of limit, we omit consideration of $f(0)$, even though f is defined at 0.)

Example 1.9. Consider the function $f(x) = x^2$ on the domain $E = \mathbb{R}$. Let $a \in \mathbb{R}$. We claim that $f(x) \to a^2$ as $x \to a$.

Note that $|x^2 - a^2| = |x - a||x + a|$. We want this to be small when x is close to a. Suppose that $|x - a| < 1$. Then

$$|x + a| = |x - a + 2a| \leq |x - a| + |2a| < 1 + 2|a|.$$

So given $\varepsilon > 0$, choose $\delta := \min\{\frac{\varepsilon}{1+2|a|}, 1\} > 0$. Then if $0 < |x - a| < \delta$ we have

$$|x^2 - a^2| \leq |x - a|(1 + 2|a|) < \delta(1 + 2|a|) \leq \varepsilon.$$

This example serves to illustrate that going back to first principles to establish the limiting value of a function may be a tedious task. Help will soon be at hand.

Remark. We saw in Example 1.9 that when considering a limit $x \to p$ we can restrict attention to x close to p, say $|x - p| < \delta_0$. Any subsequent δ that we find then just has to be replaced by $\min\{\delta, \delta_0\}$ in definition (2) to make it work for all x.

Why do we not consider $f(p)$? One of our main motivations for considering function limits stems from differential calculus. The recipe from school calculus of the derivative of f can be cast in the form

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) := \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}.$$

Clearly here we need δx to be non-zero as otherwise the quotient is undefined. To provide a uniform and consistent theory of limits that includes this case, we therefore systematically exclude $f(p)$ from consideration.

The following result validates our definitions and notation. Compare with the corresponding result for sequences and its proof.

Proposition 1.10 (Uniqueness of function limits)**.** *Let $f \colon E \to \mathbb{R}$ and p be a limit point of E. If f has a limit as $x \to p$, then this limit is unique.*

Proof. Suppose $f(x) \to \ell_1$ and also $f(x) \to \ell_2$ as $x \to p$, where $\ell_1 \neq \ell_2$. We now apply the definition of a limit with $\varepsilon := \frac{1}{2}|\ell_1 - \ell_2| > 0$:

$$\exists \delta_1 > 0 \colon \forall x \in E \colon (0 < |x - p| < \delta_1 \implies |f(x) - \ell_1| < \varepsilon),$$

$$\exists \delta_2 > 0\colon \forall x \in E\colon (0 < |x - p| < \delta_2 \Longrightarrow |f(x) - \ell_2| < \varepsilon).$$

Let $\delta := \min\{\delta_1, \delta_2\} > 0$. Since p is a limit point of E and $\delta > 0$, $\exists x \in E$ such that $0 < |x - p| < \delta$. Then for this x both $|f(x) - \ell_1| < \varepsilon$ and $|f(x) - \ell_2| < \varepsilon$ hold, and so

$$\begin{aligned}
|\ell_1 - \ell_2| &= |(f(x) - \ell_2) - (f(x) - \ell_1)| && \text{[add and subtract technique]}\\
&\leq |f(x) - \ell_2| + |f(x) - \ell_1| && \text{[triangle inequality]}\\
&< \varepsilon + \varepsilon\\
&= |\ell_1 - \ell_2|, && \text{[choice of } \varepsilon\text{]}
\end{aligned}$$

and we have a contradiction. □

Why do we need limit points? Note how the above proposition used the fact that p was a limit point of E. Indeed, if p was not a limit point then $\lim_{x\to p} f(x) = \ell$ would hold *vacuously* for *every* $\ell \in \mathbb{R}$ as we could just take δ small enough so that no point of E satisfied $0 < |x - p| < \delta$. Thus we need p to be a limit point to make the definition of limits non-trivial. In particular, when we say $f(x)$ *converges* as $x \to p$, we always insist that p is a limit point (see problem sheet 1, question 4(d), for a case when this is important).

Notice that all the examples presented so far have shown that function limits do exist. Now let's explore how to prove that a limit fails to exist. The proof of the following result illustrates how to work with the contrapositive of the limit definition. The proposition translates questions about function limits to questions about sequence limits, and vice versa, and so allows to draw on results from *Analysis I*. Note the care needed to handle the $x \neq p$ condition.

Proposition 1.11 (Function limits via sequences)**.** *Let $f\colon E \to \mathbb{R}$ where $E \subseteq \mathbb{R}$, and assume p is a limit point of E. Then the following are equivalent.*

(a) $\lim_{x\to p} f(x) = \ell$.

(b) $\lim_{n\to\infty} f(p_n) = \ell$ *for all sequences* (p_n) *with* $p_n \in E$, $p_n \neq p$ *and* $\lim_{n\to\infty} p_n = p$.

Proof. Suppose $\lim_{x\to p} f(x) = \ell$ and fix $\varepsilon > 0$. Then there exists a $\delta > 0$ such that

$$\forall x \in E\colon (0 < |x - p| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon).$$

Now suppose (p_n) is a sequence in E, with $p_n \to p$ and $p_n \neq p$. Then, taking the ε in the definition (1) of convergence of a sequence to be this δ, we have

$$\exists N \in \mathbb{N}\colon \forall n > N\colon |p_n - p| < \delta.$$

Putting the conditions together and using the fact that $p_n \in E$ and $p_n \neq p$ (so $0 < |p_n - p|$) we get

$$\exists N \in \mathbb{N}\colon \forall n > N\colon |f(p_n) - \ell| < \varepsilon.$$

As this holds for any $\varepsilon > 0$, $\lim_{n\to\infty} f(p_n) = \ell$ by definition.

Conversely, suppose $f(x) \not\to \ell$ as $x \to p$. Then[16]

$$\exists \varepsilon > 0 \colon \forall \delta > 0 \colon \exists x \in E \colon (0 < |x - p| < \delta \text{ and } |f(x) - \ell| \geq \varepsilon).$$

Fix such an $\varepsilon > 0$ and choose $\delta := \frac{1}{n}$. Then $\exists p_n \in E$, with $0 < |p_n - p| < \frac{1}{n}$ and

$$|f(p_n) - \ell| \geq \varepsilon.$$

Thus we have found a sequence $p_n \in E$, $p_n \neq p$, with $p_n \to p$ (by sandwiching), and for which $f(p_n) \not\to \ell$, as required. □

Proposition 1.11 can be used to show that a limit $\lim_{x \to p} f(x)$ does not exist by finding two rival values for the limit, assuming it did exist.

Example 1.12. Consider the function f defined in Example 1.8, namely

$$f(x) := \begin{cases} x, & \text{if } x \in \mathbb{Q},\ x \neq 0; \\ 2, & \text{if } x = 0; \\ -x, & \text{otherwise.} \end{cases}$$

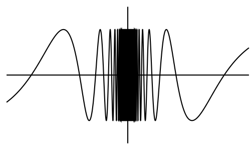
We claim that, for any $p \neq 0$, the limit $\lim_{x \to p} f(x)$ fails to exist.

Assume $p \neq 0$. Then as p is a limit point of $\mathbb{Q} \setminus \{0\}$ (Example 1.3 with trivial modification to avoid 0) there exists (by Proposition 1.4) a sequence (p_n) such that $p_n \in \mathbb{Q} \setminus \{0\}$, $p_n \neq p$ and $p_n \to p$. Similarly there exists a sequence (q_n) such that $q_n \in \mathbb{R} \setminus \mathbb{Q}$, $q_n \neq p$ and $q_n \to p$. Then

$$f(p_n) = p_n \to p \qquad \text{and} \qquad f(q_n) = -q_n \to -p.$$

Now if $\lim_{x \to p} f(x) = \ell$ then, by Proposition 1.11 and the uniqueness of sequence limits, both $\ell = p$ and $\ell = -p$ would hold, a contradiction as $p \neq 0$.

Example 1.13. To show that $\lim_{x \to 0} \sin \frac{1}{x}$ doesn't exist.
Let $f(x) = \sin \frac{1}{x}$ for $x \neq 0$. Let $p_n = \frac{1}{2\pi n}$ and $q_n = \frac{1}{2n\pi + \pi/2}$. Then both sequences (p_n) and (q_n) tend to 0 and $p_n, q_n \neq 0$, but



$$\lim_{n \to \infty} \sin \tfrac{1}{x_n} = \lim_{n \to \infty} \sin(2n\pi) = 0 \qquad \text{and} \qquad \lim_{n \to \infty} \sin \tfrac{1}{y_n} = \lim_{n \to \infty} \sin(2n\pi + \tfrac{\pi}{2}) = 1.$$

So $\lim_{x \to 0} \sin \frac{1}{x}$ cannot exist.

Generalisations to complex numbers and vectors

The definitions of limit points and limits, together with Propositions 1.4, 1.10, 1.11 and Corollary 1.5 extend immediately to $\mathbb{C}$, and indeed to vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$, with essentially

[16] Note how the negation is obtained by swapping ∀s and ∃s and negating the final statement, keeping the quantifiers in the same order.

identical proofs. We simply need to replace the real modulus with the complex modulus $|z|$, or with the length $|\mathbf{x}|$ of a vector $\mathbf{x}$ in $\mathbb{R}^n$ or $\mathbb{C}^n$, given in the usual way as

$$|\mathbf{x}| = |(x_1, \dots, x_n)| = \sqrt{|x_1|^2 + \dots + |x_n|^2}.$$

The only properties of $|.|$ that we need are $|x| \geq 0$, with equality iff $x = 0$, plus the triangle inequality (which implies the reverse triangle inequality), and these hold in all the above cases. Thus we can define limits for functions $\mathbb{C} \to \mathbb{C}$, $\mathbb{R} \to \mathbb{C}$, $\mathbb{C} \to \mathbb{R}$, $\mathbb{R} \to \mathbb{R}^n$, $\mathbb{R}^n \to \mathbb{R}^m$, etc.

It is worth remarking that functions of more than one variable, such as $f(x, y)\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ are just functions of a 'vector' $(x, y) \in \mathbb{R}^2$, and hence we have also defined multi-variable limits such as

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y).$$

As this course is principally about real functions of one variable, we will not dwell on these extensions too much in this course. One exception will be when we discuss continuity of functions of several variables. Another is when we come to power series, which are of extreme importance in complex analysis. In that case we will phrase our results in terms of complex series. Nevertheless, it is worth noting that much of the material in this course does generalise, *except* for the material in sections 4, 5, 9, 10 and 11, which are only valid for real functions of one real variable.

Infinite limits

As for sequences, we sometimes want to consider the case when the function 'tends to infinity'. Note that although it appears in our vocabulary, we have *not* given infinity the status of a number: it can only appear in certain phrases in our mathematical language which are shorthand for quite complicated statements about real numbers. Also in this case we can only consider functions whose codomain is $\mathbb{R}$ as we will need to use ordering.[17]

We follow the same idea used for sequence limits — we replace 'close to ℓ' with 'large enough'. That is, we replace

$$\forall \varepsilon > 0 \dots \implies |f(x) - \ell| < \varepsilon$$

with

$$\forall M \dots \implies f(x) > M \qquad \text{or} \qquad \forall M \dots \implies f(x) < M$$

depending on whether $\ell = \infty$ or $\ell = -\infty$. So, for example, $\lim_{x\to p} f(x) = \infty$ means

$$\forall M \in \mathbb{R}\colon \exists \delta > 0\colon \forall x \in E\colon (0 < |x - p| < \delta \implies f(x) > M),$$

and we also write this as $f(x) \to \infty$ as $x \to p$, or $\lim_{x\to p} f(x) = \infty$, or say $f(x)$ **tends to** ∞ as x tends to p.

Warning. As for sequences, we *don't* say $f(x)$ *converges* when $f(x) \to \pm\infty$. And again, as for sequences, $f(x)$ not converging does *not* imply $f(x) \to \pm\infty$ (e.g., Example 1.13).

Note that uniqueness of limits (Proposition 1.10) and limits via sequences (Proposition 1.11) extend naturally to include $\ell = \pm\infty$ with only minor changes in the proofs.

Example 1.14. $\frac{1}{x^2} \to \infty$ as $x \to 0$. Indeed, given $M \geq 1$ we can set $\delta := \frac{1}{\sqrt{M}}$ and note that $0 < |x - 0| < \delta$ implies $\frac{1}{x^2} > \frac{1}{\delta^2} = M$. On the other hand $\frac{1}{x} \not\to \infty$ (why?), but we do have $\frac{1}{|x|} \to \infty$ as $x \to 0$.

Left-hand and right-hand limits

The way we have defined limits means that statements such as $\lim_{x\to 0} \sqrt{x} = 0$ make sense, even though the domain of $\sqrt{x}$ does not include some points very close to 0 (because they are negative). However, even if a function is defined on both sides of a point p, we may sometimes wish to consider limits taking into account only the values $f(x)$ for $x < p$, or only the values $f(x)$ for $x > p$.

Definition. Let $f\colon E \to \mathbb{R}$ and let $p \in \mathbb{R}$. Then we define the **left-hand limit** (or **limit from the left**) $\lim_{x\to p^-} f(x)$ as the limit as $x \to p$, if it exists, of the function f restricted to $E \cap (-\infty, p)$. In other words,

$$\lim_{x\to p^-} f(x) = \ell \qquad \iff \qquad \lim_{\substack{x\to p\\ x\in E\cap(-\infty,p)}} f(x) = \ell.$$

Writing this out in terms of quantifiers, this is equivalent to

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (p - \delta < x < p \implies |f(x) - \ell| < \varepsilon).$$

Similarly define the **right-hand limit** (or **limit from the right**) $\lim_{x\to p^+} f(x)$ as the limit as $x \to p$, if it exists, of the function f restricted to $E \cap (p, \infty)$. In other words,

$$\lim_{x\to p^+} f(x) = \ell \qquad \iff \qquad \lim_{\substack{x\to p\\ x\in E\cap(p,\infty)}} f(x) = \ell \qquad \iff$$

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (p < x < p + \delta \implies |f(x) - \ell| < \varepsilon).$$

Naturally these definitions are only non-vacuous if p is a limit point of $E \cap (-\infty, p)$ (p is a **left limit point** of E) or $E \cap (p, \infty)$ (p is a **right limit point** of E) respectively.

Normally here E will be an interval with p in the interior, but sometimes we write, for example, $\lim_{x\to 0^+} \sqrt{x} = 0$ instead of $\lim_{x\to 0} \sqrt{x} = 0$ in cases where p is an endpoint of the domain of the function, just to emphasise that we only need the function to be defined on one side of p.

[17] Although one can always talk about $|f(z)|$ tending to infinity when $f(z)$ is complex.

Sometimes we will use the notation $f(p^-)$ and $f(p^+)$ for the left- and right-hand limits:

$$f(p^-) := \lim_{x\to p^-} f(x), \qquad f(p^+) := \lim_{x\to p^+} f(x).$$

The proof of the following claim is good practice in using the definitions.

Proposition 1.15. *Let $f\colon E \to \mathbb{R}$ and let $p \in \mathbb{R}$ be both a left and right limit point of E. Then for any $\ell \in \mathbb{R} \cup \{\pm\infty\}$ the following are equivalent:*

(a) $\lim_{x\to p} f(x) = \ell$;

(b) *Both* $\lim_{x\to p^+} f(x) = \ell$ *and* $\lim_{x\to p^-} f(x) = \ell$.

Proof. Exercise (need separate proofs for $\ell = \pm\infty$!). See also Proposition 2.14 below. □

Example 1.16. Continuing Example 1.14: $\lim_{x\to 0^+} \frac{1}{x} = +\infty$ and $\lim_{x\to 0^-} \frac{1}{x} = -\infty$.

Limits at infinity

Sometimes we want to extend the notion '$f(x) \to \ell$ as $x \to p$' to cover $p = \pm\infty$. We note that the domain E must[18] be a subset of $\mathbb{R}$ as we will be using ordering. The natural analogue of the definition of a limit is to replace 'sufficiently close to p' with 'sufficiently large', i.e., replace

$$\exists\delta > 0 \dots\ 0 < |x - p| < \delta \implies \dots$$

with

$$\exists N \dots\ x > N \implies \dots \qquad \text{or} \qquad \exists N \dots\ x < N \implies \dots$$

depending on whether $p = +\infty$ or $p = -\infty$. Thus $\lim_{x\to\infty} f(x) = \ell$ means

$$\forall\varepsilon > 0\colon \exists N \in \mathbb{R}\colon \forall x \in E\colon (x > N \implies |f(x) - \ell| < \varepsilon).$$

Note that we do not need to include the requirement that $x \neq p = \pm\infty$ here as, by assumption, f is only defined on *real* numbers $E \subseteq \mathbb{R}$.

We do have to add a condition analogous to p being a limit point so as to make the statement $\lim_{x\to\infty} f(x) = \ell$ non-vacuous. In this case we need that E **is not bounded above** so that there are always some $x \in E$ with $x > N$. Similarly, for $\lim_{x\to-\infty} f(x)$ we need that E **is not bounded below**.

The observant reader will have noticed that if $E = \mathbb{N}$ so that $f\colon \mathbb{N} \to \mathbb{R}$ is a sequence, then the definition of $\lim_{n\to\infty} f(n) = \ell$ is just the same[19] as the one given in *Analysis I*.

Example 1.17 (Integer powers)**.** Let m be a positive integer. Then, as $x \to \infty$, the power $x^m \to \infty$; and as $x \to -\infty$, $x^m \to \infty$ if m is even and $x^m \to -\infty$ if m is odd. Moreover $x^{-m} \to 0$ as $x \to \pm\infty$.

[18] One can however define $\lim_{|z|\to\infty} f(z)$ in a fairly obvious way for functions defined on $E \subseteq \mathbb{C}$. Indeed, $\lim_{z\to\infty} f(z)$ is often *defined* this way in this case, although it causes conflict in notation when $E \subseteq \mathbb{R}$.

[19] The definition given in *Analysis I* assumed $N \in \mathbb{N}$, but one can always just replace $N \in \mathbb{R}$ with $\lfloor N \rfloor$ to get an equivalent statement.

Proof. For $m > 0$ and $M \in \mathbb{R}$ we note that for $x > N := \max\{M, 1\}$ we have $x^m \geq x > M$. So by definition $x^m \to \infty$ as $x \to \infty$. Now given $\varepsilon > 0$ we note that for $x > N := \max\{\frac{1}{\varepsilon}, 1\}$ we have $|x^{-m} - 0| = \frac{1}{x^m} \leq \frac{1}{x} < \frac{1}{N} \leq \varepsilon$ so $x^{-m} \to 0$. The cases when $x \to -\infty$ are similar, but needs some care with the signs. $\square$

Remark. When considering limits as $x \to \infty$ we can restrict attention to values of x that are large enough, say $x > M_0$. Any final M that we obtain can then be replaced by $\max\{M, M_0\}$ in the definition of a limit so that it works for all x. The above proof used this to restrict to the case $x > 1$ where the inequalities were easier.

Propositions 1.4, 1.10 and 1.11 extend simply to $p = \pm\infty$ with only minor modifications: we need to replace 'p is a limit point of E' by 'E is unbounded above/below' for $p = +\infty$ or $-\infty$ respectively. We can also drop the condition $p_n \neq p$ as $p_n \in \mathbb{R}$.

2 Basic properties of limits

Our next task is to set up the basic machinery for working with function limits. The following is perhaps the most useful result.

Theorem 2.1 (Algebra of Limits (AOL) for functions)**.** *Let $E \subseteq \mathbb{R}$ and let p be a limit point of E. Let $f, g\colon E \to \mathbb{R}$ and suppose that $f(x) \to a$ and $g(x) \to b$ as $x \to p$. Then*

$$|f(x)| \to |a|, \qquad f(x) + g(x) \to a + b, \qquad f(x) - g(x) \to a - b,$$

$$f(x)g(x) \to ab \qquad \text{and} \qquad f(x)/g(x) \to a/b \quad (\text{if } b \neq 0)$$

as $x \to p$. Also, if $h(x) := c$ is a constant function on E then $h(x) \to c$ as $x \to p$.

Proof. These can all be deduced from the Algebra of Limits for Sequences using Proposition 1.11. Assume (p_n) is any sequence with $p_n \in E$, $p_n \neq p$ and $p_n \to p$. Then by Proposition 1.11,

$$f(p_n) \to a \quad \text{and} \quad g(p_n) \to b.$$

We note that if $b \neq 0$ then by taking $\varepsilon := |b| > 0$ in the definition of the limit, there is some $\delta > 0$ such that $g(x) \neq 0$ when $0 < |x - p| < \delta$. Hence $f(x)/g(x)$ is defined on some $E' \supseteq \{x \in E : 0 < |x - p| < \delta\}$, which still has p as a limit point. By AOL for sequences,

$$|f(p_n)| \to |a|, \qquad f(p_n) + g(p_n) \to a + b, \qquad f(p_n) - g(p_n) \to a - b,$$

$$f(p_n)g(p_n) \to ab, \qquad f(p_n)/g(p_n) \to a/b \quad (b \neq 0), \qquad h(p_n) \to c.$$

As this holds for all such sequences (p_n), Proposition 1.11 implies the results. $\square$

Alternatively Theorem 2.1 can be proved directly from the definitions: mimic the proofs given for sequences in *Analysis I*. (Change "$\exists N\colon \forall n\colon n > N \implies$" to "$\exists \delta\colon \forall x \in E\colon 0 < |x - p| < \delta \implies$" throughout.)

Generalisations. AOL works for complex functions with no change in the proofs. One can even extend it to functions on $\mathbb{R}^n$ or $\mathbb{C}^n$ (and so functions of several variables), or functions *to* $\mathbb{R}^n$ or $\mathbb{C}^n$, provided the statements make sense (e.g., we can't divide two vectors, but we can, for example, multiply a scalar valued function $f(x)$ by a vector valued function $\vec{g}(x)$).

AOL and infinity. AOL works when $x \to p = \pm\infty$ with only minor changes in the proof. However, for cases when the actual limits a and/or b are infinite we need to be a bit more careful. AOL works with the obvious interpretation of arithmetic operations involving $\pm\infty$, except in the **indeterminate** cases: $\infty - \infty$, $\pm\infty \cdot 0$, $(\pm\infty)/(\pm\infty)$, and *any* case of division by 0 (at least unless one makes further assumptions on f and g). See problem sheet 1 question 4 for some examples. Since it is so useful, we will state it as a theorem.

Theorem 2.2 (Extended AOL)**.** *Let $E \subseteq \mathbb{R}$ and let p be a limit point of E or let $p = \pm\infty$ with E unbounded above/below. Let $f, g\colon E \to \mathbb{R}$ and suppose that $f(x) \to a$ and $g(x) \to b$ as $x \to p$ where $a, b \in \mathbb{R} \cup \{\pm\infty\}$. Then, as $x \to p$,*

(a) *$|f(x)| \to |a|$ where we interpret $|\pm\infty| = +\infty$;*

(b) *$f(x) \pm g(x) \to a \pm b$, except when we get $\infty - \infty$ or $-\infty + \infty$. Here $a \pm b$ is interpreted as $\pm\infty$ in the obvious way when one of a or b is infinite, or both are infinite and are 'pushing' in the same direction.*

(c) *$f(x)g(x) \to ab$, except when we get $(\pm\infty)\cdot 0$ or $0 \cdot (\pm\infty)$. Here ab is interpreted as $\pm\infty$ in the obvious way when a and/or b is infinite and neither is zero.*

(d) *$f(x)/g(x) \to a/b$ provided $b \neq 0$ and except when we get $(\pm\infty)/(\pm\infty)$. Here we interpret $a/(\pm\infty) = 0$ (for a finite) and $(\pm\infty)/b = \pm\infty$ or $\mp\infty$ (for b finite, $b > 0$ or $b < 0$).*

Proof. A rather tedious exercise — there are many different cases to check! □

Example 2.3 (Polynomials)**.** Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a real polynomial with $a_n > 0$, $n > 0$. Then $p(x) \to \infty$ as $x \to \infty$; and $p(x) \to \infty$ (n even) or $p(x) \to -\infty$ (n odd) as $x \to -\infty$.

Proof. Write $p(x) = x^n q(x)$ where $q(x) = a_n + a_{n-1}x^{-1} + \cdots + a_0 x^{-n}$. As $x^{-m} \to 0$ as $x \to \infty$ for $m > 0$ (Example 1.17), repeated use of AOL gives $q(x) \to a_n$ as $x \to \infty$. Now use the Extended AOL together with Example 1.17 to show $p(x) = x^n q(x) \to \pm\infty$. □

Example 2.4 (Rational functions)**.** A **rational function** is a quotient of two polynomials:

$$f(x) = \frac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0}{b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0}.$$

Assume $a_n, b_m \neq 0$. Then when taking a limit $x \to \infty$ we can rewrite f as

$$f(x) = x^{n-m} \cdot \frac{a_n + a_{n-1}x^{-1} + \cdots + a_0 x^{-n}}{b_m + b_{m-1}x^{-1} + \cdots + b_0 x^{-m}}.$$

As $x \to \infty$, $x^{-k} \to 0$ for $k > 0$, so by AOL

$$\frac{a_n + a_{n-1}x^{-1} + \cdots + a_0 x^{-n}}{b_m + b_{m-1}x^{-1} + \cdots + b_0 x^{-m}} \to \frac{a_n}{b_m}.$$

Thus, by Extended AOL and Example 1.17,

$$\lim_{x\to\infty} f(x) = \begin{cases} 0, & \text{if } n < m; \\ \frac{a_n}{b_m}, & \text{if } n = m; \\ \pm\infty, & \text{if } n > m; \end{cases}$$

where the $\pm$ in the last case is given by the sign of a_n/b_m.

We also have the following tools, just as for sequences.

Theorem 2.5 (Limits preserve weak inequalities)**.** *Let $f, g\colon E \to \mathbb{R}$ and let p be a limit point of E. If $f(x) \le g(x)$ for all $x \in E$ and $f(x) \to a$, $g(x) \to b$ as $x \to p$, then $a \le b$.*

Theorem 2.6 (Sandwiching)**.** *Let $f, g, h\colon E \to \mathbb{R}$ and let p be a limit point of E. If for all $x \in E$, $f(x) \le g(x) \le h(x)$ and $f(x) \to \ell$, $h(x) \to \ell$ as $x \to p$ then $g(x) \to \ell$ as $x \to p$.*

Theorem 2.7 (Sandwiching, alternative form)**.** *Let $f, g, h\colon E \to \mathbb{R}$, and let p be a limit point of E. If $f(x) \to \ell$ as $x \to p$ and $|f(x) - g(x)| \le h(x)$ with $h(x) \to 0$ as $x \to p$, then $g(x) \to \ell$ as $x \to p$.*

Proofs. Exercise. Apply Proposition 1.11 to the sequence versions. □

These generalise to $E \subseteq \mathbb{C}$ etc., and to cases where p and/or ℓ are $\pm\infty$. The alternative form of sandwiching also also works when f and g are complex or vector-valued. The following can also be extended to complex or vector-valued functions. (Note the importance of condition (b)!)

Theorem 2.8 (Limits of Compositions of Functions)**.** *Suppose $f\colon E \to \mathbb{R}$ and $g\colon E' \to \mathbb{R}$ with $f(E) \subseteq E'$ (so that $g(f(x))$ is defined for all $x \in E$). Let p be a limit point of E and assume*

(a) $\lim_{x\to p} f(x) = q \in \mathbb{R}$*; and*

(b) $f(x) \ne q$ *for all* $x \in E \setminus \{p\}$.

Then q is a limit point of E'. If in addition

(c) $\lim_{y\to q} g(y) = \ell \in \mathbb{R} \cup \{\pm\infty\}$

then we have $\lim_{x\to p} g(f(x)) = \ell$.
Corresponding statements also hold when p and/or $q = \pm\infty$.

Proof. We will just prove the case when $p, q, \ell \in \mathbb{R}$ and leave the formulation and proof of the other cases as exercises.

First, as p is a limit point of E, Proposition 1.4 implies that there is a sequence $p_n \to p$ with $p_n \in E \setminus \{p\}$. But then $q_n := f(p_n) \to q$ by Proposition 1.11. But $q_n \neq q$ and $q_n \in E'$ by assumption, so q is a limit point of E', again by Proposition 1.4.

Now suppose $\lim_{y \to q} g(y) = \ell$. Fix $\varepsilon > 0$ and choose $\eta > 0$ so that $0 < |y - q| < \eta$ implies $|g(y) - \ell| < \varepsilon$. Taking this η as the ε in the definition of $\lim_{x \to p} f(x) = q$ we can find a $\delta > 0$ such that $0 < |x - p| < \delta$ implies $|f(x) - q| < \eta$. As we are assuming $f(x) \neq q$, we actually have $0 < |f(x) - q| < \eta$ and hence $|g(f(x)) - \ell| < \varepsilon$. As this holds for any $\varepsilon > 0$ we have $\lim_{x \to p} g(f(x)) = \ell$. □

Example 2.9. Theorem 2.8 may seem a bit complicated, but it often naturally appears in arguments about limits when we 'change variables'. For example, consider the statement

$$\lim_{x \to x_0} g(x) = \lim_{h \to 0} g(x_0 + h).$$

Here we take the statement to mean that if either limit exists then so does the other and they are equal. A direct proof is easy, but one can also use Theorem 2.8.

In one direction, suppose $\lim_{x \to x_0} g(x) = \ell$. Let $x = x(h) := x_0 + h$. Then we can think of $g(x_0 + h)$ as $g(x(h))$. Now $x = x(h) \to x_0$ as $h \to 0$, but $x \neq x_0$ if $h \neq 0$. Thus $\lim_{h \to 0} g(x_0 + h) = \lim_{h \to 0} g(x(h)) = \lim_{x \to x_0} g(x) = \ell$ by Theorem 2.8.

Conversely, suppose $\lim_{h \to 0} g(x_0 + h) = \ell$. Let $h = h(x) := x - x_0$. Then we can think of $g(x)$ as $g(x_0 + h(x))$, a composition of the functions $g(x_0 + \cdot)$ and $h(\cdot)$. We have $h \to 0$ as $x \to x_0$ and $h \neq 0$ for $x \neq x_0$. Thus $\lim_{x \to x_0} g(x) = \lim_{x \to x_0} g(x_0 + h(x)) = \lim_{h \to 0} g(x_0 + h) = \ell$ by Theorem 2.8.

Example 2.10. Theorem 2.8 can be used to investigate limits at ∞ of $g(x)$ by considering limits at 0 of $g(1/x)$. Write $y = y(x) := \frac{1}{x}$. Then y is defined for any sufficiently large x, $y \neq 0$ and $y \to 0$ as $x \to \infty$. So e.g., $\lim_{x \to \infty} \sin \frac{1}{x} = \lim_{x \to \infty} \sin(y(x)) = \lim_{y \to 0} \sin y = 0$. (Using standard properties of sin. In fact we can use $\lim_{y \to 0^+}$ here as we also have $y > 0$ for all large enough x.)

Example 2.11 (Real powers)**.** For $\alpha > 0$ real we have $x^\alpha \to \infty$ as $x \to \infty$. For $\alpha < 0$ we have $x^\alpha \to 0$ as $x \to \infty$.

We assume standard limits of exp and log (Proposition 5.4 below) and recall that for real α and $x > 0$ we *define* $x^\alpha := \exp(\alpha \log x)$.

Now $\log x \to \infty$ as $x \to \infty$ (Proposition 5.4), and hence for $\alpha > 0$, $\alpha \log x \to \infty$ (Extended AOL). Also $\exp y \to \infty$ as $y \to \infty$ (Proposition 5.4), so (substituting $y = \alpha \log x$) $x^\alpha = \exp(\alpha \log x) \to \infty$ (Theorem 2.8).

For $\alpha < 0$, $\alpha \log x \to -\infty$ (Extended AOL). Now $\exp y \to 0$ as $y \to -\infty$ (Proposition 5.4), so (substituting $y = \alpha \log x$) $x^\alpha = \exp(\alpha \log x) \to 0$ as $x \to -\infty$ (Theorem 2.8).

Example 2.12 (Exponentials beat powers)**.** Let $\alpha \in \mathbb{R}$ and $\beta > 0$ be constants. Then $\lim_{x \to \infty} x^\alpha e^{-\beta x} = 0$.

Proof. We may restrict attention to $x > 0$. Then, by definition of exp,

$$0 \le x^\alpha e^{-\beta x} = \frac{x^\alpha}{1 + \beta x + \cdots + (\beta x)^n/n! + \cdots} \le n!\beta^{-n}x^{\alpha-n}.$$

for any fixed n. Fix a value of $n > \alpha$. Then $0 \le n!\beta^{-n}x^{\alpha-n} \to 0$ as $x \to \infty$ by Example 2.11 and AOL. The result now follows by sandwiching. □

Remark. Working with the power series for e^x when $x > 0$, which has all terms positive, is preferable to working with it when $x < 0$, as then we have terms of alternating sign. Inequalities interact badly with expressions with mixed signs.

Example 2.13 (Powers beat logarithms)**.** For $\alpha > 0$

$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = 0 \qquad \text{and} \qquad \lim_{x\to 0^+} x^\alpha \log x = 0.$$

Proof. Write $y := \log x$ so that $\frac{\log x}{x^\alpha} = (\log x)e^{-\alpha \log x} = ye^{-\alpha y}$. Now $y = \log x \to \infty$ as $x \to \infty$ so by Theorem 2.8 and Example 2.12

$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = \lim_{y\to\infty} ye^{-\alpha y} = 0.$$

For the second statement write $y := -\log x$ so that $x^\alpha \log x = -ye^{-\alpha y}$. Now $y = -\log x \to \infty$ as $x \to 0^+$ so again so by Theorem 2.8 and Example 2.12

$$\lim_{x\to 0^+} x^\alpha \log x = -\lim_{y\to\infty} ye^{-\alpha y} = 0.$$ □

Finally we finish this section with another useful result on limits.

Proposition 2.14. *Suppose $f\colon E \to \mathbb{R}$ and let $p \in \mathbb{R}$, $\ell \in \mathbb{R} \cup \{\pm\infty\}$.*

(a) *If p is a limit point of both E_1 and E_2 where $E = E_1 \cup E_2$ then $\lim_{x\to p} f(x) = \ell$ if and only if both $\lim_{x\to p,\, x\in E_1} f(x) = \ell$ and $\lim_{x\to p,\, x\in E_2} f(x) = \ell$.*

(b) *If p is a limit point of $E_1 \subseteq E$ but not of $E \setminus E_1$ (so $E \cap (p-\delta, p+\delta) \subseteq E_1$ for some $\delta > 0$), then $\lim_{x\to p} f(x) = \ell$ if and only if $\lim_{x\to p,\, x\in E_1} f(x) = \ell$.*

In particular if p is a limit point of $E_1 \subseteq E$ then $\lim_{x\to p} f(x) = \ell$ always implies $\lim_{x\to p,\, x\in E_1} f(x) = \ell$. Similar statements hold when $p = \pm\infty$.

Proof. Exercise. □

Note that this implies Proposition 1.15 where we take $E_1 = E \cap (-\infty, p]$ and $E_2 = E \cap [p, \infty)$. Example 1.8 (and 1.12) also follows with $E_1 = \mathbb{Q}$, $E_2 = \mathbb{R} \setminus \mathbb{Q}$. See also *Analysis I*, problem sheet 4, question 2(a), for a special case of the sequence version of this result (with E_1 the set of even integers and E_2 the set of odd integers).

3 Continuity

We all have a good informal idea of what it means to say that a function has a continuous graph: we can draw it without lifting the pen from the paper. But we want now to use our precise definition of '$f(x) \to \ell$ as $x \to p$' to discuss the idea of continuity. We continue the ε-δ theme of the previous section.

Again let us consider $E \subseteq \mathbb{R}$ and $f\colon E \to \mathbb{R}$. In the definition of $\lim_{x\to p} f(x)$ in Section 1, the point p need not belong to the domain E of f. Indeed, even when $p \in E$ and $f(p)$ was defined, we steadfastly refused to acknowledge this when considering the limiting behaviour of $f(x)$ as x *approaches* p. Now we change our focus and consider the scenario in which $f(p)$ *is* defined and ask whether $\lim_{x\to p} f(x) = f(p)$.

Definition. Let $f\colon E \to \mathbb{R}$, where $E \subseteq \mathbb{R}$ and $p \in E$. We say f **is continuous at** p if

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (|x-p| < \delta \Longrightarrow |f(x) - f(p)| < \varepsilon), \tag{3}$$

otherwise we say f is **discontinuous**, or **has a discontinuity**, **at** p. We say f is **continuous on** E if f is continuous at every point $p \in E$.

Note that the 'limit' is now $f(p)$ and we do not exclude $x = p$ in (3): to do so would be neither necessary nor appropriate. We also do not require p to be a limit point of E.

Proposition 3.1 (Continuity in terms of limits)**.** *Let $f\colon E \to \mathbb{R}$, where $E \subseteq \mathbb{R}$.*

(a) *f is continuous at any isolated point*[20] *of E.*

(b) *If $p \in E$ is a limit point of E, then f is continuous at p if and only if*

$$\lim_{x\to p} f(x) \quad \textit{exists and} \quad \lim_{x\to p} f(x) = f(p).$$

Proof. (a) is immediate, since we may choose $\delta > 0$ such that $\{x \in E : 0 < |x-p| < \delta\} = \emptyset$. For such δ, we have $x \in E$ and $|x - p| < \delta$ only if $x = p$ and then $|f(x) - f(p)| < \varepsilon$, trivially.

(b): It is clear that if the continuity condition holds then the limit one does too. In the other direction, the limit condition, provided the limit is $f(p)$, gives all that we need for continuity; the inequality $|f(x) - f(p)| < \varepsilon$ holds for $0 < |x-p| < \delta$ and also trivially for $x = p$. □

Example 3.2 (Continuity of x and $|x|$)**.** Let $f(x) := x$ and $g(x) := |x|$. For f we can set $\delta := \varepsilon$ and then clearly $|x - p| < \delta$ implies $|f(x) - f(p)| = |x - p| < \varepsilon$. For g note that the reverse triangle inequality gives

$$|g(x) - g(p)| = ||x| - |p|| \le |x - p|.$$

Hence we can again take $\delta := \varepsilon$ in the ε-δ definition of continuity.

[20] Recall that an isolated point of E is a point $p \in E$ that is not a limit point of E.

Example 3.3. Let $c \in \mathbb{R}$. Consider f defined on $\mathbb{R}$ by

$$f(x) := \begin{cases} c, & \text{if } x = 0; \\ 1, & \text{otherwise.} \end{cases}$$

Then $\lim_{x\to 0} f(x) = 1$. Hence f is continuous at 0 if and only if $c = 1$. (Compare with Example 1.7.)

On the other hand, f is continuous at every point $p \neq 0$, irrespective of the value of c.

Example 3.4. Let $\alpha > 0$. The function $f(x) = |x|^\alpha \sin\frac{1}{x}$ is not defined at $x = 0$ so it makes no sense to ask if it is continuous there. In such circumstances we modify f in some suitable way. So we look at

$$g(x) := \begin{cases} |x|^\alpha \sin\frac{1}{x}, & \text{if } x \neq 0; \\ 0, & \text{if } x = 0. \end{cases}$$

Then 0 is a limit point of the domain, and we calculated before that $\lim_{x\to 0} g(x) = 0 = g(0)$, so g is continuous at 0.

The following theorem is useful in showing a function is discontinuous by considering suitable sequences of values. It follows immediately from Proposition 3.1 and the proof of Proposition 1.11. Note that we now don't need to assume $p_n \neq p$.

Theorem 3.5 (Continuity via sequences)**.** *Let $f\colon E \to \mathbb{R}$ where $E \subseteq \mathbb{R}$ and $p \in E$. Then f is continuous at p if and only if for every sequence (p_n) with $p_n \in E$ and $p_n \to p$ we have that $f(p_n) \to f(p)$ as $n \to \infty$.*

Example 3.6. Let $f(x) = 1$ when x is rational and $f(x) = 0$ when x is irrational. Since any rational p has a sequence of irrationals $p_n \to p$ we have $f(p_n) = 0 \not\to f(p) = 1$. Since any irrational p has a sequence of rationals $p_n \to p$ we have $f(p_n) = 1 \not\to f(p) = 0$. Thus f is not continuous at any point.

We can use our characterisation of continuity at limit points in terms of $\lim_{x\to p} f(x)$, together with AOL to prove that the class of functions continuous at p is closed under all the usual algebraic operations.

Theorem 3.7 (Algebra of continuous functions)**.** *Let $E \subseteq \mathbb{R}$, $p \in E$, and suppose $f, g\colon E \to \mathbb{R}$ are both continuous at p. Then the following functions are continuous at p: $|f(x)|$, $f(x) \pm g(x)$, $f(x)g(x)$, $f(x)/g(x)$ (provided $g(p) \neq 0$), and any constant function $h(x) := c$.*

Proof. This follows directly from the corresponding AOL results and Proposition 3.1. □

Example 3.8 (Polynomials and rational functions)**.** Let $f\colon \mathbb{R} \to \mathbb{R}$ be a polynomial. Then f is continuous at every point of $\mathbb{R}$. Further, consider the rational function $f(x) = \frac{r(x)}{q(x)}$, where $r, q\colon \mathbb{R} \to \mathbb{R}$ are polynomials. Then f is continuous at p provided $q(p) \neq 0$.

Proof. Example 3.2 shows that $f(x) = x$ is continuous at every point. Then Theorem 3.7 and induction on degree gives that every polynomial is continuous. Theorem 3.7 then also implies rational functions are continuous where the denominator is non-zero. □

One of the key properties of continuous functions is that they 'commute with limits'.

Theorem 3.9 (Continuous functions commute with limits)**.** *Let $f\colon E \to \mathbb{R}$ and $g\colon E' \to \mathbb{R}$ be functions with $f(E) \subseteq E'$. Suppose p is a limit point of E, or $p = \pm\infty$ and E is unbounded above/below. Suppose also that $\lim_{x\to p} f(x) = \ell \in E'$ and g is continuous at ℓ. Then*

$$\lim_{x\to p} g(f(x)) \quad \textit{exists and equals} \quad g\big(\lim_{x\to p} f(x)\big) = g(\ell).$$

Proof. Since g is continuous at ℓ, for any $\varepsilon > 0$ there is an $\eta > 0$ such that

$$\forall y \in E'\colon (|y - \ell| < \eta \implies |g(y) - g(\ell)| < \varepsilon).$$

So as $f(E) \subseteq E'$

$$\forall x \in E\colon (|f(x) - \ell| < \eta \implies |g(f(x)) - g(\ell)| < \varepsilon).$$

But $f(x) \to \ell$ as $x \to p$ so, as $\eta > 0$,

$$\exists \delta > 0\colon \forall x \in E\colon (0 < |x - p| < \delta \implies |f(x) - \ell| < \eta).$$

Combining these assertions

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (0 < |x - p| < \delta \implies |g(f(x)) - g(\ell)| < \varepsilon).$$

Hence $g(f(x)) \to g(\ell)$ as $x \to p$. The cases when $p = \pm\infty$ are similar. □

Corollary 3.10 (Composition of continuous functions)**.** *Let $f\colon E \to \mathbb{R}$ and $g\colon E' \to \mathbb{R}$ with $f(E) \subseteq E'$. If $f(x)$ is continuous at $p \in E$ and $g(x)$ is continuous at $f(p)$, then $g(f(x))$ is continuous at p.*

Proof.[21] Combine Proposition 3.1 with Theorem 3.9: if p is isolated then there is nothing to prove, and if p is a limit point of E then $\lim_{x\to p} g(f(x)) = g(\lim_{x\to p} f(x))$ by continuity of g and $\lim_{x\to p} f(x) = f(p)$ by continuity of f. Thus $\lim_{x\to p} g(f(x)) = g(f(p))$ and so $g(f(x))$ is continuous at p. □

Recall from *Analysis I* that certain functions from $\mathbb{R} \to \mathbb{R}$ — $\exp x$, $\sin x$, $\cos x$, $\sinh x$ and $\cosh x$ etc. — can be defined by power series, each of which has infinite radius of convergence. You were told in *Analysis I* that a power series defines a function which is continuous at each point within its interval of convergence. Later on (Theorem 7.13) we shall justify this claim.

For now, we shall continue to take this fact on trust. This will allow us to use the algebra of continuous functions and the composition of continuous functions to prove the continuity of a wide variety of functions.

[21] If you are asked to prove this in an exam, don't assume Theorem 3.9, but write out a direct proof using a similar argument to the one in the proof of Theorem 3.9.

Example 3.11. We claim that the function $g\colon \mathbb{R} \to \mathbb{R}$ given by

$$g(x) := \begin{cases} x \sin \frac{1}{x}, & \text{if } x \neq 0; \\ 0, & \text{if } x = 0; \end{cases}$$

is continuous at every point of $\mathbb{R}$.

We have already proved that g is continuous at 0 (special case of Example 3.4).

If $p \neq 0$: $\frac{1}{x}$ is continuous at p as $p \neq 0$ (quotient of continuous functions) and $\sin x$ is continuous at $\frac{1}{p}$ (property of sin). Hence $\sin \frac{1}{x}$ is continuous at p (composition of continuous functions). So $x \sin \frac{1}{x}$ is continuous at p (product of continuous functions).

Left-continuity and right-continuity.

The definitions of one-sided limits lead on to notions of left- and right-continuity. We say that a function f is **left-continuous** (or **continuous from the left**) **at** p if it is continuous as a function restricted to $E \cap (-\infty, p]$, namely

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (p - \delta < x \leq p \Longrightarrow |f(x) - f(p)| < \varepsilon).$$

If p is a left limit point of E then this is equivalent to $f(p^-)$ existing and $f(p^-) = f(p)$. Likewise f is **right-continuous** (or **continuous from the right**) **at** p if

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (p \leq x < p + \delta \Longrightarrow |f(x) - f(p)| < \varepsilon).$$

Proposition 3.12. *Let $f\colon E \to \mathbb{R}$ and let $p \in E$. Then the following are equivalent:*

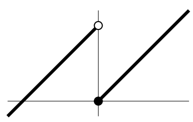
(a) *f is continuous at p;*

(b) *f is both left-continuous at p and right-continuous at p.*

Proof. Exercise. □

Example 3.13. Consider $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x, & \text{if } x \geq 0; \\ x + 1, & \text{if } x < 0. \end{cases}$$

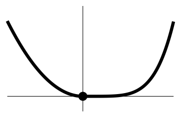


Then $f(0^+) = 0$, $f(0^-) = 1$ and $f(0) = 1$. So $\lim_{x \to 0} f(x)$ does not exist and f fails to be continuous at 0. It is right-continuous but not left-continuous at 0.

Remark. It should be clear that continuity of $f\colon E \to \mathbb{R}$ at a point p depends only on f restricted to a small region about p, say $E \cap (p - \delta, p + \delta)$. However, while continuity of f at p implies that f restricted to say $E_1 = E \cap [p, p + \delta)$ is continuous, continuity of this restricted function is not enough to imply continuity of the original f — in this case it only implies right-continuity. To get the reverse implication needs p not to be limit point of $E \setminus E_1$ (see Proposition 2.14), or equivalently $E_1 \supseteq E \cap (p - \delta, p + \delta)$ for some $\delta > 0$.

The following example shows how we can 'join' two continuous functions if their limits match up at the join.

Example 3.14. Consider $f\colon \mathbb{R} \to \mathbb{R}$ given by



$$f(x) = \begin{cases} 2x^5, & \text{if } x \geq 0; \\ 5x^2, & \text{if } x < 0. \end{cases}$$

Then $\lim_{x\to 0^+} f(x)$ and $\lim_{x\to 0^-} f(x)$ both exist and equal $f(0)$. Hence f is continuous at 0. In addition f is continuous at each point $p \in (0,\infty)$ and at each point of $p \in (-\infty, 0)$ as f is given by a polynomial in a small region around p. Therefore f is continuous on $\mathbb{R}$.

Continuity is often helpful in evaluating limits as the following example shows.

Example 3.15. $\lim_{x\to\infty} x^{1/x} = 1$.

Proof. Let $x > 0$. By definition, $x^{1/x} = e^{x^{-1}\log x}$. By Example 2.13, $x^{-1}\log x \to 0$ as $x \to \infty$. Since exp is continuous at 0, Theorem 3.9 gives

$$x^{1/x} = e^{x^{-1}\log x} \to e^0 = 1 \quad \text{as } x \to \infty.$$ □

Generalisations, continuity of functions of several variables

The definition and basic properties of continuous functions extend immediately to complex and even vector-valued functions (or functions on $\mathbb{C}$ or functions of several variables) with essentially no changes in the proofs. One useful result (which is analogous to a result on complex sequences from *Analysis I*) is the following.

Proposition 3.16. *A function $f\colon E \to \mathbb{C}$ is continuous iff both* $\mathrm{re}(f)$ *and* $\mathrm{im}(f)$ *are continuous.*

Proof. Exercise. □

Indeed, functions *to* $\mathbb{R}^n$ or $\mathbb{C}^n$ are continuous iff each coordinate is given by a continuous function. More interesting is when we consider functions *of* vectors, i.e., functions of several variables. Suppose $f\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is a function of two variables. The way we have defined continuity is that we require

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0),$$

or, to write it out more fully, for all $\varepsilon > 0$ there is a δ such that

$$|(x,y)-(x_0,y_0)| < \delta \implies |f(x,y) - f(x_0,y_0)| < \varepsilon,$$

where $|(x,y)-(x_0,y_0)|$ is the Euclidean distance from (x,y) to (x_0,y_0) in the plane.

Example 3.17. Define $f\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ by

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } (x,y)\neq(0,0); \\ 0, & \text{if } (x,y)=(0,0). \end{cases}$$

Consider $\lim_{(x,y)\to(0,0)} f(x,y)$. It helps to use polar coordinates $(x,y)=(r\cos\theta, r\sin\theta)$ here as the condition $|(x,y)-(0,0)|<\delta$ is just the condition $r<\delta$. We have $f(x,y) = \frac{1}{r^2}(r\cos\theta\cdot r\sin\theta) = \cos\theta\sin\theta$. If $\theta=\frac{\pi}{4}$, so $x=y$, then $f(x,y)=\frac{1}{2}$, while if $\theta=0$, so $y=0$, then $f(x,y)=0$. As we can find such points (x,y) with arbitrarily small r, $f(x,y)$ does not tend to a limit as $(x,y)\to(0,0)$.

Note however that for all $x\neq 0$, $\lim_{y\to 0} f(x,y) = f(x,0) = 0$ as $f(x,y)$ is continuous (rational function) of the variable y if we fix $x\neq 0$. Thus $\lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0 = \lim_{y\to 0}\lim_{x\to 0} f(x,y)$. Hence existence of iterated limits is *not* enough to imply a multi-variable limit.

There are even examples of functions which are continuous along any line $\theta=$ constant through the origin, but are not continuous at $(0,0)$. For example $f(x,y)=\frac{xy^2}{x^2+y^4}$ for $(x,y)\neq(0,0)$, $f(0,0)=0$.

Example 3.18. Now consider

$$g(x,y) := \begin{cases} \frac{x^2y}{x^2+y^2}, & \text{if } (x,y)\neq(0,0); \\ 0, & \text{if } (x,y)=(0,0). \end{cases}$$

In this case, using polar coordinates, $|g(x,y)| = |r\cos^2\theta\sin\theta|\leq r$. Hence, taking $\delta:=\varepsilon$, $|(x,y)-(0,0)|<\delta$ implies $r<\delta$ which implies $|g(x,y)-0|<\varepsilon$, so $\lim_{(x,y)\to(0,0)} g(x,y) = 0 = g(0,0)$ and g is continuous at $(0,0)$.

In the above examples we have continuity for all $(x,y)\neq(0,0)$: it is easy to see the functions $f(x,y):=x$ and $f(x,y):=y$ are continuous, so by algebra of continuous functions, (the suitable generalisation of) Theorem 3.7, any rational function $\frac{p(x,y)}{q(x,y)}$ is continuous at points where $q(x,y)\neq 0$.

4 The Boundedness Theorem and the IVT

Let $f\colon E\to\mathbb{R}$. We say that f is **bounded on** E if the image $f(E)=\{f(x): x\in E\}$ is bounded, i.e., if

$$\exists M>0\colon \forall x\in E\colon |f(x)|\leq M,$$

and similarly for bounded above/below.

When the set $f(E)$ is bounded above (and $E\neq\emptyset$), the Completeness Axiom tells us that

$$\sup f := \sup\{f(x): x\in E\}$$

exists. When $\sup f \in f(E)$ we say that f **attains its sup(remum)**. Corresponding definitions apply to real-valued functions which are bounded below.

While the notion of boundedness is also available for a complex valued function f, the notions of $\sup f$ and $\inf f$ make sense only when f is *real-valued.*

Here is the first Big Theorem of the course.

Theorem 4.1 (Boundedness Theorem)**.** *Suppose $a < b$ and $f\colon [a,b] \to \mathbb{R}$ is a continuous function on the closed bounded interval $[a,b]$. Then*

(a) *f is bounded.*

(b) *f attains its* sup *and its* inf*. That is, there exist points ξ_1 and ξ_2 in $[a,b]$ such that*

$$f(\xi_1) = \sup_{x\in[a,b]} f(x) \qquad \text{and} \qquad f(\xi_2) = \inf_{x\in[a,b]} f(x).$$

Note that in general ξ_1 and ξ_2 will not be unique.

Proof. (a): Argue by contradiction. Suppose f were unbounded. Then for any $n \in \mathbb{N}$, there exists $x_n \in [a,b]$ such that $|f(x_n)| > n$. Since (x_n) is bounded ($x_n \in [a,b]$), by the Bolzano–Weierstrass Theorem, there exists a subsequence (x_{s_n}) converging to p, say. As $[a,b]$ is closed and $x_{s_n} \in [a,b]$ we must have $p \in [a,b]$. Now f is continuous at p and hence

$$\lim_{n\to\infty} f(x_{s_n}) = f(p),$$

so in particular the sequence $(f(x_{s_n}))$ is convergent, and hence bounded. But $|f(x_{s_n})| > s_n \geq n$, so $(f(x_{s_n}))$ is unbounded, a contradiction. Therefore f must be bounded.

(b): Let $M = \sup_{x\in[a,b]} f(x)$. Then by the approximation property of the supremum, for all $n \geq 1$ there exists an $x_n \in [a,b]$ with $M - \frac{1}{n} < f(x_n) \leq M$. Since (x_n) is bounded, by the Bolzano–Weierstrass Theorem, there exists a subsequence (x_{s_n}) converging to p, say. Then $p \in [a,b]$ as $[a,b]$ is closed. Now f is continuous at p and hence

$$\lim_{n\to\infty} f(x_{s_n}) = f(p),$$

But $M - \frac{1}{s_n} < f(x_{s_n}) \leq M$, so by sandwiching $f(p) = \lim_{n\to\infty} f(x_{s_n}) = M$.

A similar argument deals with the infimum, or we can apply what we have done to $-f$ and get the result at once since for any bounded non-empty subset S of $\mathbb{R}$,

$$\inf\{s : s \in S\} = -\sup\{-s : s \in S\}. \qquad \square$$

Example 4.2. Let $E = (0,1]$ and $f\colon E \to \mathbb{R}$ be given by $f(x) = \frac{1}{x}$. Then f is bounded below and attains its inf: $\inf f = f(1)$. On the other hand f is not bounded above: $f(x) \to \infty$ as $x \to 0$. Hence the requirement that E is *closed* in Theorem 4.1 is necessary.

Example 4.3. Let $E = \mathbb{R}$ and let $f(x) = e^x$. Then $\inf f = 0$, but is not attained, and f is not bounded above as $f(x) \to \infty$ as $x \to \infty$. Hence the requirement that E is *bounded* in Theorem 4.1 is necessary.

Example 4.4. Let $E = [0,1]$ and $f(x) = \frac{1}{x}$ for $0 < x < 1$ and $f(0) = f(1) = 2$. Then f is unbounded above and $\inf f = 1$ is not attained. Hence the requirement that f is *continuous* in Theorem 4.1 is necessary.

Remark. On the other hand, looking at the proof of the Boundedness Theorem, one sees that all we needed about the domain of f was that it was closed and bounded — it did *not* need to be an *interval*. In fact it did not even need to be real. Any closed and bounded subset of either $\mathbb{R}$ or $\mathbb{C}$ would do[22]. Such a subset is called **compact**, a concept that will be of great importance in later courses.

Example 4.5. Assume f is a continuous complex-valued function defined on $[a, b]$. Then $|f|$ is continuous and real-valued and the Boundedness Theorem applies to $|f|$. Hence f is bounded. Part (b) of the theorem involves order notions: we can no longer define $\sup f$ and $\inf f$ when f is complex-valued.

So far we have concentrated on extreme values, the supremum and the infimum of a continuous real-valued function on a closed bounded interval. What can we say about possible values between these? Here is the second of our Big Theorems.

Theorem 4.6 (Intermediate Value Theorem (IVT))**.** *Assume $a < b$, $f \colon [a, b] \to \mathbb{R}$ is continuous, and let c be a real number between $f(a)$ and $f(b)$. Then there is at least one point $\xi \in [a, b]$ such that $f(\xi) = c$.*

Note that the restriction that f be real-valued is essential. Also, ξ need not be unique.

Proof. (Divide and Conquer.) By replacing f with $-f$ if necessary, we may assume $f(a) \leq c \leq f(b)$. We shall inductively define a nested sequence of intervals $[a_n, b_n]$, $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$, with $f(a_n) \leq c \leq f(b_n)$ and $b_n - a_n \to 0$.

We start with $[a_0, b_0] = [a, b]$. Now, having defined a_n and b_n, let $m_n = \frac{1}{2}(a_n + b_n)$ be the midpoint of the interval $[a_n, b_n]$. If $f(m_n) \leq c$, let $[a_{n+1}, b_{n+1}] = [m_n, b_n]$; otherwise let $[a_{n+1}, b_{n+1}] = [a_n, m_n]$. Then in either case we have $f(a_{n+1}) \leq c \leq f(b_{n+1})$. Also $b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$, so by induction $b_n - a_n = \frac{1}{2^n}(b - a) \to 0$.

Now (a_n) is clearly increasing and bounded above (by b), so tends to a limit $\xi \in [a, b]$. Similarly (b_n) is clearly decreasing and bounded below (by a), so tends to a limit $\xi' \in [a, b]$. But $b_n - a_n \to 0$, so by AOL we have $\xi = \xi'$. Now by continuity of f and preservation of weak inequalities by limits we have

$$f(\xi) = f\big(\lim_{n\to\infty} a_n\big) = \lim_{n\to\infty} f(a_n) \leq c.$$

Similarly

$$f(\xi) = f\big(\lim_{n\to\infty} b_n\big) = \lim_{n\to\infty} f(b_n) \geq c.$$

Thus $f(\xi) = c$. □

Note that this proof gives an effective algorithm (known as the **bisection method**) for homing in on a root of any continuous equation as $\xi \in [a_n, b_n]$ for all n and $b_n - a_n \to 0$.

[22] Or a closed and bounded subset of $\mathbb{R}^n$ or $\mathbb{C}^n$ — Bolzano–Weierstrass works in these cases too.

Proof. (Alternative inf/sup proof.) Again, by considering $-f$ if necessary we may assume that $f(a) \le c \le f(b)$. Define

$$S := \{x \in [a, b] : f(x) \le c\}.$$

Then $a \in S$ so $S \neq \emptyset$ and S is bounded above by b. So, by the Completeness Axiom, $\xi := \sup S$ exists.[23] Since $a \in S$ we have $\xi = \sup S \ge a$ and since b is an upper bound for S we have $\xi = \sup S \le b$. Therefore $\xi \in [a, b]$.

By the approximation property of sup there exists $x_n \in S$ with $\xi - \frac{1}{n} < x_n \le \xi$. Then $x_n \to \xi$ so continuity of f together with preservation of weak inequalities gives

$$f(\xi) = \lim_{n\to\infty} f(x_n) \le c.$$

Assume $\xi < b$ and pick $y_n \to \xi$ with $\xi < y_n < b$. As $y_n > \xi$ we have $y_n \notin S$ and so $f(y_n) > c$. As $y_n \to \xi$, continuity of f and preservation of weak inequalities gives

$$f(\xi) = \lim_{n\to\infty} f(y_n) \ge c.$$

On the other hand, if $\xi = b$ then clearly $f(\xi) = f(b) \ge c$. Hence $f(\xi) = c$. □

Example 4.7. There exists a unique positive number ξ such that $\xi^2 = 2$.

To prove this we consider $f\colon [1, 2] \to \mathbb{R}$ defined by $f(x) = x^2$. Note that $f(1) = 1 < 2 < 4 = f(2)$ and also, as f is a polynomial, it is continuous. Thus, by the IVT, there exists $\xi \in [1, 2]$ such that $f(\xi) = 2$, as required. Uniqueness can be proved as in *Analysis I*.

Remark. The proof of existence of $\sqrt{2}$ given in *Analysis I* relied crucially on the Completeness Axiom and on a trichotomy argument, as did our proofs of the IVT.

Corollary 4.8 (Continuous image of an interval)**.** *If I is an interval and $f\colon I \to \mathbb{R}$ is continuous, then the image $f(I) = \{f(x) : x \in I\}$ is also an interval.*

Proof. Pick $x \le y \le z$ with $x, z \in f(I)$, say $x = f(a)$, $z = f(b)$, $a, b \in I$. Then as I is an interval we have $[a, b] \subseteq I$ (or $[b, a] \subseteq I$), so we can consider f as a continuous function on $[a, b]$. By the IVT there exists a ξ between a and b with $f(\xi) = y$, so $y \in f(I)$. Thus $f(I)$ has the interval property. □

Corollary 4.9 (Continuous image of a closed bounded interval)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then $f([a, b]) = [c, d]$ for some $c, d \in \mathbb{R}$.*

Proof. By the Boundedness Theorem, part (a), we can define

$$c := \inf_{x\in[a,b]} f(x) \qquad \text{and} \qquad d := \sup_{x\in[a,b]} f(x).$$

Clearly $f([a, b]) \subseteq [c, d]$.

By the Boundedness Theorem, part (b), there exist $\alpha \in [a, b]$ and $\beta \in [a, b]$ such that $f(\alpha) = c$ and $f(\beta) = d$. Hence $c, d \in f([a, b])$.

But $f([a, b])$ is an interval by Corollary 4.8, so $[c, d] \subseteq f([a, b])$. Hence $f([a, b]) = [c, d]$. □

[23] We can also consider $\inf\{x : f(x) \ge c\}$ and construct a similar proof.

Remark. It is not necessarily the case that c or d is $f(a)$ or $f(b)$. Consider, for example, $\sin x$ on $[0, 2\pi]$.

Remark. In the Part A *Topology* course you will find out more about continuity and how to capture this property more elegantly than with the ε-δ definition. You will also encounter more general definitions of compact sets (in $\mathbb{R}$ these are just closed and bounded sets) and connected sets (in $\mathbb{R}$ these are just intervals). The Boundedness Theorem is a special case of the general result that a continuous image of a compact set is compact. The IVT (or its equivalent reformulation, Corollary 4.8) is a special case of the general result that a continuous image of a connected set is connected.

5 Monotonic functions and the Continuous IFT

Definition. Let $E \subseteq \mathbb{R}$ and $f\colon E \to \mathbb{R}$. We say that f is **increasing** (respectively **decreasing**, **strictly increasing**, **strictly decreasing**) if for all $x, y \in E$ with $x < y$ we have $f(x) \leq f(y)$ (respectively $f(x) \geq f(y)$, $f(x) < f(y)$, $f(x) > f(y)$). A function is called (**strictly**) **monotonic** or (**strictly**) **monotone** on E if it is (strictly) increasing or decreasing on E.

Note that a function which is strictly monotonic is injective: $x \neq y \implies f(x) \neq f(y)$.

Recall (from the *Introduction to University Mathematics* course) that a function $f\colon X \to Y$ has an **inverse** $f^{-1}\colon Y \to X$ if and only if f is bijective, i.e., it is both injective and surjective. If we consider f as a function from X to its image $f(X)$, then it is by definition surjective. Hence any injective function $f\colon X \to Y$ has an inverse $f^{-1}\colon f(X) \to X$ defined on the image $f(X)$ of f.

We are now ready to prove the next Big Theorem of this course. It will tell us that a continuous, strictly monotonic function on an interval has a *continuous* inverse.

Theorem 5.1 (The Continuous Inverse Function Theorem (C-IFT))**.** *Let $f\colon I \to \mathbb{R}$ be a strictly monotonic and continuous function on the interval I. Then*

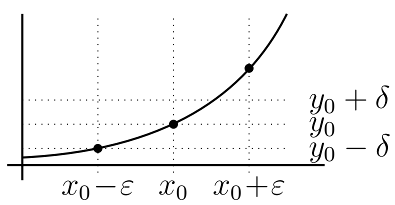
(a) *f is a bijection from I to the interval $f(I)$; and*

(b) *the inverse map to f, $f^{-1}\colon f(I) \to I$, is also strictly monotonic and continuous.*

Proof. Assume without loss of generality that f is strictly increasing. We know from Corollary 4.8 that its image $f(I)$ is an interval. As f is strictly increasing, it is injective and hence gives a bijection from I to $f(I)$. Hence the inverse function $f^{-1}\colon f(I) \to I$, defined by $f^{-1}(y) = x$ when $f(x) = y$, is well-defined. It is also strictly increasing as if $y_1 = f(x_1)$, $y_2 = f(x_2)$, then $x_1 > x_2$ implies $y_1 > y_2$ and $x_1 = x_2$ implies $y_1 = y_2$. Hence if $y_1 < y_2$ we must have $f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2)$ by trichotomy.

It only remains to show that f^{-1} is continuous. Fix $y_0 = f(x_0) \in f(I)$ and $\varepsilon > 0$. Assume first that $x_0 \pm \varepsilon \in I$ and let

$$\delta := \min\{f(x_0) - f(x_0 - \varepsilon), f(x_0 + \varepsilon) - f(x_0)\}.$$



Note that f is strictly increasing, so $\delta > 0$. Also, if $y \in f(I)$ and $|y - y_0| < \delta$ then

$$f(x_0 - \varepsilon) \leq f(x_0) - \delta = y_0 - \delta < y < y_0 + \delta = f(x_0) + \delta \leq f(x_0 + \varepsilon).$$

As f^{-1} is strictly increasing, this implies $x_0 - \varepsilon < f^{-1}(y) < x_0 + \varepsilon$ and hence $|f^{-1}(y) - f^{-1}(y_0)| < \varepsilon$ as required.

If either of $x_0 \pm \varepsilon \notin I$ then one can either reduce ε until it is, in which case the δ found for this smaller ε suffices, or if x_0 is an endpoint of I just remove the undefined term in the minimum defining δ. For example, if $x_0 = \min I$ and $x_0 + \varepsilon \in I$, then set $\delta := f(x_0 + \varepsilon) - f(x_0)$. Now for $|y - y_0| < \delta$ we have as above that $f^{-1}(y) < x_0 + \varepsilon$. But $f^{-1}(y) \geq x_0$ as $f^{-1}(y) \in I$ and $x_0 = \min I$. So again $|f^{-1}(y) - f^{-1}(y_0)| < \varepsilon$. □

Remark. Problem sheet 3 question 4 asks you to prove that, if $f: I \to \mathbb{R}$ is a continuous, injective function with $f(a) < f(b)$ for some $a < b$, then f is strictly increasing on I. So in the statement of the C-IFT it is sufficient that $f: I \to \mathbb{R}$ is continuous and injective.

Example 5.2. For any integer $n \geq 1$ there exists a continuous, strictly increasing nth root function $\sqrt[n]{\ }: [0, \infty) \to \mathbb{R}$ (general n) or $\sqrt[n]{\ }: \mathbb{R} \to \mathbb{R}$ (n odd).

Indeed, the power function $f(x) := x^n$ is continuous and strictly increasing on $[0, \infty)$. Its image is unbounded above and $0^n = 0$, so as the image is an interval it must be $[0, \infty)$. If n is odd then f is strictly increasing on the whole of $\mathbb{R}$ and its image is unbounded in both directions, so must be $\mathbb{R}$.

⚠ **Warning.** If you choose to use the notation f^{-1} for the inverse function of f, when this exists, then you must make very clear what you intend the domains and codomains of f and f^{-1} to be.

Example 5.3. $\sin x$ is strictly increasing on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ with image $[-1, 1]$. Hence we can define $\arcsin: [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$. If we chose a different interval on which to define sin we might either not have sin strictly monotonic (e.g., on $[0, \pi]$) or if it is, we might get a very different definition for arcsin (e.g., if we consider sin on $[\frac{\pi}{2}, \frac{3\pi}{2}]$).

Exponentials and Logarithms

Your likely first encounter with inverse functions would have occurred when you were introduced to the (natural) logarithm function as the inverse of the exponential function. Here we show how to exploit the C-IFT to establish the existence and basic properties of $\log x$ (or $\ln x$ as you may have known it at school). However, before that we need some properties of the exponential function.

We define $\exp(x)$, also written e^x, by

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{4}$$

The most important property of the exponential is that for all $x, y \in \mathbb{C}$,

$$\exp(x + y) = \exp(x)\exp(y). \tag{5}$$

We will not give a proof here, but will prove it later (for real x and y only). If you wish to see a proof of (5) that uses only *Analysis I* material and works for complex x and y, see the supplementary material on exponentials on the website. However, all the other properties of exp that we shall need are fairly easy to deduce from (4) and (5)

Proposition 5.4. $\exp\colon \mathbb{R} \to \mathbb{R}$ *is a continuous, strictly increasing function on* $\mathbb{R}$ *with image* $(0, \infty)$*. As a result, it has a strictly increasing continuous inverse* $\log\colon (0, \infty) \to \mathbb{R}$ *which satisfies*

$$\log(xy) = \log(x) + \log(y)$$

for all $x, y > 0$*. Moreover, we have the limits*

$$\lim_{x\to\infty} \exp x = \infty, \qquad \lim_{x\to-\infty} \exp x = 0, \qquad \lim_{x\to\infty} \log x = \infty, \qquad \lim_{x\to 0^+} \log x = -\infty,$$

and the useful inequality $\exp(x) \geq 1 + x$ *for all* $x \in \mathbb{R}$*.*

Proof. We will prove later that any function defined by a power series is continuous, and indeed differentiable, inside its radius of convergence, but for now we provide a simple direct proof that works just for exp.

Claim 1: $\exp(x) > 0$ and $\exp(x) \geq 1 + x$ for all $x \in \mathbb{R}$.

Proof. For $x \geq 0$ this is clear from the definition $\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots$ as all remaining terms are non-negative. In particular $\exp(x) > 0$ for all $x \geq 0$. Now taking $y = -x$ in (5) gives $\exp(-x) = 1/\exp(x) > 0$ for all $x \geq 0$. Also, for $x \in [0, 1)$, $\exp(x) \leq 1 + x + x^2 + \cdots = \frac{1}{1-x}$, and so $\exp(-x) \geq 1 - x$ for $x \in [0, 1)$. As this also holds trivially for $x \geq 1$, $\exp(-x) \geq 1 - x$ for all $x \geq 0$ and so $\exp(x) \geq 1 + x$ for all $x \in \mathbb{R}$. □

Claim 2: exp is continuous on $\mathbb{R}$.

Proof. For $|x| < 1$ we have $1+x \leq \exp(x) = 1/\exp(-x) \leq \frac{1}{1-x}$ by Claim 1, so sandwiching and AOL gives $\exp(x) \to 1 = \exp(0)$ as $x \to 0$. Now, by Example 2.9, (5) and AOL,

$$\lim_{x\to x_0} \exp(x) = \lim_{h\to 0} \exp(x_0 + h) = \lim_{h\to 0} \exp(x_0)\exp(h) = \exp(x_0) \cdot \lim_{h\to 0} \exp(h) = \exp(x_0). \quad \square$$

Claim 3: exp has image $(0, \infty)$, $\lim_{x\to\infty} \exp(x) = \infty$, $\lim_{x\to-\infty} \exp(x) = 0$.

Proof. By Claim 1, $\exp(x) \geq 1 + x$, so $\exp(x) \to \infty$ as $x \to \infty$ by sandwiching. Hence $\exp(x) = 1/\exp(-x) \to 0$ as $x \to -\infty$ by Extended AOL. As exp is continuous its image must be an interval. The only possibility is $(0, \infty)$ as it is unbounded above, contains points arbitrarily close to 0, but only contains positive numbers. □

Claim 4: exp is strictly increasing.

Proof. If $x < y$ then $\exp(y) = \exp(x)\exp(y - x)$, but $\exp(y - x) \geq 1 + (y - x) > 1$ and $\exp(x) > 0$, so $\exp(y) > \exp(x)$. □

The first part of the proposition now follows from the C-IFT and applying log to the equation

$$\exp(\log(xy)) = xy = \exp(\log x)\exp(\log y) = \exp(\log x + \log y).$$

The limits for log follow from monotonicity: given M set $N := e^M$, then for $x > N$, $\log x > M$. Given M set $\delta := e^{-M}$, then for $0 < x < \delta$, $\log x < -M$. □

Corollary 5.5. *For any $\alpha \in \mathbb{R}$ the function $x \mapsto x^\alpha$ is continuous on $(0, \infty)$.*

Proof. $x^\alpha := \exp(\alpha \log x)$ is a composition of continuous functions. □

More on monotonic functions

The material in this section is non-examinable for *Analysis II*. In it we consider the situation for an arbitrary monotonic function, not assumed to be continuous. We start with a function analogue to the results from *Analysis I* on monotonic sequences.

Theorem 5.6 (One-sided limits of increasing functions)**.** *Let $f : E \to \mathbb{R}$ be increasing. If $p \in E$ is a left limit point of E then the left-hand limit $f(p^-)$ of f at p exists and $f(p^-) = \sup\{f(x) : x < p,\ x \in E\} \le f(p)$. Similarly if $p \in E$ is a right limit point of E then $f(p^+) = \inf\{f(x) : x > p,\ x \in E\} \ge f(p)$.*

Proof. The set $\{f(x) : x < p,\ x \in E\}$ is non-empty as p is a left limit point of E, and is bounded above by $f(p)$ since f is increasing. Therefore by the Completeness Axiom $\ell := \sup\{f(x) : x < p,\ x \in E\}$ exists and $\ell \le f(p)$. We have to show that $f(p^-) = \ell$. Let $\varepsilon > 0$. By the Approximation Property for sup, there exists $x_\varepsilon \in E$, $x_\varepsilon < p$, such that

$$\ell - \varepsilon < f(x_\varepsilon) \le \ell.$$

Choose $\delta := p - x_\varepsilon$. Then $\delta > 0$ as $x_\varepsilon < p$. Also, as f is increasing,

$$p - \delta = x_\varepsilon < x < p \Longrightarrow \ell - \varepsilon < f(x_\varepsilon) \le f(x) \le \ell.$$

By definition $f(p^-) = \ell$ and we are done.

The result for $f(p^+)$ can be obtained by a similar argument, or by applying what we have done to the function $-f(-x)$ on $(-b, -a)$ and juggling with the inequalities. □

Corollary 5.7. *If $f : E \to \mathbb{R}$ is increasing and discontinuous at $p \in E$ then either*

(a) *p is a left limit point of E and $f(p^-) < f(p)$; or*

(b) *p is a right limit point of E and $f(p^+) > f(p)$.*

Proof. If f is discontinuous at p then by Proposition 3.12 it must fail either left-continuity or right-continuity there. Suppose it is not left-continuous, i.e., it is not continuous as a function on $E \cap (-\infty, p]$. Then p cannot be isolated in $E \cap (-\infty, p]$, and thus must be a left limit point. By Theorem 5.6 $f(p^-)$ exists and $f(p^-) \le f(p)$. As f is not left-continuous $f(p^-) \ne f(p)$, so (a) holds. A similar argument implies (b) when f is not right-continuous at p. □

We say that a function $f\colon E \to \mathbb{R}$ has a **jump discontinuity** at p if p is both a left and right limit point of E and $f(p^+)$ and $f(p^-)$ both exist, but $f(p^+) \neq f(p^-)$. For monotonic functions these are the only type of discontinuity we could have, say, in the interior of an interval. On the other hand, monotonic functions can have infinitely many such discontinuities. Indeed, given any countable set $S \subseteq \mathbb{R}$, it is possible to construct an increasing functions on $\mathbb{R}$ whose set of points of discontinuity is exactly S. See problem sheet 4 question 10.

Corollary 5.8. *Any monotonic function has at most countably many points of discontinuity.*

Proof. We may assume without loss of generality that f is increasing. So if f is discontinuous at p then Corollary 5.7, one of the intervals $(f(p^-), f(p))$ or $(f(p), f(p^+))$ must be non-trivial. But distinct 'jump' intervals must be disjoint as if $p < q$ then $f(p^-) \leq f(p) \leq f(p^+) \leq f(x) \leq f(q^-) \leq f(q) \leq f(q^+)$ for any $x \in (p, q)$. Also, each non-trivial interval must contain a rational number, so we can construct an injective map from the set of discontinuities of f to $\mathbb{Q}$, which is countable. □

We stress that the behaviour of monotonic functions is very special. Consider, for example, $f\colon (0,1) \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q}; \\ 0, & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Then the left-hand and right-hand limits $f(p^-)$ and $f(p^+)$ fail to exist for every $p \in (0,1)$. Moreover f is discontinuous at every point of the uncountable set $(0,1)$.

Corollary 5.9. *Suppose I is an interval and $f\colon I \to \mathbb{R}$ is monotonic. Then f is continuous if and only if the image $f(I)$ is an interval.*

Proof. We have already seen that if f is continuous then $f(I)$ is an interval (even for a non-monotonic function). Now suppose without loss of generality that f is increasing and is discontinuous at a point $p \in I$. Then either $f(p^-) < f(p)$ or $f(p) < f(p^+)$. Suppose without loss of generality that $f(p^-) < f(p)$. Then (as $f(p^-)$ is defined), there exists a $q < p$ with $q \in I$ and $f(q) \leq f(p^-) < f(p)$. But any point in $(f(p^-), f(p))$ lies between $f(q)$ and $f(p)$, but is not in the image of f. Hence $f(I)$ is not an interval. □

Remark. We note that this gives an alternative way of showing f^{-1} is continuous in the proof of the Continuous IFT: $f^{-1}(f(I)) = I$ is an interval, so f^{-1} must be continuous.

6 Uniform continuity

This section and the next one are unashamedly technical. In them we look closely at conditions for continuity of functions and at convergence of sequences of functions. The pay-off will be theorems which are important throughout analysis.

Definition. Let $f\colon E \to \mathbb{R}$ or $\mathbb{C}$. Then f is **uniformly continuous on** E if

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall p \in E\colon \forall x \in E\colon (|x - p| < \delta \implies |f(x) - f(p)| < \varepsilon).$$

Compare this with the definition of f being **continuous on** E, i.e., at every $p \in E$:

$$\forall p \in E\colon \forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in E\colon (|x - p| < \delta \implies |f(x) - f(p)| < \varepsilon).$$

The difference between the two statements is in *the order of the quantifiers*. Swapping $\forall$'s doesn't affect the meaning, but swapping the order in which $\forall p$ and $\exists \delta$ occur does change the meaning. Read the expressions from left to right. For uniform continuity on E we need a δ which works universally — that is, for all p in E at the same time. For continuity on E we first choose any p and then find δ that works just for that choice of p: in this case δ is allowed to depend on p.

Of course if $f\colon E \to \mathbb{R}$ is uniformly continuous on E then f is continuous on E. The converse is false, as we now demonstrate.

Example 6.1. Consider $f(x) = \sin \frac{1}{x}$ on $E = (0, 1]$. Certainly f is continuous on E. We shall show that f fails to be uniformly continuous on E.

Take $\varepsilon = 1$. We show that there is no $\delta > 0$ that works in the condition for uniform continuity.

Take sequences $x_n = \frac{1}{2\pi n + \pi/2}$ and $p_n = \frac{1}{2\pi n + 3\pi/2}$. Then $|f(x_n) - f(p_n)| = |1 - (-1)| = 2$, but $|x_n - p_n| \to 0$. So for any $\delta > 0$, there exists p_n and x_n such that $|x_n - p_n| < \delta$ but $|f(x_n) - f(p_n)| \not< 1$. So f is not uniformly continuous.

This example demonstrates an effective strategy for showing a function is not uniformly continuous: find sequences x_n and y_n with $|x_n - y_n| \to 0$ but $|f(x_n) - f(y_n)| \not\to 0$.

Example 6.2. Consider $f(x) = \cos(x^2)$ on $\mathbb{R}$. Take sequences $x_n = \sqrt{2n\pi}$ and $y_n = \sqrt{(2n+1)\pi}$. Then $|x_n - y_n| = |x_n^2 - y_n^2|/|x_n + y_n| = \pi/|x_n + y_n| \to 0$ as $n \to \infty$ but $|f(x_n) - f(y_n)| = |1 - (-1)| = 2 \not\to 0$. Hence f is continuous, but not uniformly continuous on $\mathbb{R}$.

Uniform continuity is a condition that is found to be necessary in certain technical proofs in analysis which involve continuous functions.[24] So the following theorem is important beyond the present course.

Theorem 6.3 (Continuity implies uniform continuity on closed bounded intervals)**.** *If $f\colon [a, b] \to \mathbb{R}$ is continuous, then f is uniformly continuous on $[a, b]$.*

Proof. Suppose for a contradiction that f were not uniformly continuous. By the contrapositive of the uniform continuity condition there would exist $\varepsilon > 0$ such that for

[24] For example, it will be used in *Analysis III* to show one can always integrate a continuous function on a closed bounded interval.

any $\delta > 0$ — which we choose as $\delta = \frac{1}{n}$ for arbitrary n — there exists a pair of points $x_n, y_n \in [a, b]$, such that

$$|x_n - y_n| < \tfrac{1}{n} \qquad \text{but} \qquad |f(x_n) - f(y_n)| \geq \varepsilon.$$

Since each $x_n \in [a, b]$, the sequence (x_n) is bounded, and by the Bolzano–Weierstrass Theorem there exists a subsequence (x_{s_n}) which converges to some p. Now p must be a limit point of $[a, b]$, so $p \in [a, b]$. But

$$y_{s_n} = (y_{s_n} - x_{s_n}) + x_{s_n} \to 0 + p = p$$

by AOL, so by continuity at p we have

$$|f(x_{s_n}) - f(y_{s_n})| \leq |f(x_{s_n}) - f(p)| + |f(y_{s_n}) - f(p)| \to 0 \text{ as } n \to \infty.$$

This gives the required contradiction as we assumed $|f(x_n) - f(y_n)| \geq \varepsilon$ for all n. □

Remark. We note that uniform continuity, unlike continuity, is a *global* property: in the examples above $\sin \frac{1}{x}$ is uniformly continuous on all intervals of the form $[\varepsilon, 1]$, $\varepsilon > 0$, but not on $(0, 1]$, while $\cos(x^2)$ is uniformly continuous on all intervals of the form $[0, N]$, but not on $[0, \infty)$. Also, these examples show that both the conditions of closed and bounded are required in Theorem 6.3. Note also that f itself being bounded did not help at all when it came to uniform continuity in Examples 6.1 and 6.2.

The following is a very special class of functions that *are* uniformly continuous.

Definition. We say that f is **Lipschitz continuous on** E if there exists a constant $K > 0$ such that

$$\forall x, y \in E \colon |f(x) - f(y)| \leq K|x - y|.$$

Assume f satisfies this condition. Given $\varepsilon > 0$ choose $\delta := \frac{\varepsilon}{K}$. Then $\delta > 0$ and for $x, y \in E$ for which $|x - y| < \delta$,

$$|f(x) - f(y)| \leq K|x - y| < \varepsilon.$$

Thus f is uniformly continuous on E.

Later we will see (via the Mean Value Theorem) that 'bounded derivative' is enough to imply Lipschitz, and hence uniform continuity. However, not all Lipschitz functions are differentiable.

Example 6.4. $f(x) = \sqrt{x}$ is Lipschitz continuous on $[1, \infty)$, but not on $[0, 1]$. It is however uniformly continuous on the whole of $[0, \infty)$

To obtain the Lipschitz condition on $[1, \infty)$ note that, for all $x, y \geq 1$.

$$|\sqrt{x} - \sqrt{y}| = \frac{|x - y|}{\sqrt{x} + \sqrt{y}} \leq \tfrac{1}{2}|x - y|,$$

so $K = \frac{1}{2}$ works. However, $|\sqrt{x} - \sqrt{0}| \le K|x - 0|$ fails to hold when $x < 1/K^2$, so $\sqrt{x}$ is not Lipschitz on $[0, 1]$.

Now $\sqrt{x}$ is continuous on $[0, 1]$ (as it is the inverse of the strictly increasing continuous function $x^2 \colon [0, 1] \to [0, 1]$), so it is uniformly continuous on $[0, 1]$ and (by the above) also on $[1, \infty)$. We now stitch these two together to establish uniform continuity on $[0, 1] \cup [1, \infty)$. However, this takes a bit of care.

We know

$$\forall \varepsilon > 0 \colon \exists \delta_1 > 0 \colon \forall x, y \in [0, 1] \colon (|x - y| < \delta_1 \implies |\sqrt{x} - \sqrt{y}| < \tfrac{1}{2}\varepsilon)$$

and

$$\forall \varepsilon > 0 \colon \exists \delta_2 > 0 \colon \forall x, y \in [1, \infty) \colon (|x - y| < \delta_2 \implies |\sqrt{x} - \sqrt{y}| < \tfrac{1}{2}\varepsilon).$$

Choose $\delta = \min\{\delta_1, \delta_2\} > 0$. Suppose that $|x - y| < \delta$. If $x, y \ge 1$ or $x, y \le 1$ we are done. So suppose (wlog) that $x \in [1, \infty)$ and $y \in [0, 1]$ and $|x - y| < \delta$. Then $|x - 1| < \delta$ and $|1 - y| < \delta$ so that

$$|\sqrt{x} - \sqrt{y}| \le |\sqrt{x} - \sqrt{1}| + |\sqrt{1} - \sqrt{y}| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

Hence $|\sqrt{x} - \sqrt{y}| < \varepsilon$ whenever $x, y \in [0, \infty)$ are such that $|x - y| < \delta$. By definition, $f(x) = \sqrt{x}$ is uniformly continuous on $[0, \infty)$.

Remark. In general, if f is uniformly continuous on intervals I and J and $I \cap J \ne \emptyset$, then f is uniformly continuous on the interval $I \cup J$. However this does *not* apply to the union of infinitely many intervals: f uniformly continuous on $[n, n+1]$ for each n does *not* imply f is uniformly continuous on $[1, \infty)$ as we saw with the $\cos(x^2)$ example.

In the case of an interval that is not closed one can still give a simple condition for uniform continuity. Proving the following is question 3 on problem sheet 4.

Proposition 6.5. *Assume $f \colon (a, b] \to \mathbb{R}$ is continuous. Then f is uniformly continuous if and only if $\lim_{x \to a^+} f(x)$ exists.*

7 Uniform convergence

In analysis one often wants to know how different limiting processes interact with one other. In particular, does a limiting process, such as that involved in continuity, commute with another type of limit? Sadly, however, the answer in general is 'No'. This leads us to try to find sufficient conditions under which the answer will be 'Yes'. In this section we take a first excursion into problems of this kind.

Pointwise convergence

Initially, we want to consider a sequence (f_n) of functions, where $E \subseteq \mathbb{R}$ and $f_n \colon E \to \mathbb{R}$ for $n \in \mathbb{N}$. Observe that, for each *fixed* $x \in E$, the sequence $(f_n(x))$ is a sequence of real numbers, whose behaviour we can analyse by the techniques of *Analysis I*.

We say (f_n) **converges** (**pointwise**) to the function $f\colon E \to \mathbb{R}$ (and write $f = \lim f_n$ or $f_n \to f$ on E) if for each $x \in E$ the sequence $(f_n(x))$ converges to $f(x)$. That is,

$$\forall x \in E\colon \forall \varepsilon > 0\colon \exists N \in \mathbb{N}\colon \forall n > N\colon |f_n(x) - f(x)| < \varepsilon. \qquad \textbf{(pointwise convergence)}$$

Note that here N is allowed to depend on both x and ε.

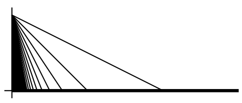
Pointwise convergence is nothing unfamiliar. In saying, for example,

$$e^x = 1 + x + \tfrac{x^2}{2!} + \cdots \qquad \text{on } \mathbb{R}$$

we mean precisely that the partial sums of the series on the right-hand side converge pointwise to the exponential function for each $x \in \mathbb{R}$.

Example 7.1. Consider the sequence of functions (f_n), where $f_n\colon [0,1] \to \mathbb{R}$ is given by

$$f_n(x) := \begin{cases} 1 - nx, & \text{if } 0 \le x < \frac{1}{n}; \\ 0, & \text{if } x \ge \frac{1}{n}. \end{cases}$$



Consider also the function $f\colon [0,1] \to \mathbb{R}$ given by

$$f(x) := \begin{cases} 1, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases}$$

What happens as n increases? Note that for each fixed $x \in [0,1]$ we have $f(x) = \lim_{n\to\infty} f_n(x)$ (separate cases $x \neq 0$ and $x = 0$). Hence (f_n) converges pointwise to f.

Note that although all the f_n are continuous, the pointwise-limit function f is *not* continuous at 0. Spelling this out,

$$\lim_{x\to 0}\lim_{n\to\infty} f_n(x) = \lim_{x\to 0} f(x) = 0 \qquad \text{but} \qquad \lim_{n\to\infty}\lim_{x\to 0} f_n(x) = \lim_{n\to\infty} 1 = 1.$$

The order in which the limits are taken affects the value of the iterated limit.

Moral: in general, iterated limits may squabble. They must be handled with care.

Uniform continuity leads to stronger results than continuity one point at a time. The idea in the definition of uniform continuity was to require a 'universal δ'. There is a parallel with the key definition of this section, which we now give.

Uniform convergence

Definition. Let (f_n) be a sequence of functions $f_n\colon E \to \mathbb{R}$ or $\mathbb{C}$. Then (f_n) **converges uniformly to f on E** if[25]

$$\forall \varepsilon > 0\colon \exists N \in \mathbb{N}\colon \forall n > N\colon \forall x \in E\colon |f_n(x) - f(x)| < \varepsilon. \qquad \textbf{(uniform convergence)}$$

[25] The order of $\forall n$ and $\forall x$ does not matter here, so could be swapped to make the correspondence with the definition of uniform continuity clearer. However this form is slightly more convenient.

If this holds we write $f_n \xrightarrow{u} f$ on E. Note that specifying the set E is an integral part of the definition. The order of the quantifiers matters: the uniform convergence condition demands a universal N which is *independent* of x (although it may still depend on ε).

It is immediate from the definitions that if $f_n \xrightarrow{u} f$ on E then (f_n) converges pointwise to f on E.

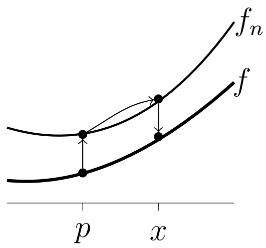
The next theorem gives a reason why uniform convergence is a Good Thing.

Theorem 7.2 (Uniform limits preserve continuity)**.** *Let (f_n) be a sequence of continuous functions on E which converges uniformly to f on E. Then f is continuous on E.*

Proof. To prove continuity of f we first fix some $p \in E$ and $\varepsilon > 0$.

By uniform convergence we can find $N \in \mathbb{N}$ such that

$$n > N \Longrightarrow \forall x \in E\colon |f_n(x) - f(x)| < \tfrac{\varepsilon}{3}.$$



Fix an $n > N$. Then by continuity of f_n at p there exists $\delta > 0$ such that

$$|x - p| < \delta \Longrightarrow |f_n(x) - f_n(p)| < \tfrac{\varepsilon}{3}$$

(δ depending on n — but n is fixed). Hence for $|x - p| < \delta$,

$$|f(x) - f(p)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(p)| + |f_n(p) - f(p)| < \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} = \varepsilon.$$

This suffices to prove our claim. Note that uniformity of convergence is needed to handle the first term simultaneously for every relevant x. □

Remark. The proof actually shows the slightly stronger statement: if $f_n \xrightarrow{u} f$ on E and each f_n is continuous at $p \in E$, then f is continuous at p.

We now convert the uniform convergence condition into a more amenable form.

Proposition 7.3 (Testing for uniform convergence)**.** *Assume $f, f_n\colon E \to \mathbb{R}$ or $\mathbb{C}$. Then the following statements are equivalent:*

(a) *$f_n \xrightarrow{u} f$ on E;*

(b) *for each sufficiently large n the set $\{|f_n(x) - f(x)| : x \in E\}$ is bounded and*

$$s_n := \sup_{x \in E} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty.$$

Proof. Assume (a). Then, given $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for $n > N$ and for all $x \in E$ we have $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$. So the first condition in (b) holds for such n and hence s_n is well defined. Fix n and take the supremum over $x \in E$ to get, for $n > N$,

$$0 \leq s_n = \sup_{x \in E} |f_n(x) - f(x)| \leq \tfrac{\varepsilon}{2} < \varepsilon.$$

Hence $s_n \to 0$.

Conversely, assume (b). Given $\varepsilon > 0$, choose N so that $n > N$ implies $s_n < \varepsilon$. Then, for all $n > N$ and all $x \in E$,

$$|f_n(x) - f(x)| \leq s_n < \varepsilon.$$

Hence $f_n \xrightarrow{u} f$. □

A few comments on working with Proposition 7.3 are in order. First of all, it allows us to reduce testing for uniform convergence of (f_n) on E to three steps:

Step 1: find the pointwise limit.

With $x \in E$ fixed, find $f(x) := \lim_{n\to\infty} f_n(x)$, or show it fails to exist (of course, if the pointwise limit fails to exist for any $x \in E$, then certainly (f_n) does not converge uniformly and we proceed no further). Look out for values of x which need special attention.

Step 2: calculate (or find bounds for) s_n.

Assuming all f_n and f are continuous and E is an interval $[a, b]$ (the most common scenario), the Boundedness Theorem applied to the continuous function $|f_n - f|$ tells us the sup is attained, so we want to know the maximum value of $|f_n - f|$. Frequently $f_n - f$ will be of constant sign so we can get rid of the modulus signs. Then, if the functions f_n and f are differentiable the supremum (or infimum) of $f_n - f$ will be achieved either at a or at b or at some interior point where $f_n'(x) - f'(x) = 0$. It is fine to use school calculus to find maxima and minima by differentiation, when the derivative exists — we'll validate this technique later. See examples below for illustrations.

Step 3: see if s_n tends to 0.

Now (s_n) is a sequence of real numbers. We are back in *Analysis I* territory, and can use standard techniques and standard limits from that course.

Note that in Step 1 we work with fixed x and in Step 2 we work with fixed n (and in Step 3 we don't have x anymore): we never need to consider both x and n varying at the same time.

Example 7.4. Let

$$f_n(x) := \begin{cases} 1 - nx, & \text{if } 0 \leq x < \frac{1}{n}; \\ 0, & \text{if } x \geq \frac{1}{n}. \end{cases}$$

Step 1: Fix x. Suppose first that $x \neq 0$. Then $\exists N \in \mathbb{N}$ such that $0 < \frac{1}{N} < x$ (Archimedean Property). This implies $f_n(x) = 0$ for all $n > N$. Therefore $f_n(x) \to 0$ as $n \to \infty$ whenever $x \neq 0$. If $x = 0$, then $f_n(0) = 1$ and so $f_n(0) \to 1$.

We deduce that the pointwise limit indeed exists and equals f , where

$$f(x) := \begin{cases} 1, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Step 2: Now fix n and calculate s_n.

$$s_n := \sup_{x\in[0,1]} |f_n(x) - f(x)| = \sup_{x\in(0,1/n]} |1 - nx| = 1.$$

Step 3: Trivially, $s_n \to 1 \neq 0$. Hence (f_n) is not uniformly convergent.

Of course the contrapositive of Theorem 7.2 gives an alternative proof that convergence cannot be uniform.

Example 7.5. Let $E = [0, 1)$ and let $f_n(x) = x^n$.

Step 1: For fixed $x \in [0, 1)$, we have $x^n \to 0$ as $n \to \infty$. Hence the pointwise limit is $f = 0$.

Step 2: Trivially, $s_n = \sup_{x\in[0,1)} |x^n| = 1$. Indeed $x \in [0, 1)$ implies $|x^n| \leq 1$, but $x^n \to 1$ as $x \to 1^-$.

Step 3: Hence $s_n \not\to 0$, so convergence is not uniform.

Now consider what happens if, with f_n as before, we work on $[0, b]$, where b is a constant with $0 \leq b < 1$. The pointwise limit is unchanged but now

$$s_n = \sup_{x\in[0,b]} x^n = b^n \to 0 \text{ as } n \to \infty.$$

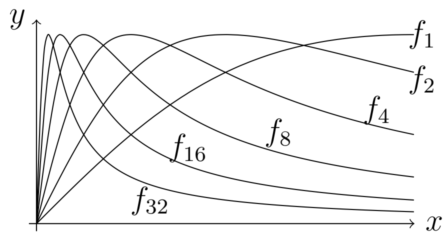
Hence convergence is uniform on $[0, b]$ for each fixed $b < 1$.

This example highlights that uniform convergence, or not, depends on the set E. It makes no sense to say '(f_n) converges uniformly' without specifying the set E on which the functions are considered. Also being uniformly convergent on each $E_n = [0, 1 - \frac{1}{n}]$ does *not* imply uniform convergence on $\bigcup_n E_n = [0, 1)$.

Example 7.6. Let $E = [0, 1]$ and let

$$f_n(x) := \frac{nx}{1 + n^2x^2}.$$

Clearly $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in [0, 1]$. But $f_n(\frac{1}{n}) = \frac{1}{2}$, so that



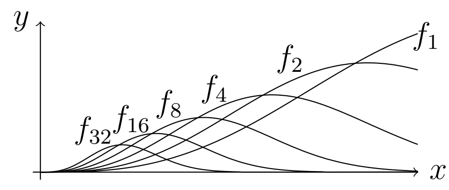
$$\sup_{x\in[0,1]} |f_n(x) - f(x)| \geq \tfrac{1}{2} \not\to 0 \quad \text{as } n \to \infty$$

and so (f_n) converges to 0 but *not uniformly* on $[0, 1]$.

In fact $x = \frac{1}{n}$ is the point at which $f_n(x)$ is maximal, so $s_n = \frac{1}{2}$. One can find this point by setting $f_n' = 0$. Note that we don't need to justify this, it is enough that $x = \frac{1}{n}$ 'breaks' uniform convergence.

Example 7.7. Let $E = [0,1]$ and consider

$$f_n(x) := nx^3 e^{-nx^2}.$$



Step 1: Fix x. For $x = 0$, $f_n(x) = 0$ for all n. For $x > 0$ we have, from the exponential series,

$$0 \le f_n(x) = \frac{nx^3}{1 + nx^2 + \frac{(nx^2)^2}{2!} + \cdots} \le \frac{2}{nx} \to 0 \quad \text{as } n \to \infty.$$

So, by sandwiching, $f_n(x) \to 0$ and this is true for $x = 0$ too, trivially.

Step 2: Fix n and compute $s_n := \sup\{nx^3 e^{-nx^2} : x \in [0,1]\}$. We have

$$\frac{\mathrm{d}}{\mathrm{d}x} nx^3 e^{-nx^2} = 3nx^2 e^{-nx^2} - 2n^2x^4 e^{-nx^2} = nx^2(3 - 2nx^2)e^{-nx^2}$$

and this is zero when $x = 0$ (giving a minimum) and when $2nx^2 = 3$ (giving a maximum). Hence

$$s_n = nx^3 e^{-nx^2}\Big|_{x=\sqrt{3/2n}} = n(3/2n)^{3/2} e^{-3/2} = C/\sqrt{n}$$

where C is a constant independent of n.

If you prefer a proof that does not rely on calculus one can note that for $x > \frac{1}{\sqrt{n}}$, $0 \le f_n(x) \le \frac{2}{nx} \le \frac{2}{\sqrt{n}}$ and for $x \le \frac{1}{\sqrt{n}}$, $0 \le f_n(x) \le nx^3 \le \frac{1}{\sqrt{n}}$, hence $s_n \le \frac{2}{\sqrt{n}}$. [If one needs separate arguments to bound a function in different ranges, it is often easiest to split at a point (here $x = \frac{1}{\sqrt{n}}$) that is close to the maximum.]

Step 3: From Step 2, $s_n \to 0$ as $n \to \infty$. Therefore $f_n \xrightarrow{u} 0$ on $[0,1]$.

Example 7.8 (Partial sums of the geometric series)**.** On $E = (-1,1)$ consider (f_n) given by

$$f_n(x) := 1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}$$

Step 1: Fix x with $|x| < 1$ and let $n \to \infty$. Then $f_n(x) \to f(x) := \frac{1}{1-x}$.

Step 2: Fix n. Here

$$\left\{\left|\tfrac{1-x^{n+1}}{1-x} - \tfrac{1}{1-x}\right| : |x| < 1\right\} = \left\{\tfrac{|x|^{n+1}}{1-x} : |x| < 1\right\}$$

is not bounded above. To see this, consider what happens as $x \to 1^-$. Hence the sequence is not uniformly convergent on $(-1,1)$.

Just as we found for sequences of real numbers, there is a characterisation of uniform convergence which does not depend on knowing the limit function.

Theorem 7.9 (Cauchy Criterion for uniform convergence of sequences)**.** *For $n \in \mathbb{N}$ let $f_n \colon E \to \mathbb{R}$ or $\mathbb{C}$. Then (f_n) converges uniformly on E if and only if*[26]

$$\forall \varepsilon > 0 \colon \exists N \in \mathbb{N} \colon \forall n, m > N \colon \forall x \in E \colon |f_n(x) - f_m(x)| < \varepsilon.$$

[26]Note that this is just the pointwise Cauchy criterion but with the $\forall x$ moved to after the $\exists N$.

Proof. $\Longrightarrow$: Suppose (f_n) converges uniformly on E with limit function f. Then

$$\forall \varepsilon > 0 \colon \exists N \in \mathbb{N} \colon \forall n > N \colon \forall x \in E \colon |f_n(x) - f(x)| < \tfrac{\varepsilon}{2}.$$

So, for all $\varepsilon > 0$ there exists an N such that

$$\forall n, m > N \colon \forall x \in E \colon |f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Hence the uniform Cauchy criterion holds.

$\Longleftarrow$: Suppose the uniform Cauchy criterion holds. Then for each $x \in E$, $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$, so it is convergent. Let us denote its limit by $f(x)$. Now

$$\forall \varepsilon > 0 \colon \exists N \in \mathbb{N} \colon \forall n, m > N \colon \forall x \in E \colon |f_n(x) - f_m(x)| < \tfrac{\varepsilon}{2}.$$

Fix $\varepsilon > 0$, $N \in \mathbb{N}$, $n > N$ and $x \in E$, and let $m \to \infty$ in the above inequality. By AOL and the preservation of weak inequalities[27],

$$|f_n(x) - f(x)| = \lim_{m \to \infty} |f_n(x) - f_m(x)| \leq \tfrac{\varepsilon}{2} < \varepsilon.$$

As this holds for all $n > N$ and all $x \in E$, $f_n \xrightarrow{u} f$ on E. □

An important application of the Cauchy criterion is to series where we often do not know what the limit should be. Indeed, we often use series to *define* a function.

As usual, we handle a series by considering its sequence of partial sums. Accordingly, given a sequence (u_k) of functions defined on a set E we say that the series $\sum u_k$ **converges pointwise** (**uniformly**) **on** E if (f_n) converges pointwise (uniformly) on E, where

$$f_n(x) := u_1(x) + u_2(x) + \cdots + u_n(x) = \sum_{k=1}^{n} u_k(x).$$

Assume each u_k is continuous on E. Then each f_n is also continuous on E. As a corollary of Theorem 7.2 we deduce that if $\sum u_k$ converges uniformly on E then $\sum_{k=1}^{\infty} u_k$ is continuous on E. So we need some way of determining when the convergence is uniform.

Corollary 7.10 (Cauchy Criterion for uniform convergence of series)**.** *Let (u_k) be a sequence of functions on E. Then $\sum u_k$ converges uniformly on E if and only if*

$$\forall \varepsilon > 0 \colon \exists N \in \mathbb{N} \colon \forall n > m > N \colon \forall x \in E \colon |u_{m+1}(x) + \cdots + u_n(x)| < \varepsilon.$$

Proof. Apply Theorem 7.9 to the sequence of partial sums given by $f_n := \sum_{k=1}^{n} u_k$. □

There is a more user-friendly sufficient condition for uniform convergence of a series. It is not a *necessary* condition however.

Theorem 7.11 (Weierstrass' M-test)**.** *Suppose there exist real constants M_k such that*

$$\forall k \colon \forall x \in E \colon |u_k(x)| \leq M_k \qquad \textit{and} \qquad \sum M_k \textit{ converges.}$$

Then the series $\sum u_k(x)$ converges uniformly on E.

Remark. It is critically important in the M-test that M_k is a convergent series of *constants*: M_k must be *independent of* x.

Proof. Apply the Cauchy criterion (Theorem 0.2) to the partial sums of $\sum M_k$:

$$\forall \varepsilon > 0 \colon \exists N \in \mathbb{N} \colon \forall n > m > N \colon \left| \sum_{k=1}^{n} M_k - \sum_{k=1}^{m} M_k \right| = M_{m+1} + \cdots + M_n < \varepsilon.$$

Thus we have for each $x \in E$ and all $n > m > N$,

$$|f_m(x) - f_n(x)| = |u_{m+1}(x) + \cdots + u_n(x)| \leq M_{m+1} + \cdots + M_n < \varepsilon. \tag{6}$$

Hence[28] for each fixed x, $(f_n(x))$ satisfies the Cauchy criterion, and so converges to $f(x)$ say. Thus the series $\sum u_k$ converges *pointwise*.

To check that convergence is uniform, take the limit as $n \to \infty$ in (6) (with x and m fixed) to get that for all $m > N$ and $x \in E$,

$$|f_m(x) - f(x)| \leq \varepsilon.$$

As ε was arbitrary and N did not depend on x, f_m converges to f uniformly as $m \to \infty$. □

Example 7.12. On $E = [0, 1]$ and for $k \geq 1$, let $u_k(x) = \frac{x^p}{1+k^2x^2}$ where p is a constant.

Assume $p \geq 2$. Then, for $x \in [0, 1]$,

$$|u_k(x)| \leq \frac{x^{p-2}}{k^2} \leq M_k := \frac{1}{k^2}. \tag{7}$$

Since $\sum k^{-2}$ converges, $\sum u_k(x)$ converges uniformly on $[0, 1]$ by the M-test.

Now assume $1 < p < 2$. The choice of M_k we used in (7) no longer works. Note that $u_k(x) \geq 0$ so, for fixed k, let's find the maximum value of $u_k(x)$ on $[0, 1]$ by differentiation. We have

$$u_k'(x) = \frac{px^{p-1}(1 + k^2x^2) - 2k^2x^{p+1}}{(1 + k^2x^2)^2}$$

and we see that the maximum of u_k on $[0, 1]$ is achieved at $x_k \in [0, 1]$ where $x_k = \sqrt{p/(2-p)}/k$. We deduce that, for all $x \in [0, 1]$,

$$0 \leq u_k(x) \leq u_k(x_k) \leq M_k := \frac{C}{k^p},$$

where C is a positive constant depending on p but *independent of* x.

[Alternatively: if $x < \frac{1}{k}$, $u_k(x) \leq x^p \leq \frac{1}{k^p}$; while if $x \geq \frac{1}{k}$, $u_k(x) \leq \frac{x^p}{k^2x^2} = \frac{1}{k^p}(\frac{x}{k})^{p-2} \leq \frac{1}{k^p}$.]

The series $\sum \frac{1}{k^p}$ converges for $p > 1$ by the Integral Test. Hence $\sum u_k$ converges uniformly on $[0, 1]$ by the M-test.

Remark. The M-test is useful when it works, but is not infallible. It investigates the maximum of each term separately rather than of the expression arising in the uniform Cauchy criterion, Corollary 7.10. See problem sheet 4 question 8.

[27] Note how $< \frac{\varepsilon}{2}$ changed to $\leq \frac{\varepsilon}{2}$ here.

[28] Actually we are now done by Corollary 7.10, but if you are asked to prove the M-test in an exam you should write out the details as I have done here.

Power series

We now reach another Big Theorem.

Theorem 7.13 (Uniform convergence and continuity of power series)**.** *Let $\sum c_k x^k$ be a real or complex power series with radius of convergence $R \in (0, \infty]$.*

(a) *$\sum c_k x^k$ converges uniformly on $\{x : |x| \le \rho\}$ for any (finite) ρ with $0 < \rho < R$.*

(b) *$f(x) := \sum_{k=0}^{\infty} c_k x^k$ defines a continuous function f on $\{x : |x| < R\}$.*

Proof. (a) Let $M_k = |c_k \rho^k|$. Then as $\rho < R$, $\sum c_k \rho^k$ converges absolutely, and so $\sum M_k$ converges. For $|x| \le \rho$, $|c_k x^k| \le M_k$, so $\sum c_k x^k$ converges uniformly on $\{x : |x| \le \rho\}$ by the M-test.

(b) Fix x_0 with $|x_0| < R$ and choose ρ so that $|x_0| < \rho < R$. By (i), $\sum c_k x^k$ converges uniformly on $\{x : |x| \le \rho\}$ and, as polynomials are continuous, Theorem 7.2 implies that the limit $f(x)$ is continuous on $\{x : |x| \le \rho\}$. Hence f is continuous at x_0. □

Remark. We needed $|x_0| < \rho$ in the proof of (b). If $|x_0| = \rho$ we would only be able to deduce some sort of one-sided continuity of f from continuity on $\{x : |x| \le \rho\}$.

Corollary 7.14. *The following functions, given by power series with infinite radius of convergence, are continuous on $\mathbb{R}$:*

$$\exp x,\ \sin x,\ \cos x,\ \sinh x,\ \cosh x.$$

Functions derived from these via reciprocal and quotient, such as

$$\operatorname{cosec} x,\ \sec x,\ \tan x,\ \cot x$$

are continuous on any set on which the denominator is never zero.

Functions which can be derived from the above functions by application of the Continuous Inverse Function Theorem are themselves continuous. This includes $\log x$ on $(0, \infty)$ and $\arctan x$ on $(-\infty, \infty)$.

Warning. We cannot stress too strongly that Theorem 7.13 is subtle and needs applying with care. Let $\sum c_k x^k$ be a power series with radius of convergence $R > 0$. In general $\sum c_k x^k$ will *not* converge uniformly on $\{x : |x| < R\}$. Indeed, Example 7.8 shows that $\sum x^k$ is not uniformly convergent on $(-1, 1)$. It does however converge uniformly any any interval $[-\rho, \rho]$ with $0 < \rho < 1$, and the limit is continuous on the whole of $(-1, 1)$. Remember that uniform convergence (and uniform continuity) are *global* properties, they depend on the whole of E. Pointwise convergence and continuity are *local* properties — for them to hold on E one just needs to check what happens at or near each $x_0 \in E$.

Example 7.15. Consider the series $\sum_{k=0}^{\infty} x^k \cos(kx^2)$ on $E = [0, 1)$. By the Comparison Test the series converges for each fixed $x \in [0, 1)$.

Indeed, for any η with $0 < \eta < 1$,

$$\forall x \in [0, \eta] \colon |x^k \cos(kx^2)| \le M_k := |\eta|^k \qquad \text{and} \qquad \sum M_k \text{ converges.}$$

By the M-test, the series converges uniformly on $[0,\eta]$.

We don't have a candidate for M_k which would show that the series is *uniformly* convergent on $[0,1)$. Nonetheless we claim that $f(x) := \sum_{k=0}^{\infty} x^k \cos(kx^2)$ defines a function which is continuous on $[0,1)$. To do this, fix p with $0 \le p < 1$ and choose $\eta > 0$ with $p < \eta < 1$. Then the series converges uniformly on $[0,\eta]$. Since each function $x^k \cos(kx^2)$ is continuous on $[0,\eta]$, Theorem 7.2 implies that f is continuous on $[0,\eta]$ and hence is continuous at p.

Example 7.16. Consider the series

$$\sum_{k=0}^{\infty} \frac{k^2 x}{1 + k^4 x^2}.$$

We claim that this converges uniformly on $[\delta,1]$ for each δ with $0 < \delta < 1$. Let $M_k := k^{-2}\delta^{-1}$. Then, on $[\delta,1]$

$$\Big|\frac{k^2 x}{1 + k^4 x^2}\Big| \le \Big|\frac{k^2 x}{k^4 x^2}\Big| \le k^{-2}\delta^{-1} = M_k.$$

Since $\sum M_k$ converges, we do indeed have uniform convergence on each interval $[\delta,1]$. We shall now show that the series is *not* uniformly convergent on the interval $(0,1]$. [Note: failing to find an appropriate M_k is not enough — the M-test is sufficient but not necessary for uniform convergence.]

If the series were uniformly convergent, the uniform Cauchy criterion would show that, for any $\varepsilon > 0$ there exists N such that for all $x \in (0,1]$, and all $n > N$,

$$\Big|\sum_{k=n}^{n} \frac{k^2 x}{1 + k^4 x^2}\Big| = \Big|\frac{n^2 x}{1 + n^4 x^2}\Big| < \varepsilon$$

But for $x = \frac{1}{n^2}$ this would give $\frac{1}{2} < \varepsilon$ for every $\varepsilon > 0$, a contradiction. [More generally: if $\sum u_k(x)$ converges uniformly on E then $u_k(x) \to 0$ uniformly on E.]

But, localising to a point $p \in (0,1]$ and choosing δ such that $0 < \delta < p$, we see that the series defines a function which is continuous on $(0,1]$.

8 Differentiation

In this section we look at differentiation, making use of the machinery of function limits which we have developed. We rediscover all the familiar differentiation rules from school calculus and start to explore examples of functions which are and are not differentiable. Major theorems on differentiable functions come in the next section.

Definition. Let $f\colon E \to \mathbb{R}$, and let $x_0 \in E$ be a limit point of $E \subseteq \mathbb{R}$. We say f **is differentiable at** x_0 if the following limit exists:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

When it exists we denote the limit by $f'(x_0)$ and we call it the **derivative of f at x_0**. We say that f **is differentiable on** E if f is differentiable at every point of E.

Alternative notations. We shall, as convenient, adopt the various different ways of writing derivatives with which you'll be already familiar: for a differentiable function $y = y(x)$:

$$y' \qquad \text{or} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} \qquad \text{or} \qquad \frac{\mathrm{d}}{\mathrm{d}x}y(x).$$

We next present a reformulation of the definition of differentiability as a point. The central idea is to avoid the need for division, which often simplifies the algebra[29]

Proposition 8.1 (Alternative formulation of differentiability)**.** *Let $f\colon E \to \mathbb{R}$ and let x_0 be a limit point of E. Then the derivative $f'(x_0)$ exists and equals ℓ iff one can write*

$$f(x_0 + h) = f(x_0) + \ell h + \varepsilon(h)h \tag{8}$$

with $\varepsilon(h) \to 0$ as $h \to 0$.

Proof. Note that for any $x = x_0 + h \neq x_0$, $f(x_0 + h) = f(x_0) + \ell h + \varepsilon(h)h$ is equivalent to

$$\varepsilon(h) = \frac{f(x_0 + h) - f(x_0)}{(x_0 + h) - x_0} - \ell.$$

Thus the definition of the derivative being equal to ℓ is precisely the condition (after the change of variable $x = x_0 + h$ and AOL) that $\varepsilon(h) \to 0$ as $h \to 0$. $\square$

Example 8.2. It is immediate that f given by $f(x) = x$ is differentiable on $\mathbb{R}$ with $f'(x) = 1$. Indeed, we can take $\ell = 1$, $\varepsilon(h) = 0$ in Proposition 8.1. Slightly more interestingly, $f(x) = x^2$ is differentiable with $f'(x) = 2x$: take $\ell = 2x_0$ and $\varepsilon(h) = h$ in Proposition 8.1.

Another easy consequence is that differentiability implies continuity.

Proposition 8.3 (Differentiability implies continuity)**.** *Let $f\colon E \to \mathbb{R}$ and let x_0 be a limit point of E. If f is differentiable at x_0, then it is continuous at x_0.*

Proof. $\lim_{h\to 0} f(x_0 + h) = f(x_0)$ is immediate from Proposition 8.1 and AOL.
[Alternatively: $\lim_{x\to x_0} f(x) - f(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \cdot \lim_{x\to x_0}(x - x_0) = f'(x_0)\cdot 0$.] $\square$

[29] Also, with minor changes, it allows for differentiation of functions defined on vectors, or multi-variable functions. More on this in the Part A course *Multidimensional Analysis and Geometry*.

Generalisations

Generalisations to functions $\mathbb{C} \to \mathbb{C}$ and $\mathbb{R} \to \mathbb{C}$ are straightforward. We can't extend to functions $\mathbb{C} \to \mathbb{R}$. (Why: firstly f' would have to be in $\mathbb{C}$ anyway since we need to divide by $x - x_0 \in \mathbb{C}$, but for a more fundamental problem wait for the Part A course *Metric spaces and Complex Analysis* — it turns out that f would have to be constant for f' to exist in any reasonable subset of $\mathbb{C}$.) Extensions to vector-valued function are also straightforward, but basically just amount to doing everything coordinatewise. Function of several variables or functions of vectors are a bit more complicated (see *Multivariable Calculus* or, better, the Part A course *Multidimensional Analysis and Geometry*).

Big-O and little-o notation

When expressing error terms, it is often convenient to use Landau's big-O/little-o notation that was introduced in *Analysis I*.

Definition. If $f, g \colon E \to \mathbb{R}$ we say $f(x) = O(g(x))$ as $x \to p$ if there is a constant M such that $|f(x)/g(x)| \leq M$ for x sufficiently close to p. We say $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to p$.

Here we include the possibility that $p = \pm\infty$ in which case 'sufficiently close to p' means 'sufficiently large' (or 'sufficiently large and negative' when $p = -\infty$).

Example 8.4. We have $x^2 = o(x)$, $\sin x = O(x)$, $\sin x = o(1)$ as $x \to 0$. We have $x = o(x^2)$, $\sin x = O(1)$, $\frac{1}{x} = o(1)$, $\log x = O(x)$ as $x \to \infty$.

Example 8.5. We can write the condition for differentiability in Proposition 8.1 as $f(x_0 + h) = f(x_0) + f'(x_0)h + o(h)$ as $h \to 0$.

Remark. Writing $f(x) = O(g(x))$ or $f(x) = o(g(x))$ is slight abuse of notation as the RHS is really a set of possible functions, one of which matches the LHS. In particular, $o()$ or $O()$ should only appear on the RHS in any equation.[30] It would be very confusing to write e.g., $o(x) = O(\sqrt{x})$.

One sided derivatives

If $E = [a, b]$ then for f to be differentiable at a or b involves taking a one-sided limit. More generally, sometimes it is helpful or necessary to consider one-sided versions of derivatives even when we are not at one end of the domain. We say that f **has a right-derivative** at x_0 if

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$

[30] Or in a multi-line sequence of equations, the O-terms on the RHSs should get progressively weaker (or equivalent) on each successive line.

exists. This is equivalent to asking for the function f restricted to $[x_0, b)$ to have a derivative at x_0. When it does, we denote the limit by $f'_+(x_0)$ (Alternative notation: $f'(x_0^+)$, but this can be confused with $\lim_{x \to x_0^+} f'(x)$, which is not the same thing![31]) Similarly, f **has a left-derivative** at x_0 if

$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}$$

exists, in which case we write it as $f'_-(x_0)$.

The following result is immediate from Proposition 1.15.

Proposition 8.6. *Let $f \colon E \to \mathbb{R}$ and assume $x_0 \in E$ is both a left and right limit point of E. Then the following are equivalent:*

(a) *f is differentiable at x_0;*

(b) *f has both left- and right-derivatives at x_0 and they are equal.*

Example 8.7. Consider $f(x) = |x|$ on $\mathbb{R}$. Here f is differentiable at any $x_0 \neq 0$. At 0 we have one-sided derivatives $f'_-(0) = -1$ and $f'_+(0) = 1$, so $f'(0)$ fails to exist.

This example shows that a function which is continuous at a point x_0 need not be differentiable at x_0.

Example 8.8. Define $f \colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^{3/2}, & \text{for } x > 0; \\ 0, & \text{for } x \leq 0. \end{cases}$$

Then $f'_-(0)$ exists and equals 0, obviously. Also

$$f'_+(0) = \lim_{x \to 0+} \frac{x^{3/2} - 0}{x - 0} = \lim_{x \to 0+} \sqrt{x} = 0.$$

Hence, by Proposition 8.6, $f'(0)$ exists and equals 0. Alternatively, we can give a direct sandwiching argument:

$$\left| \frac{f(x) - f(0)}{x - 0} - 0 \right| \leq \frac{|x|^{3/2}}{|x|} = \sqrt{|x|} \to 0 \text{ as } x \to 0.$$

Now we start assembling the rules of differential calculus as you learned them at school, but now obtained as consequences of AOL for function limits.

Theorem 8.9 (Algebraic properties of differentiation)**.** *Assume that $f, g \colon E \to \mathbb{R}$ are both differentiable at the limit point $x_0 \in E$, and that $a, b \in \mathbb{R}$. Then the following hold.*

(a) **Linearity:** *$af(x) + bg(x)$ is differentiable at x_0 with derivative $af'(x_0) + bg'(x_0)$.*

[31] Although it turns out that if f is continuous and $f'(x_0^+) = \lim_{x \to x_0^+} f'(x)$ exists and then so does $f'_+(x_0)$ and they are equal. See Problem Sheet 6 question 7.

(b) **Product Rule:** *$f(x)g(x)$ is differentiable at x_0 with derivative $f'(x_0)g(x_0)+f(x_0)g'(x_0)$.*

(c) **Quotient Rule:** *Assume $g(x_0) \neq 0$. Then $f(x)/g(x)$ is differentiable at x_0 with derivative*

$$\frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Proof. (a)&(b) We have

$$\begin{aligned} f(x_0 + h) &= f(x_0) + f'(x_0)h + \varepsilon_1(h)h, \\ g(x_0 + h) &= g(x_0) + g'(x_0)h + \varepsilon_2(h)h, \end{aligned}$$

where $\varepsilon_1(h), \varepsilon_2(h) \to 0$ as $h \to 0$. Then

$$af(x_0 + h) + bg(x_0 + h) = af(x_0) + bg(x_0) + \big(af'(x_0) + bg'(x_0)\big)h + \big[a\varepsilon_1(h) + b\varepsilon_2(h)\big]h$$

and

$$\begin{aligned} f(x_0 + h)g(x_0 + h) = f(x_0)g(x_0) &+ \big(f(x_0)g'(x_0) + f'(x_0)g(x_0)\big)h \\ &+ \big[f(x_0)\varepsilon_2(h) + g(x_0)\varepsilon_1(h) + (f'(x_0) + \varepsilon_1(h))(g'(x_0) + \varepsilon_2(h))h\big]h. \end{aligned}$$

By standard AOL for function limits the expressions in square brackets tend to 0 as $h \to 0$. Now by Proposition 8.1 we deduce that $af(x)+bg(x)$ and $f(x)g(x)$ are differentiable at x_0, with derivatives $af'(x_0) + bf'(x_0)$ and $f(x_0)g'(x_0) + f'(x_0)g(x_0)$ respectively.

[If one wanted to write these proofs out using o-notation, one could write:

$$\begin{aligned} af(x_0 + h) + bg(x_0 + h) &= af(x_0) + bg(x_0) + af'(x_0)h + bg'(x_0)h + o(ah) + o(bh) \\ &= af(x_0) + bg(x_0) + (af'(x_0) + bg'(x_0))h + o(h), \\ f(x_0 + h)g(x_0 + h) &= \big(f(x_0) + f'(x_0)h + o(h)\big)\big(g(x_0) + g'(x_0)h + o(h)\big) \\ &= f(x_0)g(x_0) + f'(x_0)g(x_0)h + f(x_0)g'(x_0)h + f'(x_0)g'(x_0)h^2 \\ &\quad + o(f(x_0)h) + o(f'(x_0)h^2) + o(g(x_0)h) + o(g'(x_0)h^2) + o(h^2) \\ &= f(x_0)g(x_0) + (f'(x_0)g(x_0) + f(x_0)g'(x_0))h + o(h). \end{aligned}$$
]

(c) We first give the result when $f(x) := 1$. Note that

$$\frac{1/g(x) - 1/g(x_0)}{x - x_0} = \frac{-1}{g(x)g(x_0)} \cdot \frac{g(x) - g(x_0)}{x - x_0}.$$

Taking limits as $x \to x_0$ and using AOL and continuity of g at x_0 gives that $(1/g)'(x_0)$ exists and

$$\left(\frac{1}{g}\right)'(x_0) = \frac{-1}{g(x_0)^2} \cdot g'(x_0).$$

The general quotient rule can then be obtained by combining this with the product rule:

$$\left(\frac{f}{g}\right)'(x_0) = f'(x_0) \cdot \frac{1}{g(x_0)} + f(x_0) \cdot \frac{-g'(x_0)}{g(x_0)^2} = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$ □

Example 8.10. The power function x^n is differentiable at all points, for $n \in \mathbb{N}$, and so are polynomials, and rational functions at points where the denominator is non zero.

Higher Derivatives

Suppose that $f\colon (a,b) \to \mathbb{R}$ is differentiable at every point of (a,b), then it makes sense to ask if f' is differentiable at $x_0 \in (a,b)$. If it is differentiable then we denote its derivative by $f''(x_0)$

We can seek to iterate this process. Write $f^{(0)} = f$, $f^{(1)} = f'$, and suppose $f^{(0)} = f$, $f^{(1)} = f'$, ..., $f^{(n)}$ have been defined recursively at every point of (a,b) (we make this assumption to simplify matters). If $f^{(n)}$ is differentiable at $x_0 \in (a,b)$ then we say f **is $(n+1)$-times differentiable at** x_0 and we write $f^{(n+1)}(x_0) := (f^{(n)})'(x_0)$.

If f has derivatives of all orders on (a,b) (that is, $f^{(n)}(x_0)$ exists for each $x_0 \in (a,b)$ and for each $n = 1, 2, \ldots$, we say it is **infinitely differentiable on** (a,b).

The following is proved by an easy induction using Linearity and the Product Rule. (Compare with the proof of the binomial expansion of $(1+x)^n$ for n a positive integer.)

Proposition 8.11 (Leibniz' Formula). *Let $f, g\colon (a,b) \to \mathbb{R}$ be n-times differentiable on (a,b). Then $x \mapsto f(x)g(x)$ is n-times differentiable and*

$$(fg)^{(n)}(x) = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x) g^{(n-j)}(x).$$

Proof. Exercise. □

Chain Rule

Theorem 8.12 (Chain Rule). *Assume that $f\colon E \to \mathbb{R}$ and that $g\colon E' \to \mathbb{R}$ with $f(E) \subseteq E'$ (so that $g \circ f\colon E \to \mathbb{R}$ is defined). Suppose further that f is differentiable at the limit point $x_0 \in E$ and that g is differentiable at $f(x_0)$. Then $g \circ f$ is differentiable at x_0 and*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. For convenience write $y_0 = f(x_0)$. Then by Proposition 8.1 we have

$$\begin{aligned} f(x_0 + h) &= f(x_0) + f'(x_0)h + \varepsilon_1(h)h, \\ g(y_0 + \eta) &= g(y_0) + g'(y_0)\eta + \varepsilon_2(\eta)\eta, \end{aligned}$$

where $\varepsilon_1(h), \varepsilon_2(\eta) \to 0$ as $h, \eta \to 0$. We define $\varepsilon_2(0) = 0$ so that ε_2 is continuous at 0 and note that the above also holds for $\eta = 0$. Now set

$$\eta := f(x_0 + h) - f(x_0) = f'(x_0)h + \varepsilon_1(h)h$$

so that

$$\begin{aligned} g(f(x_0 + h)) &= g(y_0 + \eta) = g(y_0) + g'(y_0)\eta + \varepsilon_2(\eta)\eta \\ &= g(y_0) + g'(y_0)f'(x_0)h + \big[g'(y_0)\varepsilon_1(h) + \varepsilon_2(\eta)(f'(x_0) + \varepsilon_1(h))\big]h. \end{aligned}$$

Now $\eta = f'(x_0)h + \varepsilon_1(h)h \to 0$ as $h \to 0$. Thus[32] $\varepsilon_2(\eta) \to 0$ as $h \to 0$. So, by AOL, the expression in square brackets tends to 0 as $h \to 0$. Thus $g(f(x))$ is differentiable at x_0 and the derivative is $g'(y_0)f'(x_0) = g'(f(x_0))f'(x_0)$. □

Example 8.13. Let $f(x) = x^2 \cos\frac{1}{x}$ for $x \neq 0$ and $f(0) = 0$. We shall assume that cos and sin are differentiable with the expected derivatives. This will follow from the Differentiation Theorem for power series (Theorem 8.16).

On $\mathbb{R} \setminus \{0\}$ we can apply the standard differentiation rules, including the Chain Rule, and we get, for $x \neq 0$,

$$f'(x) = 2x\cos\tfrac{1}{x} + \sin\tfrac{1}{x}. \tag{9}$$

Now consider 0: for $x \neq 0$,

$$\left|\frac{f(x) - f(0)}{x - 0}\right| = |x\cos\tfrac{1}{x}| \leq |x| \to 0 \text{ as } x \to 0.$$

Therefore $f'(0)$ exists and equals 0.

Note that (9) shows that $\lim_{x\to 0} f'(x)$ fails to exists (the first term tends to 0, the second one does not have a limit as $x \to 0$, so the sum cannot tend to a limit). We deduce that f' is not continuous at 0. By the contrapositive of Proposition 8.3, $f''(0)$ cannot exist. (Note that f is infinitely differentiable on $\mathbb{R} \setminus \{0\}$.)

Inverse functions

Like the other main results in this section, our final theorem tells us how to build new differentiable functions.

Theorem 8.14 (The Inverse Function Theorem[33] (IFT))**.** *Suppose I is a non-trivial interval and $f\colon I \to \mathbb{R}$ is a strictly monotonic continuous function with inverse function $g\colon f(I) \to I$. Assume that f is differentiable at $x_0 \in I$ and that $f'(x_0) \neq 0$. Then g is differentiable at $f(x_0)$ and*

$$g'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Proof. The statement includes all the assumptions we imposed for the Continuous IFT. Hence $f(I)$ is an interval and $g\colon f(I) \to I$ is continuous and strictly monotonic. Now let $y_0 = f(x_0)$. Then

$$g'(f(x_0)) = \lim_{y\to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \lim_{y\to y_0} \frac{x - x_0}{f(x) - f(x_0)},$$

[32]Note that we could have $\eta = 0$, so it is important that we defined $\varepsilon_2(0) = 0$.

[33]The IFT is usually quoted as saying $f'(x_0) \neq 0$ and f' continuous at x_0 implies f is invertible near x_0, the inverse having the appropriate derivative. But $f'(x_0) \neq 0$ and f' continuous imply f' has a constant sign near x_0 and as we will see later this will imply monotonicity near x_0. The version given here therefore implies the standard form of the IVT, and is in fact stronger.

provided this last limit exists, and where we have defined $x = g(y)$. But g is continuous, so $x \to x_0$ (and $x \neq x_0$ by injectivity of g) as $y \to y_0$ and

$$\lim_{y\to y_0} \frac{x - x_0}{f(x) - f(x_0)} = \lim_{x\to x_0} \frac{x - x_0}{f(x) - f(x_0)} = \lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0}\right)^{-1} = \frac{1}{f'(x_0)}$$

by Theorem 2.8 and AOL. □

Still assuming the Differentiation Theorem for power series and its consequences for the elementary functions, we deduce that the following are differentiable and have the expected derivatives

$$\begin{array}{ll} \log\colon (0,\infty) \to \mathbb{R} & \log'(y) = \frac{1}{y}, \\ \arctan\colon \mathbb{R} \to \mathbb{R} & \arctan'(y) = \frac{1}{1+y^2}. \end{array}$$

To confirm the result for $g(y) = \log y$, note that, for fixed $y_0 \in (0, \infty)$, Theorem 8.14 can be applied with $f(x) = \exp x$. Write $x_0 = \log y_0$ so $y_0 = \exp x_0$. The formula in the theorem gives

$$\log'(y_0) = \frac{1}{\exp'(x_0)} = \frac{1}{\exp(x_0)} = \frac{1}{y_0}.$$

The derivative of arctan is handled similarly, making use of standard trigonometric identities.

Differentiation of power series

Our objective in this section is to prove the Differentiation Theorem for power series which was introduced, but not proved, in *Analysis I*, and states that one can differentiate a power series 'term-by-term' provided one is strictly inside the radius of convergence of the power series.

We will prove this in a manner that works for complex power series as it is an important result in Complex Analysis as well, and the proof is identical.

We first show that the result of term-by-term differentiation is well-defined inside the radius of convergence of the original series.

Lemma 8.15 (ROC of derivative power series)**.** *Suppose the power series $\sum_{k\geq 0} c_k x^k$ has radius of convergence $R \in [0, \infty]$. Then the power series $\sum_{k\geq 1} k c_k x^{k-1}$ also has radius of convergence R.*

Proof. Suppose $|x| < R$. Then by the definition of R there exists y such that $|x| < |y| < R$ and $\sum c_k y^k$ converges. But then $c_k y^k \to 0$ as $k \to \infty$ and in particular the sequence $(c_k y^k)$ is bounded, say $|c_k y^k| \leq M$. Then $|k c_k x^{k-1}| \leq M|y|^{-1} \cdot k(|x|/|y|)^{k-1}$. Now $\sum k(|x|/|y|)^{k-1}$ converges by e.g., the Ratio Test. Thus by the Comparison Test $\sum k c_k x^{k-1}$ is (absolutely) convergent.

Conversely, if $|x| > R$ we know $c_k x^k \not\to 0$, but then clearly $k c_k x^{k-1} \not\to 0$, so $\sum k c_k x^{k-1}$ is divergent. □

Theorem 8.16 (Differentiation Theorem for power series)**.** *Let the real or complex power series $f(x) := \sum_{k=0}^{\infty} c_k x^k$ have radius of convergence $R \in (0, \infty]$. Then f is differentiable in $\{x : |x| < R\}$ and f' is given by term-by-term differentiation:*

$$f'(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}.$$

Proof. Fix $x_0 \in \mathbb{C}$ with $|x_0| < R$ and fix $\rho \in \mathbb{R}$ with $|x_0| < \rho < R$. By Lemma 8.15, $g(x) := \sum_{k=1}^{\infty} k c_k x^{k-1}$ has radius of convergence R and hence $g(x_0)$ is well defined. We also observe, applying Lemma 8.15 again, that $\sum_{k=2}^{\infty} k(k-1) c_k x^{k-2}$ has ROC R, and so converges absolutely at $\rho < R$. In particular

$$M := \sum_{k=2}^{\infty} k(k-1)|c_k|\rho^{k-2} < \infty.$$

Now to show $f'(x_0) = g(x_0)$ it is enough to bound

$$\Big|\frac{f(x) - f(x_0)}{x - x_0} - g(x_0)\Big| = \Big|\sum_{k=1}^{\infty} c_k\Big(\frac{x^k - x_0^k}{x - x_0} - k x_0^{k-1}\Big)\Big| \tag{10}$$

when x is sufficiently close to x_0, so wlog $|x| < \rho$. Summing a geometric series we have

$$\frac{x^k - x_0^k}{x - x_0} = x_0^{k-1} + x_0^{k-2}x + \cdots + x^{k-1},$$

so for $k = 1$ we have $\frac{x^k - x_0^k}{x - x_0} = k x_0^{k-1}$ and for $k \geq 2$ we get

$$\begin{aligned} \frac{x^k - x_0^k}{x - x_0} - k x_0^{k-1} &= (x_0^{k-1} + x_0^{k-2}x + \cdots + x^{k-1}) - (x_0^{k-1} + x_0^{k-1} + \cdots + x_0^{k-1}) \\ &= x_0^{k-1}(1-1) + x_0^{k-2}(x - x_0) + x_0^{k-3}(x^2 - x_0^2) + \cdots + (x^{k-1} - x_0^{k-1}) \\ &= (x - x_0)\cdot\big(x_0^{k-2} + x_0^{k-1}(x + x_0) + \cdots + (x^{k-2} + \cdots + x_0^{k-2})\big). \end{aligned}$$

Hence if $|x_0|, |x| < \rho$ we have

$$\Big|\frac{x^k - x_0^k}{x - x_0} - k x_0^{k-1}\Big| \leq |x - x_0|\big(\rho^{k-2} + 2\rho^{k-2} + \cdots + (k-1)\rho^{k-2}\big) = |x - x_0| \cdot \tfrac{k(k-1)}{2}\rho^{k-2}.$$

But then, by (10),

$$\Big|\frac{f(x) - f(x_0)}{x - x_0} - g(x_0)\Big| \leq |x - x_0| \sum_{k=2}^{\infty} \tfrac{1}{2}k(k-1)|c_k|\rho^{k-2} = \tfrac{M}{2}|x - x_0| \to 0 \text{ as } x \to x_0. \quad \square$$

Example 8.17. The series defining exp, cos, sin, cosh, sinh all have infinite radius of convergence. The Differentiation Theorem gives, for $x \in \mathbb{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\exp(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{k=0}^{\infty}\frac{x^k}{k!} \overset{*}{=} \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^k}{k!} = \sum_{k=1}^{\infty}\frac{x^{k-1}}{(k-1)!} = \exp(x);$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = \mathrm{d}\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k}}{(2k)!} \overset{*}{=} \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}\frac{(-1)^k x^{2k}}{(2k)!} = \sum_{k=1}^{\infty}\frac{(-1)^k x^{2k-1}}{(2k-1)!} = -\sin(x);$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k+1}}{(2k+1)!} \overset{*}{=} \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}\frac{(-1)^k x^{2k+1}}{(2k+1)!} = \sum_{k=1}^{\infty}\frac{(-1)^k x^{2k}}{(2k)!} = \cos(x);$$

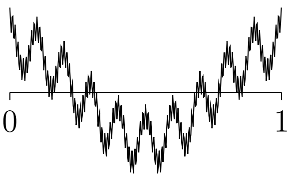
and likewise for $\cosh x$ and $\sinh x$. The occurrences of $\overset{*}{=}$ show the points at which we have differentiated term by term, as the Differentiation Theorem tells us we may.

A continuous but nowhere differentiable function

In this (non-examinable) section we construct a function that is continuous on $\mathbb{R}$, but not differentiable at *any* point. This construction might seem pathological, but in some sense 'most' continuous functions are like this[34].

Define

$$f(x) := \sum_{k=0}^{\infty} 2^{-k}\cos(10^k \cdot 2\pi x).$$



Note that $f(x)$ converges uniformly on $\mathbb{R}$ by the M-test (with $M_k = 2^{-k}$). Hence f is continuous on $\mathbb{R}$. It is also periodic[35] with period 1.

Now comes the difficult bit: showing f is not differentiable anywhere.

Pick $x_0 \in \mathbb{R}$ and define

$$y_n = 10^{-n}\lfloor 10^n x_0\rfloor, \qquad \text{and} \qquad z_n = 10^{-n}\big(\lfloor 10^n x_0\rfloor + \tfrac{1}{2}\big).$$

In other words, y_n is x_0 'rounded down' to n decimal places and z_n appends the digit 5 at the $n+1$st place after the decimal point. Now summing from $k = n$ onwards we have

$$\sum_{k=n}^{\infty} 2^{-k}\cos(10^k \cdot 2\pi y_n) - \sum_{k=n}^{\infty} 2^{-k}\cos(10^k \cdot 2\pi z_n) = 2^{-n}(1-(-1)) + 0 + \cdots = 2\cdot 2^{-n}, \quad (11)$$

as $10^k \cdot 2\pi y_n$ and $10^k \cdot 2\pi z_n$ are an even and odd multiple of π respectively for $k = n$ and both are even multiples of π for all $k > n$. Also, for any x, y, $|\cos(x) - \cos(y)| = |2\sin\frac{x+y}{2}\sin\frac{x-y}{2}| \le 2\cdot 1\cdot |\frac{x-y}{2}| = |x-y|$, so for the first n terms of the sum we have

$$\Big|\sum_{k=0}^{n-1} 2^{-k}\cos(10^k \cdot 2\pi y_n) - \sum_{k=0}^{n-1} 2^{-k}\cos(10^k \cdot 2\pi z_n)\Big| \le \sum_{k=0}^{n-1} 2^{-k}10^k \cdot 2\pi|y_n - z_n|$$
$$= (1 + 5 + \cdots + 5^{n-1}) \cdot 2\pi \cdot \tfrac{1}{2}\cdot 10^{-n} = \tfrac{5^n-1}{5-1}\cdot \pi \cdot 10^{-n} \le 2^{-n}. \quad (12)$$

[34] In the Part B course *Continuous Martingales and Stochastic Calculus* one constructs *Brownian motion*, which is a model of a random continuous function. It turns out that with probability 1 it is nowhere differentiable.

[35] In fact it is a key result in Fourier analysis that any periodic continuous function can be written as an infinite series of trigonometric functions. Thus the form of f is not particularly special.

Hence, by combining (11) and (12) and using the reverse triangle inequality, we have

$$|f(z_n) - f(y_n)| \geq 2 \cdot 2^{-n} - 2^{-n} = 2^{-n}.$$

Now suppose f were differentiable at x_0. Then

$$\begin{aligned} f(y_n) &= f(x_0) + f'(x_0)(y_n - x_0) + o(y_n - x_0), \\ f(z_n) &= f(x_0) + f'(x_0)(z_n - x_0) + o(z_n - x_0). \end{aligned}$$

But then

$$|f(z_n) - f(y_n)| \leq |f'(x_0)||z_n - y_n| + o(|y_n - x_0|) + o(|z_n - x_0|) \leq K \cdot 10^{-n}$$

for any $K > |f'(x_0)|$ when n is sufficiently large as $|y_n - x_0|, |z_n - x_0|, |z_n - y_n| \leq 10^{-n}$. But for large n this contradicts the fact that $|f(z_n) - f(y_n)| \geq 2^{-n}$. Hence $f'(x_0)$ does not exist.

Remark. This example also shows that a uniform limit of differentiable functions is not necessarily differentiable.

9 The Mean Value Theorem

In this section we shall restrict attention to real-valued functions defined on intervals in $\mathbb{R}$. While many of the results we obtained in the previous section for real-valued functions of a real variable have obvious analogues when $\mathbb{R}$ is replaced by $\mathbb{C}$, the theory of differentiability of complex valued functions on the complex plane turns out to be very different from that in the real case and is much more powerful. Complex Analysis is covered within the Part A Core. The results in this section however rely heavily on the fact that the functions are real-valued.

Definition. Let $E \subseteq \mathbb{R}$ and $f \colon E \to \mathbb{R}$.

(a) $x_0 \in E$ is a **local maximum** of f if there exists a $\delta > 0$ such that $f(x) \leq f(x_0)$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap E$.

(b) $x_0 \in E$ is a **local minimum** of f if there exists a $\delta > 0$ such that $f(x) \geq f(x_0)$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap E$.

A local maximum or minimum is called a **local extremum**. If the inequality is strict (for $x \neq x_0$) we will say that the extremum is **strict**.

Here is the crucial property.

Theorem 9.1 (Fermat's Theorem on Extrema). *Let $f \colon (a, b) \to \mathbb{R}$ and suppose that $x_0 \in (a, b)$ is a local extremum and f is differentiable at x_0. Then $f'(x_0) = 0$.*

Proof. If x_0 is a local maximum, then there exists $\delta > 0$ such that whenever $0 < x - x_0 < \delta$ and $x \in (a, b)$,

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

so that

$$f'(x_0) = f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

On the other hand, there exists $\delta > 0$ such that whenever $-\delta < x - x_0 < 0$ and $x \in (a, b)$,

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

so that

$$f'(x_0) = f'_-(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

We conclude that $f'(x_0) = 0$.

A similar argument applies when x_0 is a local minimum. (Or apply the above to $-f$.) □

Remark. In Fermat's theorem it is essential that the interval (a, b) is open. Why?

We now apply Fermat's Theorem to obtain a simple criterion for the existence of a point where $f' = 0$.

Theorem 9.2 (Rolle's Theorem)**.** *Let $a < b$ and $f\colon [a, b] \to \mathbb{R}$. Assume that*

(a) *f is continuous on $[a, b]$;*

(b) *f is differentiable on (a, b);*

(c) *$f(a) = f(b)$.*

Then there exists $\xi \in (a, b)$ such that $f'(\xi) = 0$.

Proof. As f is continuous on $[a, b]$ it is bounded and attains its maximum and minimum on $[a, b]$ (by the Boundedness Theorem). If $f(x_0) > f(a)$ for some $x_0 \in [a, b]$ let $\xi \in [a, b]$ be such that $f(\xi) = \sup_{x \in [a,b]} f(x)$. As $f(\xi) \geq f(x_0) > f(a) = f(b)$, $\xi \in (a, b)$. Also ξ is a clearly a local maximum of f and so by Fermat's result $f'(\xi) = 0$. Similarly if $f(x_0) < f(a)$ for some $x_0 \in [a, b]$ we can take $\xi \in [a, b]$ such that $f(\xi) = \inf_{x \in [a,b]} f(x)$. The only remaining case is if $f(x_0) = f(a)$ for all $x_0 \in [a, b]$. But then $f(x)$ is a constant and so $f'(\xi) = 0$ for any $\xi \in (a, b)$. □

Need continuity on all of $[a, b]$. | Need differentiability on all of (a, b). | Need $f(a) = f(b)$. | All conditions satisfied.

When using the theorem remember to check *all* conditions including the continuity and differentiability conditions. For example, $f\colon [0, 1] \to \mathbb{R}$ defined by $f(x) = x$ for $x \in$

$[0,1)$ and $f(1) = 0$ satisfies all the conditions except continuity at 1. The function $g\colon [-1,1] \to \mathbb{R}$ given by $g(x) = |x|$ satisfies all conditions except that g is not differentiable at $x = 0$. And the function $h\colon [0,1] \to \mathbb{R}$ given by $h(x) = x$ satisfies all conditions except $h(0) = h(1)$. But in all three cases there is no point at which the derivative is zero.

Remember that f is differentiable implies that f is continuous. Thus the hypotheses (a) and (b) would be satisfied if f was differentiable on $[a,b]$ (with one-sided derivatives at the endpoints). However, often it is important that Rolle holds under the given weaker conditions.

One way of expressing Rolle's Theorem informally is by saying

'Between any two zeros of f there is a zero of f'.'

The following is an example where Rolle's Theorem is applied several times in this form.

Example 9.3. Assume that the real-valued function f is twice differentiable on $[0,1]$ and that f''' exists in $(0,1)$. Assume in addition that $f(0) = f'(0) = f(1) = f'(1) = 0$. To prove: that there exists a point $\xi \in (0,1)$ at which $f'''(\xi) = 0$.

The conditions are satisfied to apply Rolle's Theorem to f on $[0,1]$ and so there exists $\alpha \in (0,1)$ such that $f'(\alpha) = 0$. Now the conditions are satisfied to apply Rolle's Theorem to f' on each of $[0,\alpha]$ and $[\alpha,1]$ to obtain β_1 and β_2 with $0 < \beta_1 < \alpha < \beta_2 < 1$ and $f''(\beta_1) = f''(\beta_2) = 0$. Finally, since $\beta_1, \beta_2 \in (0,1)$ on which f''' is given to exist, we know f'' is continuous on $[\beta_1,\beta_2]$ and differentiable on (β_1,β_2), so we can apply Rolle's Theorem to f'' on $[\beta_1,\beta_2]$ to obtain the required point $\xi \in (\beta_1,\beta_2)$ with $f'''(\xi) = 0$.

The next Big Theorem is one of the most important and useful in the course. Is is easily derived for Rolle's Theorem by adding a suitable linear function to f to make the endpoints agree.

Theorem 9.4 (Mean Value Theorem (MVT))**.** *Let $a < b$ and $f\colon [a,b] \to \mathbb{R}$. Assume*

(a) *f is continuous on $[a,b]$; and*

(b) *f is differentiable on (a,b).*

Then there exists $\xi \in (a,b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Define $F(x) := f(x) - f(a) - K(x-a)$ where K is chosen so that $F(b) = F(a) = 0$, namely

$$K := \frac{f(b) - f(a)}{b - a}.$$

Certainly $F\colon [a,b] \to \mathbb{R}$ is continuous, F is differentiable on (a,b) and, by choice of K, $F(a) = F(b)$. Thus Rolle's Theorem applies, and so $F'(\xi) = 0$ for some $\xi \in (a,b)$. But $F'(x) = f'(x) - K$ so

$$f'(\xi) = K = \frac{f(b) - f(a)}{b - a}. \qquad \square$$

[For examples showing that all the conditions in the MVT are required, take the counterexamples following Rolle's Theorem and tilt your page/screen a bit ☺.]

The following is a surprisingly useful generalisation of the Mean Value Theorem with a very similar proof.

Theorem 9.5 (Cauchy's MVT *or* Generalised MVT)**.** *Let $a < b$ and $f, g\colon [a,b] \to \mathbb{R}$. Assume*

(a) *f, g are continuous on $[a,b]$; and*

(b) *f, g are differentiable on (a,b).*

Then there exists $\xi \in (a,b)$ such that

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

If in addition $g'(x) \neq 0$ for all $x \in (a,b)$, then $g(b) \neq g(a)$ and the conclusion can be written

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Remark. We cannot obtain this result by applying the MVT to f and g individually since that way we'd obtain two 'ξ's, one for f and one for g, and these would in general not be equal.

Proof. Suppose first that $g(b) \neq g(a)$. Define $F(x) := f(x) - f(a) - K(g(x) - g(a))$, where K is chosen so that $F(b) = F(a) = 0$, namely

$$K := \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Then F is continuous on $[a,b]$, differentiable on (a,b) and $F(a) = F(b)$. Hence by Rolle's theorem there exists $\xi \in (a,b)$ such that

$$F'(\xi) = f'(\xi) - Kg'(\xi) = 0,$$

or equivalently

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a))$$

as required.

If $g(b) = g(a)$ then by Rolle's theorem there is a point $\xi \in (a,b)$ with $g'(\xi) = 0$, and this ξ satisfies the required equation. Thus if $g'(x) \neq 0$ for all $x \in (a,b)$ then we must have $g(b) \neq g(a)$ and the last statement of the theorem follows by simple algebra. □

Here is one of the most useful corollaries of the MVT.

Theorem 9.6 (Constancy Theorem)**.** *Let I be an interval and $f\colon I \to \mathbb{R}$ be differentiable with $f'(x) = 0$ for all $x \in I$. Then f is constant on I.*

Note that the interval I need not be bounded or closed, but it *does* need to be an interval: $f\colon (1,2) \cup (3,4) \to \mathbb{R}$ defined by $f(x) = 1$ for $x \in (1,2)$ and $f(x) = 2$ for $x \in (3,4)$ is clearly differentiable with zero derivative for all $x \in (1,2) \cup (3,4)$, but is also not constant.

Proof. For any $a, b \in I$ with $a < b$ apply the MVT to f on $[a,b]$. (Note that f is differentiable on I implies that f is continuous on $[a,b] \subseteq I$.) Then $f(b) - f(a) = f'(\xi)(b-a)$ for some $\xi \in (a,b) \subseteq I$. But $f'(\xi) = 0$, so that $f(b) = f(a)$. Since this holds for all $a < b$ with $a, b \in I$, f is constant on I. □

The following examples illustrate a method of using the Constancy Theorem to solve certain differential equations. The 'trick' is to manipulate them so that they look like $\frac{\mathrm{d}}{\mathrm{d}x}F = 0$ for some function F.

Example 9.7. Suppose that f is a function on an interval I whose derivative is x^2. Then there exists a constant C such that, for all $x \in I$, $f(x) = \frac{1}{3}x^3 + C$.

Let $F(x) := f(x) - \frac{1}{3}x^3$. Then F is differentiable and $F'(x) = x^2 - x^2 = 0$. By the Constancy Theorem $F(x) = C$ for some constant C and hence $f(x) = \frac{1}{3}x^3 + C$.

Example 9.8 ($\exp(x+y) = \exp(x)\exp(y)$)**.** Fix a constant c and consider $F(x) = \exp(x)\exp(c-x)$. Then using the Chain rule, Product rule, and $\exp'(x) = \exp(x)$ (obtained by the Differentiation Theorem for power series) we obtain

$$F'(x) = \exp(x)\exp(c-x) - \exp(x)\exp(c-x) = 0.$$

We deduce that $F(x)$ is a constant: $\exp(x)\exp(c-x) = F(x) = F(0) = 1 \cdot \exp(c)$. Substituting $c = x+y$ now gives $\exp(x+y) = \exp(x)\exp(y)$ for all $x, y \in \mathbb{R}$.

Note that similar methods allow for proofs of all the usual trigonometric identities, at least for real numbers.

Example 9.9 (Trigonometric addition formulae)**.** Recall that $\sin(x)$ and $\cos(x)$ are defined via power series on the whole of $\mathbb{R}$ and that $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$ followed from the Differentiation Theorem for power series. Fix a constant c and consider $F(x) = \cos(x)\cos(c-x) - \sin(x)\sin(c-x)$. Then using the chain rule and product rule

$$F'(x) = -\sin(x)\cos(c-x) + \cos(x)\sin(c-x) - \cos(x)\sin(c-x) + \sin(x)\cos(c-x) = 0.$$

We deduce that $F(x)$ is a constant: $\cos(x)\cos(c-x) - \sin(x)\sin(c-x) = F(x) = F(0) = \cos(c)$. Substituting $c = x+y$ now gives

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y). \tag{13}$$

Similarly (or by differentiation w.r.t. x)

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y). \tag{14}$$

Substituting $y = -x$ into the formula for $\cos(x+y)$ and noting that $\sin(-x) = -\sin(x)$ also gives the well-known formula

$$\cos^2 x + \sin^2 x = 1. \tag{15}$$

for all real x. All these also hold for complex x and y — see the supplementary material on the exponential function on the website.

Example 9.10. We shall show that the general solution of the equation $f'(x) = \lambda f(x)$ for all $x \in \mathbb{R}$, is $f(x) = ae^{\lambda x}$ where a is a constant. (That is, every solution is of this form.)

We spot that $e^{\lambda x}$ is a solution, so consider $F(x) := f(x)/e^{\lambda x} = e^{-\lambda x} f(x)$. Then $F'(x) = f'(x)e^{-\lambda x} - f(x)\lambda e^{-\lambda x} = 0$. Hence, by the Constancy Theorem $F(x)$ is constant, $F(x) = a$; that is all solutions are of the form $f(x) = ae^{\lambda x}$.

Corollary 9.11 (Derivatives and monotonicity)**.** *Let I be an interval and let $f \colon I \to \mathbb{R}$ be differentiable.*

(a) *If $f'(x) \geq 0$ for all $x \in I$ then f is increasing on I.*

(b) *If $f'(x) \leq 0$ for all $x \in I$ then f is decreasing on I.*

(c) *If $f'(x) > 0$ for all $x \in I$ then f is strictly increasing on I.*

(d) *If $f'(x) < 0$ for all $x \in I$ then f is strictly decreasing on I.*

Proof. Simply fix $a, b \in I$ with $a < b$ and apply MVT to f on $[a, b]$ to get $f(b) - f(a) = f'(\xi)(b - a)$ for some $\xi \in (a, b) \subseteq I$. □

⚠ **Remark.** x^3 is strictly increasing on $\mathbb{R}$ but has derivative 0 at $x = 0$, so the converses to (c) and (d) do not hold.

Example 9.12 (Alternating bounds on sin and cos)**.** By (15) we have

$$\cos x \leq 1.$$

If we set $F(x) := x - \sin x$ then $F'(x) = 1 - \cos x \geq 0$. Hence, by Corollary 9.11, F is increasing: $F(x) \geq F(0) = 0$ for all $x \geq 0$. Thus

$$\sin x \leq x \qquad \text{for } x \geq 0.$$

If we set $F(x) := 1 - \frac{x^2}{2} - \cos x$ then $F'(x) = \sin x - x \leq 0$ for $x \geq 0$ so, by Corollary 9.11, F is decreasing: $F(x) \leq F(0) = 0$ for all $x \geq 0$. Thus

$$\cos x \geq 1 - \frac{x^2}{2} \qquad \text{for } x \geq 0.$$

If we set $F(x) := x - \frac{x^3}{6} - \sin x$ then $F'(x) = 1 - \frac{x^2}{2} - \cos x \leq 0$ for $x \geq 0$ so, by Corollary 9.11, F is decreasing: $F(x) \leq F(0) = 0$ for all $x \geq 0$. Thus

$$\sin x \geq x - \frac{x^3}{6} \qquad \text{for } x \geq 0.$$

If we set $F(x) := 1 - \frac{x^2}{2} + \frac{x^4}{24} - \cos x$ then $F'(x) = \sin x - x + \frac{x^3}{6} \geq 0$ for $x \geq 0$ so, by Corollary 9.11, F is increasing: $F(x) \geq F(0) = 0$ for all $x \geq 0$. Thus

$$\cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24} \qquad \text{for } x \geq 0.$$

And so on and so on ... Inductively we can bound both sin and cos above and below by their series terminated at odd or even numbers of terms respectively (exercise: write out a proof of this).

Example 9.13 (π)**.** We can define π as twice the smallest positive solution of $\cos x = 0$. Indeed, $1 - \frac{x^2}{2} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{24}$ by Example 9.12, so $\cos x > 0$ for $x < \sqrt{2}$, but $\cos 2 \le -\frac{1}{3}$. Thus by the IVT, there exists some π, $2\sqrt{2} < \pi < 4$ with $\cos\frac{\pi}{2} = 0$. Moreover, this value is unique as $\cos' x = -\sin x$ and $\sin x \ge x - \frac{x^3}{6} = x(1 - \frac{x^2}{6}) > 0$ for $0 < x < 2$, so $\cos x$ is strictly decreasing on $[0, 2]$. As $\sin\frac{\pi}{2} > 0$ and $\cos\frac{\pi}{2} = 0$ we deduce from (15) that $\sin\frac{\pi}{2} = 1$. Then from (13) and (14) we deduce that

$$\cos(x + \tfrac{\pi}{2}) = -\sin(x) \qquad \sin(x + \tfrac{\pi}{2}) = \cos(x).$$

Applying this four times gives $\cos(x + 2\pi) = \cos x$ and $\sin(x + 2\pi) = \sin x$.

Example 9.14 (Lipschitz functions revisited)**.** Suppose I is an interval and $f\colon I \to \mathbb{R}$ is differentiable with bounded derivative, $|f'(x)| \le M$ for all $x \in I$. Then f is Lipschitz continuous: by the MVT $|f(x) - f(y)| = |f'(\xi)||x - y| \le M|x - y|$ for some ξ between x and y.

If $I = [a, b]$ and in addition f' is continuous, then f' is bounded by the Boundedness Theorem. Hence any continuously differentiable[36] function on a closed bounded interval is Lipschitz continuous.

Warning. $f\colon [0, 1] \to \mathbb{R}$ defined by $f(x) = \sqrt{x}$ does not satisfy these conditions even though f is continuously differentiable on $(0, 1)$. We need the derivatives at the endpoints as well here.

Example 9.15 (Bernoulli's inequality)**.** In *Analysis I* you met the useful inequality

$$(1 + x)^r \ge 1 + rx \qquad \text{for } x > -1,\ r \in \mathbb{N}.$$

This was proved by induction on r. We now prove it for all *real* $r \ge 1$. First we note that the standard formula for the derivative of a power still holds:

$$\frac{\mathrm{d}}{\mathrm{d}x}x^r = \frac{\mathrm{d}}{\mathrm{d}x}\exp(r\log x) = \frac{r}{x}\exp(r\log x) = r\exp(r\log x - \log x) = rx^{r-1}$$

for $x > 0$ and any $r \in \mathbb{R}$. Now consider $F(x) = (1 + x)^r - (1 + rx)$. Then $F'(x) = r(1 + x)^{r-1} - r = r((1 + x)^{r-1} - 1)$. Then for $r \ge 1$ and $x \ge 0$, $(1 + x)^{r-1} \ge 1$ ($\exp((r - 1)\log x)$ is increasing in x), so $F'(x) \ge 0$ and hence F is increasing for $x \ge 0$. Thus $F(x) \ge F(0) = 0$ for $x \ge 0$. Similarly $(1 + x)^{r-1} \le 1$ for $x \in (-1, 0]$, so $F'(x) \le 0$ there and so $F(x) \ge F(0) = 0$ for $x \in (-1, 0]$.

Example 9.16 (Jordan's inequality)**.** $\frac{2}{\pi} \le \frac{\sin x}{x} \le 1$ for $x \in (0, \frac{\pi}{2}]$.

Proof. We have already proved the second inequality in Example 9.12 and to prove the first it is enough to show that $F(x) := \frac{\sin x}{x}$ is decreasing on $(0, \frac{\pi}{2}]$ as $F(\frac{\pi}{2}) = \frac{\pi}{2}$. Differentiation gives

$$F'(x) = \frac{x\cos x - \sin x}{x^2}.$$

So let's consider the derivative of $G(x) := x\cos x - \sin x$ on $(0, \frac{\pi}{2}]$. We have $G'(x) = -x\sin x + \cos x - \cos x = -x\sin x \le 0$ as $\sin x > 0$ on $(0, \frac{\pi}{2}]$. Hence G is decreasing so $G(x) \le G(0) = 0$ on $(0, \frac{\pi}{2}]$. Hence $F'(x) \le 0$ and so $F(x)$ is decreasing on $(0, \frac{\pi}{2}]$. □

[36] A **continuously differentiable** function is, of course, a differentiable function f for which f' is continuous.

10 Taylor's Theorem

Our objective in this section to investigate how a real-valued function may be approximated by a polynomial. We emphasise that our methods rely on Rolle's Theorem and the Mean Value Theorem. This means that the results of this section are for *real-valued functions only.*

We begin by noting that the very definition of differentiability concerns the approximation of a function by a linear function. Indeed $f'(x_0)$ exists if and only if we can write

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \varepsilon(h)h$$

for some $\varepsilon(h) \to 0$ as $h \to 0$. Using Landau's notation this is equivalent to

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h) \quad \text{as } h \to 0.$$

The Mean Value Theorem gives another approximation, but with the added assumption that f' exists in an interval. We have

$$f(x_0 + h) = f(x_0) + f'(\xi)h$$

for some ξ between x_0 and $x_0 + h$.

Suppose we wanted a better approximation to f near x_0. A natural generalization would be to approximate f with a quadratic, say

$$f(x_0 + h) \approx f(x_0) + f'(x_0)h + Kh^2.$$

Assuming f has a second derivative, if would seem reasonable to choose K so the second derivatives matched. (Then the first derivatives of both sides would agree with just an $o(h)$ error, and integrating this over a length h would give an error of $o(h^2)$.) This suggests that we should take $K = \frac{1}{2}f''(x_0)$ and that

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \tfrac{f''(x_0)}{2}h^2 + o(h^2).$$

More generally we could imaging higher and higher degree polynomial approximations to f, assuming f has derivatives we can match to sufficiently high order. Even better would be an extension of the MVT as this gives more control over the error, possibly something like

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \tfrac{f''(\xi)}{2}h^2.$$

Taylor's Theorem gives such an extension. We phrase the following in a similar way to the MVT so as to give a natural generalisation of Theorem 9.4.

Theorem 10.1 (Taylor's Theorem)**.** *Let $a < b$ and $f\colon [a,b] \to \mathbb{R}$. Let $n \geq 0$ be such that*

(a) *f, f', ..., $f^{(n)}$ exist and are continuous on $[a,b]$;*

(b) *$f^{(n+1)}$ exists on (a,b).*

Then there exists $\xi \in (a,b)$ such that

$$f(b) = f(a) + f'(a)(b-a) + \tfrac{f''(a)}{2!}(b-a)^2 + \cdots + \tfrac{f^{(n)}(a)}{n!}(b-a)^n + \tfrac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1}.$$

The same holds with $b < a$ using intervals $[b,a]$ and (b,a) in place of $[a,b]$ and (a,b).

Proof. We will use induction on n. The case $n = 0$ is precisely the MVT: $f(b) = f(a) + f'(\xi)(b-a)$ for some $\xi \in (a,b)$.

Now assume $n > 0$ and define $F\colon [a,b] \to \mathbb{R}$ by

$$F(x) := f(x) - f(a) - f'(a)(x-a) - \cdots - \tfrac{f^{(n)}(a)}{n!}(x-a)^n - \tfrac{K}{(n+1)!}(x-a)^{n+1}.$$

where K is a constant chosen so that $F(b) = 0$. We also clearly have $F(a) = 0$ and, by assumption, F is continuous on $[a,b]$ and differentiable on (a,b). Hence by Rolle's Theorem there exists $c \in (a,b)$ such that $F'(c) = 0$. Now

$$F'(x) = f'(x) - 0 - f'(a) - f''(a)(x-a) - \cdots - \tfrac{f^{(n)}(a)}{(n-1)!}(x-a)^{n-1} - \tfrac{K}{n!}(x-a)^n$$

and by induction, applying the $n-1$ case of the theorem to f' on $[a,c]$, we have

$$f'(c) = f'(a) + f''(a)(c-a) + \cdots + \tfrac{f^{(n)}(a)}{(n-1)!}(c-a)^{n-1} + \tfrac{f^{(n+1)}(\xi)}{n!}(c-a)^n$$

for some $\xi \in (a,c) \subseteq (a,b)$. But then

$$\begin{aligned} 0 = F'(c) &= f'(c) - f'(a) - f''(a)(c-a) - \cdots - \tfrac{f^{(n)}(a)}{(n-1)!}(c-a)^{n-1} - \tfrac{K}{n!}(c-a)^n \\ &= \tfrac{f^{(n+1)}(\xi)}{n!}(c-a)^n - \tfrac{K}{n!}(c-a)^n. \end{aligned}$$

Thus $K = f^{(n+1)}(\xi)$. Recalling that we chose K so that $F(b) = 0$, the required result drops out.

The case when $b < a$ is similar, or can be deduced from the above result by applying it to $f(-x)$ considered as a function $[-a,-b] \to \mathbb{R}$ and carefully tracking all the sign changes. □

We can write Taylor's theorem in a form that matches our previous discussion by taking $a = x_0$ and $b = x_0 + h$:

$$f(x_0+h) = f(x_0) + f'(x_0)h + \cdots + \tfrac{f^{(n)}(x_0)}{n!}h^n + \tfrac{f^{(n+1)}(x_0+\theta h)}{(n+1)!}h^{n+1}$$

where $0 < \theta < 1$, h can be either positive or negative and $f, f', \ldots, f^{(n+1)}$ are assumed to exist in the appropriate ranges.

It is important to realise that the number θ here depends on h (and on x_0, which we regard as fixed). We have in general no information on how θ varies with h, though it may sometimes be possible to get information in the limit as $h \to 0$ (see problem sheet 7).

The further $x_0 + h$ is from x_0 the less likely the polynomial part is to give a good approximation to $f(x_0 + h)$. Moreover it may be hard in specific cases to find a tight estimate of the size of the error term $\frac{h^{n+1}}{(n+1)!} f^{(n+1)}(x_0 + \theta h)$ especially since the value of θ is not known, so that we need a global upper bound covering all possible values of $x_0 + \theta h$. However, on the assumption that $f^{(n+1)}$ is bounded on $[x_0, x_0 + h]$ (which would follow if it were continuous there) we do have

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \cdots + \tfrac{f^{(n)}(x_0)}{n!} h^n + O(h^{n+1}).$$

Example 10.2. Consider the expansion of $f(x) = \log(1 + x)$ around $x = 0$. We have $f'(x) = \frac{1}{1+x}$, and by induction

$$f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n} \qquad \text{for all } n \geq 1.$$

This gives

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + (-1)^n \frac{x^{n+1}}{(n+1)(1+\theta x)^{n+1}}. \tag{16}$$

We note that $1 + \theta x$ lies between 1 and $1 + x$, so for example, with $n = 2$ we have

$$x - \frac{x^2}{2} + \frac{x^3}{3(1+x)^3} \leq \log(1 + x) \leq x - \frac{x^2}{2} + \frac{x^3}{3},$$

for $x > -1$ (consider $x \in (-1, 0)$ and $x \geq 0$ separately).

Infinite Taylor series

A natural question is whether we can just let $n \to \infty$ in Taylor's Theorem and obtain an infinite power series for f. The answer is unfortunately 'No' in general.

One obvious obstruction is that the higher derivatives may simply not exist. We have seen examples of continuous functions that are continuous but not differentiable at a point. It is relatively easy to construct examples that are n times differentiable but not $n + 1$ times differentiable. One such example is

$$f(x) = |x|^{n+1/2}$$

which is n but not $n+1$ times differentiable at $x = 0$. (One can even get examples where this happens at every x. For example, one can integrate the example on page 57 n times.)

But let's assume f is **infinitely differentiable**, that is $f^{(n)}(x)$ exists for all $n \geq 0$ and all x in the domain of f. Is this enough to get the Taylor series to converge to f? Again, the answer is 'No' in general, however often it works. To see when it works, write

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} x^k + E_n(h),$$

where $E_n(h) = \frac{f^{(n+1)}(x_0+\theta h)}{(n+1)!}h^{n+1}$ is the error term in Taylor's Theorem. By AOL

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}x^k = \lim_{n\to\infty}(f(x_0+h) - E_n(h)) = f(x_0+h) - \lim_{n\to\infty} E_n(h),$$

if this last limit exists. Thus $f(x_0+h)$ is given by the infinite power series if and only if $E_n(h) \to 0$ as $n \to \infty$ (with x_0 and h fixed).

Example 10.3. Continuing the example of $\log(1+x)$, we construct the infinite Taylor series

$$f(x) := x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k}.$$

To determine whether or not this is really $\log(1+x)$ we look at the error term in (16)

$$E_n = \frac{(-1)^n}{n+1}\Big(\frac{x}{1+\theta_n x}\Big)^{n+1}.$$

Note that $\theta = \theta_n$ depends on n (as well as x). As $0 < \theta_n < 1$ we see that if $x \in [-\frac{1}{2}, 1]$ we have $|x/(1+\theta_n x)| \leq 1$ (for negative x we need $1 + \theta_n x \geq 1 - |x|$ to be at least $|x|$, so $x \geq -\frac{1}{2}$). Thus for $x \in [-\frac{1}{2}, 1]$, $|E_n| \leq \frac{1}{n+1} \to 0$ as $n \to \infty$ and so $f(x) = \log(1+x)$.

For $x > 1$ the series $f(x)$ does not converge (by e.g., the Ratio Test), so we don't have an infinite power series for $\log(1+x)$, despite the fact that $\log(1+x)$ is perfectly well defined and infinitely differentiable between 0 and x.

For $x \leq -1$ we could not hope for a series expression for $\log(1+x)$ as $\log(1+x)$ is not defined.

This leaves the cases when $-1 < x < -\frac{1}{2}$ where the series $f(x)$ happily converges, but it is not clear whether or not it converges to $\log(1+x)$ as we do not have enough control over the error term E_n.

In this case it turns out that $f(x)$ does indeed equal $\log(1+x)$. We can use the Differentiation Theorem for power series to deduce that

$$f'(x) = 1 - x + x^2 - x^3 + \cdots = \frac{1}{1+x}$$

for $|x| < 1$ (the radius of convergence of f is $R = 1$). Thus $g(x) := f(x) - \log(1+x)$ has derivative 0 in $|x| < 1$ and so by the Constancy Theorem $g(x)$ is a constant for $|x| < 1$. As clearly $g(0) = 0$ we have

$$f(x) = \log(1+x) \qquad \text{for } x \in (-1, 1).$$

We note that Taylor's theorem also gave this for $x = 1$, so we deduce that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \qquad \text{for } -1 < x \leq 1.$$

Note that Taylor's theorem failed to prove this for $x \in (-1, -\frac{1}{2})$, although only because we did not have good enough bounds on θ_n. On the other hand the Constancy Theorem approach failed at $x = 1$, while the Taylor's Theorem approach worked there.

[The case $x = 1$ is also a spin-off of the definition of the Euler–Mascheroni constant, see the *Analysis I* notes page 100. It is also a consequence of Abel's Continuity Theorem, the (non-examinable) Theorem 12.9 below.]

The above example shows that the infinite Taylor series may fail to converge even when the function is infinitely differentiable in the appropriate range. Could it be therefore that it is just convergence of the power series that we need? Unfortunately the answer is again 'No' in general. It is possible that $E_n(h)$ might converge to a non-zero value and so the Taylor series converges, but to the wrong value!

Example 10.4. Consider $f\colon R \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} e^{-1/x^2}, & \text{for } x \neq 0; \\ 0, & \text{for } x = 0. \end{cases}$$

Some experimentation shows that we expect

$$f^{(k)}(x) := \begin{cases} Q_k(1/x)e^{-1/x^2}, & \text{for } x \neq 0; \\ 0, & \text{for } x = 0. \end{cases}$$

for some polynomial Q_k of degree $3k$. We can prove this by induction: at points $x \neq 0$ this is routine use of linearity, the product rule and the chain rule. But at $x = 0$ we need to take more care, and use the definition:

$$\frac{f^{(k)}(x) - f^{(k)}(0)}{x - 0} = x^{-1}Q_k(1/x)e^{-1/x^2}$$

which we must prove tends to zero as $x \to 0$. Change the variable to $t = 1/x$, then we have $tQ_k(t)e^{-t^2}$ which is a finite sum of terms like $t^s e^{-t^2}$, which we know tend to zero as $|t|$ tends to infinity.

So for this function f the series $\sum \frac{f^{(k)}(0)}{k!} x^k = 0$, so converges to 0 at every x. But the error term $E_n(x)$ is the same for all n (it equals $f(x)$) and so does not tend to 0 at any point except 0.

Note that we can add this function to $\exp x$ and $\sin x$ and so on, and get functions with the same set of derivatives at 0 as these functions, so that they will have the same Taylor polynomials—but are different functions.

Example 10.5. We can even construct infinitely differentiable functions whose Taylor series have zero radius of convergence. For example, let

$$f(x) := \sum_{k=1}^{\infty} \frac{\sin(k^3 x)}{k^k}.$$

We note that this converges (very quickly). With a bit of work one can show that

$$f'(x) = \sum_{k=1}^{\infty} \frac{\cos(k^3 x)}{k^{k-3}}.$$

This is not as easy as it looks! Here is one approach. Set $g(x) := \sum_{k=1}^{\infty} \frac{\cos(k^3 x)}{k^{k-3}}$. Then by applying Taylor's theorem to $\sin(k^3 x)$ we have

$$f(x+h) = \sum_{k=1}^{\infty} \frac{\sin(k^3 x) + k^3 \cos(k^3 x)h - \frac{1}{2}k^6 \sin(k^3 \xi_k)h^2}{k^k} = f(x) + g(x)h + \varepsilon(h)h,$$

for some ξ_k between x and $x+h$ and where

$$|\varepsilon(h)| = \Big| \sum_{k=1}^{\infty} \frac{-\frac{1}{2}k^6 \sin(k^3 \xi_k)h}{k^k} \Big| \le |h| \sum_{k=1}^{\infty} \frac{1}{2k^{k-6}}.$$

But $\sum_{k=1}^{\infty} \frac{1}{2k^{k-6}}$ converges to a constant, so $\varepsilon(h) \to 0$ as $h \to 0$.

In general

$$f^{(n)}(x) = \pm \sum_{k=1}^{\infty} \frac{\sin(k^3 x)}{k^{k-3n}} \quad \text{or} \quad \pm \sum_{k=1}^{\infty} \frac{\cos(k^3 x)}{k^{k-3n}}.$$

Now $|f^{(n)}(0)| = \sum_k k^{3n-k} \ge n^{3n-n} = n^{2n}$ for n odd. Thus as $n! \le n^n$, $|f^{(n)}(0)/n!| \ge n^n$. One can then deduce that the series $\sum \frac{f^{(n)}(0)}{n!} x^n$ has zero radius of convergence (as for $x \neq 0$ the terms don't tend to 0).

Example 10.6 (Real power series)**.** Suppose we have a function defined by

$$f(x) := \sum_{k=0}^{\infty} c_k x^k \qquad |x| < R$$

where $R > 0$ is the radius of convergence of $\sum c_k x^k$. Then the Differentiation Theorem for power series tells us that f has derivatives of all orders. Moreover, by induction,

$$f^{(n)}(x) = \sum_{k=n}^{\infty} c_k k(k-1)\cdots(k-n+1)x^{k-n} \qquad |x| < R,$$

so in particular $f^{(n)}(0) = n! c_n$. Therefore $c_n = f^{(n)}(0)/n!$ and so by definition

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(k)}(0)}{k!}x^k + \cdots \qquad |x| < R.$$

So *if* we knew f could be expressed as a power series, then the infinite Taylor expansion would be that power series.

Example 10.7. Suppose $f\colon \mathbb{R} \to \mathbb{R}$ has the property that for all x, $f'(x) = f(x)$ and $f(0) = 1$. Assuming such an f exists, and without knowing anything about the exponential function, we deduce that $f^{(n)}(x) = f(x)$ exists and is continuous for all n

(continuous as f' exists). But then $f^{(n)}(x)$ is bounded on any fixed interval $[-N, N]$, say $|f^{(n)}(x)| = |f(x)| \le M$ with M independent of n. Hence by Taylor's theorem we deduce that

$$f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + E_n(x)$$

where $|E_n(x)| \le \frac{Mx^{n+1}}{(n+1)!}$. As $E_n(x) \to 0$ as $n \to \infty$, we deduce that in fact $f(x)$ is given by the infinite Taylor series

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

As the Differentiation Theorem for power series shows that in fact this power series differentiates to itself, we deduce that in fact such a function f does exist (and is probably interesting enough to give a name to!).

Many other differential equations can be 'solved' in a similar manner.

In fact, a power series $f(x)$ can be expressed as an infinite Taylor series about *any* point x_0 strictly inside its radius of convergence.

Theorem 10.8. *Suppose $f(x) = \sum c_k x^k$ is a real or complex power series with radius of convergence $R \in (0, \infty]$. Then for $|x_0| < R$*

$$f(x_0 + h) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} h^k$$

for all h with $|h| + |x_0| < R$.

Proof. See Part A *Metric spaces and Complex Analysis.* □

We say a function f is **analytic** at a point x_0 if there exists some $\delta > 0$ such that one can write $f(x_0 + h)$ as a power series for $|h| < \delta$. For example, most standard functions such sin, log, etc., are analytic over much if not all of their domains. By the Differentiation Theorem this implies that f is infinitely differentiable. By Example 10.6 it is also equivalent (for real-valued functions) to the Taylor series of f about x_0 converging to the function, at least when h is sufficiently small. Theorem 10.8 states that any power series is analytic within its radius of convergence. Examples 10.4 and 10.5 give functions that are infinitely differentiable at 0 but not analytic there.

A brief aside

We have focused exclusively on use of the Taylor polynomials as polynomial approximations to a given function f on some closed interval, with suitable assumptions on existence of derivatives f', f'', ... There are other possibilities that may be appropriate in certain contexts. For example, one might want to construct a polynomial approximation which agrees with f at some specified finite set of n points (a curve-fitting problem). This requires **Lagrange interpolation** to obtain an approximating polynomial of degree $n - 1$.

Then one can use repeated applications of Rolle's Theorem on a suitably defined function — a strategy akin to that we used to prove Taylor's Theorem. This and other similar problems are taken up in the Part A course *Numerical Analysis*.

There are different versions of Taylor's Theorem valid under different technical assumptions and with the remainder term expressible in different ways. An illustration can be found on problem sheet 7.

On the positive side we record that the picture changes radically when one considers complex valued functions of a complex variable. Then condition of differentiability is much stronger, and any complex-valued function differentiable on an open disc in $\mathbb{C}$ is in fact analytic, so infinitely differentiable there. This will be covered in the Part A course *Metric Spaces and Complex Analysis*.

11 L'Hôpital's Rule

We have already indicated how the MVT and Taylor's theorem leads to useful inequalities involving the elementary functions and we have given examples of standard limits that can be obtained by basic AOL-style arguments. However, there are examples that cannot be obtained by these simple methods.

It should be apparent that what prevents us from using e.g., AOL directly to find a limit is that we encounter one of the indeterminate forms not handled by Theorem 2.2. For example, trying to find the limit of a quotient $f(x)/g(x)$ as $x \to p$, say, when the individual limits, $\lim_{x\to p} f(x)$ and $\lim_{x\to p} g(x)$ are both 0.

What we are contending with here are limits which involve what are known generically as **indeterminate forms**. They come in a variety of flavours, and our examples so far illustrate how to deal, albeit in a somewhat ad hoc way, with many of the limits that crop up frequently in practice. Can we more systematic and can we invoke theoretical tools to extend our catalogue of examples? The answer to both questions is a qualified 'Yes'.

In the remainder of this section we discuss a technique known as **L'Hôpital's Rule** (or maybe it should be referred to as L'Hôpital's Rules). It is not our intention to provide a comprehensive handbook of the various scenarios to which the L'Hôpital technique can be adapted. In any case, indeterminate limits arising in applications often require special treatment and call for ingenuity.

Let's consider first a simple case of a limit of a quotient of two functions

$$\lim_{x\to p} \frac{f(x)}{g(x)}.$$

If $f(x) \to a$ and $g(x) \to b$ with a, b finite and $b \neq 0$ then we can use AOL. We can also use Extended AOL for certain forms such as a/∞ ($a \neq \pm\infty$). Cases of $\pm\infty/b$ when

$b \neq \pm\infty$ and $a/0$ when $a \neq 0$ are guaranteed not to converge (see problem sheet 1), but what about $0/0$ or ∞/∞?

A trick that one can use when $f(x)$ and $g(x)$ are differentiable at p and $f(p) = g(p) = 0$ is use the definition of differentiability to evaluate the limit:

$$\lim_{x\to p} \frac{f(x)}{g(x)} = \lim_{x\to p} \frac{\frac{f(x)-f(p)}{x-p}}{\frac{g(x)-g(p)}{x-p}} = \frac{\lim_{x\to p} \frac{f(x)-f(p)}{x-p}}{\lim_{x\to p} \frac{g(x)-g(p)}{x-p}} = \frac{f'(p)}{g'(p)},$$

provided $g'(p) \neq 0$. We obtain the following.

Proposition 11.1 (Simple L'Hôpital Rule)**.** *Let $f, g\colon E \to \mathbb{R}$ at let $p \in E$ be a limit point of E. Assume that*

(a) $f(p) = g(p) = 0$*;*

(b) $f'(p)$ *and* $g'(p)$ *exist;*

(c) $g'(p) \neq 0$.

Then

$$\lim_{x\to p} \frac{f(x)}{g(x)} \quad \textit{exists and equals} \quad \frac{f'(p)}{g'(p)}.$$

Example 11.2. Given that the Differentiation Theorem for power series tells us that $\sin x$ is differentiable with derivative $\cos x$, we can immediately see that

$$\lim_{x\to 0} \frac{\sin x}{x} = \frac{\cos 0}{1} = 1.$$

Other examples include

$$\lim_{x\to 0} \frac{\log(1+x)}{\sin x} = \frac{1/(1+x)|_{x=0}}{\cos x|_{x=0}} = \frac{1}{1} = 1$$

and

$$\lim_{x\to 0} \frac{x^{3/2}}{\tan x} = \frac{3/2 \cdot x^{1/2}|_{x=0}}{\sec^2 x|_{x=0}} = \frac{0}{1} = 0.$$

Other indeterminate forms such as 1^∞ or ∞^0 or 0^0 can often be handled by writing out the power in terms of exp and using continuity and limit properties of the exponential function.

Example 11.3 (Euler's limit)**.** $\lim_{x\to 0}(1+x)^{1/x} = \lim_{x\to\infty}(1+\frac{1}{x})^x = e$.

Proof. $\lim_{x\to 0} \frac{\log(1+x)}{x} = \frac{(1+x)^{-1}|_{x=0}}{1} = 1$, Hence $(1+x)^{1/x} = e^{\log(1+x)/x} \to e^1 = e$ by continuity of exp at 1. For the second limit, substitute $y = 1/x$ and let $x \to \infty$. □

Suppose now we wish to evaluate the limit

$$\lim_{x\to 0} \frac{1-\cos x}{x^2},$$

say. Unfortunately, in this case $g(x) = x^2$ has derivative 0 at $x = 0$, and so we can't apply the above result. The following gives a way of proceeding in this case. However, because the proof uses the MVT, we do need much stronger conditions on both the numerator and denominator functions.

Theorem 11.4 (L'Hôpital's Rule, $\frac{0}{0}$ form)**.** *Suppose f and g are real-valued functions defined in some interval $(a, a+\delta)$, $\delta > 0$. Assume that*

(a) *f and g are differentiable in $(a, a+\delta)$;*

(b) $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$*;*

(c) *$g'(x) \neq 0$ on $(a, a+\delta)$;*

(d) $\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ *exists (in $\mathbb{R} \cup \{\pm\infty\}$).*

Then $g(x) \neq 0$ on $(a, a+\delta)$ and

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} \text{ exists and equals } \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

For the case of left-hand limits replace $(a, a+\delta)$ by $(a-\delta, a)$.
For the case of a two-sided limits replace $(a, a+\delta)$ by $(a-\delta, a+\delta) \setminus \{a\}$.

Proof. We have opted to prove the version for a right-hand limit since we can do this without the distraction of having to bother about the sign of $x-a$ when working with the Cauchy MVT. The left-hand limit version is proved likewise and the two-sided version then follows from Proposition 1.15.

So assume conditions (a)–(d) hold as set out for the right-hand limit version. By (b) we may (re)define $g(a) = f(a) = 0$ so that f and g are continuous on $[a, a+\delta)$. We also know by Rolle's Theorem for g that $g(x) = g(x) - g(a) \neq 0$ for $x \in (a, a+\delta)$ as g is continuous on $[a, x]$ and differentiable with $g' \neq 0$ on (a, x). Now apply the Cauchy MVT to obtain $\xi_x \in (a, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

Since $a < \xi_x < x$, necessarily $x \to a$ forces $\xi_x \to a$. The result now follows from (d) and Theorem 2.8. □

Remark. Usually proving (d) gives (c) as a byproduct (possibly after reducing δ). For example if we used another application of L'Hôpital to determine the limit of $f'(x)/g'(x)$. However there are situations where algebraic cancellation can occur in $f'(x)/g'(x)$ hiding a sequence of points sneakily tending to p where $g' = 0$. One can't use the theorem in this case, and indeed the conclusion can be false, so (c) does need to be checked.

Example 11.5. $\lim_{x\to 0} \frac{1-\cos x}{x^2}$. As both $1 - \cos x$ and x^2 are both differentiable and equal to zero at $x = 0$, we can apply L'Hôpital to get

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x}$$

provided the RHS limit exists and $2x \neq 0$ for $x \neq 0$ near 0. But L'Hôpital can be applied again as $\sin x$ and $2x$ are both are differentiable and equal to zero at $x = 0$. Thus

$$\lim_{x\to 0} \frac{\sin x}{2x} = \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

As this last limit exists (and $2 \neq 0$ near $x = 0$), so does the original limit (and $2x \neq 0$ for $x \neq 0$ near 0) and we finally deduce that $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$.

Note that, as in the above example, it is quite common to apply L'Hôpital more than once. However the logic is somewhat backwards. Strictly speaking we should start with the evaluation of $\lim_{x\to 0} \frac{\sin x}{2x}$, as until we know that that limit exists, we do not know the original limit exists. However it is easier to write the argument as follows with the later lines justifying the earlier ones.

First note that $1 - \cos x$ and x^2 are both infinitely differentiable, and so the derivative condition (a) in L'Hôpital holds throughout. At each stage we just need to check that numerator and denominator are both zero at $x = 0$ and that the denominator is non-zero nearby 0 (which, except at the end, is implied by the next application of L'Hôpital, and at the end is usually implied by continuity of the non-zero denominator). So

$$\begin{aligned}
\lim_{x\to 0} \frac{1-\cos x}{x^2} &= \lim_{x\to 0} \frac{\sin x}{2x} && \text{L'Hôpital } \tfrac{0}{0}\text{, provided RHS exists and denom} \neq 0 \text{ near } 0 \\
&= \lim_{x\to 0} \frac{\cos x}{2} && \text{L'Hôpital } \tfrac{0}{0}\text{, provided RHS exists and denom} \neq 0 \text{ near } 0 \\
&= \frac{1}{2} && \text{Continuity of cos and AOL, and yes, } 2 \neq 0 \text{ near } 0
\end{aligned}$$

Example 11.6. $\lim_{x\to 0} \frac{\sin x - x}{\sinh^3 x}$. Applying the method in the previous example we get

$$\begin{aligned}
\lim_{x\to 0} \frac{\sin x - x}{\sinh^3 x} &= \lim_{x\to 0} \frac{\cos x - 1}{3\sinh^2 x \cosh x} && \text{L'H } \tfrac{0}{0}\text{, provided...} \\
&= \lim_{x\to 0} \frac{-\sin x}{6\sinh x \cosh^2 x + 3\sinh^3 x} && \text{L'H } \tfrac{0}{0}\text{, provided...} \\
&= \lim_{x\to 0} \frac{-\cos x}{6\cosh^3 x + 12\sinh^2 x \cosh x + 9\sinh^2 x \cosh x} && \text{L'H } \tfrac{0}{0}\text{, provided...} \\
&= \frac{-1}{6} && \text{Continuity + AOL}
\end{aligned}$$

Note however that the differentiation was beginning to get rather tedious. Indeed, one should avoid just simply applying L'Hôpital multiple times without thought. Often the calculations can be simplified by combining with AOL or other techniques. For example:

$$\begin{aligned}
\lim_{x\to 0}\frac{\sin x - x}{\sinh^3 x} &= \lim_{x\to 0}\frac{\cos x - 1}{3\sinh^2 x\cosh x} && \text{L'H } \tfrac{0}{0}\text{, provided...}\\
&= \lim_{x\to 0}\frac{1}{3\cosh x}\cdot\lim_{x\to 0}\frac{\cos x - 1}{\sinh^2 x} && \text{AOL}\\
&= \frac{1}{3}\lim_{x\to 0}\frac{-\sin x}{2\sinh x\cosh x} && \text{L'H } \tfrac{0}{0}\text{, provided...}\\
&= \frac{1}{3}\lim_{x\to 0}\frac{-1}{2\cosh x}\cdot\lim_{x\to 0}\frac{\sin x}{\sinh x} && \text{AOL}\\
&= \frac{-1}{6}\lim_{x\to 0}\frac{\cos x}{\cosh x} && \text{L'H } \tfrac{0}{0}\text{, provided...}\\
&= -\frac{1}{6} && \text{Continuity}
\end{aligned}$$

Again the justification is that each line holds provided the RHS limits exist and the denominator is non-zero nearby $x = 0$, and thus the last line inductively justifies all the previous ones. One needs to be a bit more careful that the factors we are taking out are not hiding a sequence of zeros on the denominator, causing (c) to fail.

Of course, it is sometimes just easier to use Taylor's Theorem:

$$\begin{aligned}
\lim_{x\to 0}\frac{\sin x - x}{\sinh^3 x} &= \lim_{x\to 0}\frac{(x - \frac{1}{3!}x^3 + O(x^5)) - x}{(x + O(x^3))^3}\\
&= \lim_{x\to 0}\frac{-\frac{1}{6}x^3 + O(x^5)}{(x(1 + O(x^2)))^3}\\
&= \lim_{x\to 0}\frac{-\frac{1}{6} + O(x^2)}{(1 + O(x^2))^3}\\
&= -\frac{1}{6}.
\end{aligned}$$

Again, we emphasise that one should be on the lookout for AOL and other methods to simplify things, rather than just applying L'Hôpital multiple times on autopilot. For example,

$$\lim_{x\to 0}\frac{\sin^3 x}{x^3 + x^4} = \lim_{x\to 0}\frac{1}{1+x}\cdot\left(\lim_{x\to 0}\frac{\sin x}{x}\right)^3 = 1\cdot 1^3 = 1.$$

Extensions

One can extend L'Hôpital's rule to the case when the limit is as $x \to \pm\infty$ fairly easily by replacing x with $1/x$ (see problem sheet 8). One can also extend L'Hôpital's rule to the case when $f(x), g(x) \to \pm\infty$ as $x \to a$, although this requires a bit more work.

Theorem 11.7 (L'Hôpital's Rule, $\frac{\infty}{\infty}$ form)**.** *Suppose f and g are real-valued functions defined in some interval $(a, a+\delta)$, $\delta > 0$. Assume that*

(a) *f and g are differentiable in $(a, a+\delta)$;*

(b) $\lim_{x\to a^+} |f(x)| = \lim_{x\to a^+} |g(x)| = \infty$;

(c) $g'(x) \neq 0$ *on* $(a, a+\delta)$.

(d) $\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ *exists.*

Then

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} \text{ exists and equals } \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

For the case of left-hand limits replace $(a, a+\delta)$ *by* $(a-\delta, a)$.
For the case of a two-sided limits replace $(a, a+\delta)$ *by* $(a-\delta, a+\delta) \setminus \{a\}$.

Remark. We can't just replace $\frac{f}{g}$ with $\frac{1/g}{1/f}$ and apply the $\frac{0}{0}$ form. (Why?)

Proof. We will just prove the right-hand limit version, the other cases follow quite easily. So assume conditions (a)–(d) hold as set out for the right-hand limit version. We also know by Rolle's Theorem for g that $g(x) - g(c) \neq 0$ for $a < x < c < a+\delta$. Apply the Cauchy MVT to obtain $\xi_{x,c} \in (x, c)$ such that

$$\frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f'(\xi_{x,c})}{g'(\xi_{x,c})}.$$

Now if $\frac{f'(x)}{g'(x)} \to \ell \in \mathbb{R}$ as $x \to a^+$ we can't deduce that $\xi_{x,c}$ converges (as it is only restricted to be between a and c). However, given $\varepsilon > 0$ we can find a $\delta' \in (0, \delta)$ such that

$$\Big|\frac{f(x) - f(c)}{g(x) - g(c)} - \ell\Big| = \Big|\frac{f'(\xi_{x,c})}{g'(\xi_{x,c})} - \ell\Big| < \varepsilon \tag{17}$$

for all $a < x < c < a+\delta'$ (as then $a < \xi_{x,c} < a+\delta'$). We want $|\frac{f(x)}{g(x)} - \ell|$ small so we need to do some algebraic manipulation on (17). Clearing the fraction in (17) gives

$$|f(x) - f(c) - \ell g(x) + \ell g(c)| < \varepsilon |g(x) - g(c)|,$$

so by the triangle inequality

$$|f(x) - \ell g(x)| < \varepsilon |g(x) - g(c)| + |f(c) - \ell g(c)|.$$

Hence

$$\Big|\frac{f(x)}{g(x)} - \ell\Big| < \varepsilon\Big|1 - \frac{g(c)}{g(x)}\Big| + \Big|\frac{f(c) - \ell g(c)}{g(x)}\Big|. \tag{18}$$

Now fix c and let $x \to a^+$. As $|g(x)| \to \infty$ we see the RHS of (18) tends to $\varepsilon \cdot 1 + 0 = \varepsilon$ as $x \to a^+$. Thus for x sufficiently close to a we have

$$\Big|\frac{f(x)}{g(x)} - \ell\Big| < 2\varepsilon.$$

As this holds for any $\varepsilon > 0$, $\frac{f(x)}{g(x)} \to \ell$. Similar (easier) arguments apply when $\ell = \pm\infty$. $\square$

12 The Binomial Expansion

By simple induction we can prove that for any natural number n (including 0) we have for all real or complex x that

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k = \sum_{k=0}^{\infty} \binom{n}{k} x^k,$$

where the coefficient $\binom{n}{k}$ of x^k can be proved to be

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)\cdots 1} \qquad (= 0 \text{ if } k > n).$$

We want to extend this result. We have also seen in our work on sequences and series that

$$(1+x)^{-1} = \sum_{k=0}^{\infty} (-1)^k x^k \qquad \text{for all } |x| < 1$$

and here the coefficient of x^k can be written as

$$(-1)^k = \frac{(-1)(-2)\cdots(-k)}{k(k-1)\cdots 1},$$

and we can prove by induction (for example using differentiation term by term) that for all $n \in \mathbb{N}$ we have that

$$(1+x)^{-n} = \sum_{k=0}^{\infty} \frac{(-n)(-n-1)\cdots(-n-k+1)}{k(k-1)\cdots 1} x^k \qquad \text{for all } |x| < 1,$$

so the binomial theorem above holds for all *integers* n if we define

$$\binom{n}{k} := \frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)\cdots 1}.$$

We are going to generalise this — in the case of some real values of x — to all values of n, not just integers. Note that this is altogether deeper: $(1+x)^p$ is defined for non-integral p, and for (real) $x > -1$, to be the function $\exp(p \log(1+x))$.

Definition. For all $p \in \mathbb{R}$ and all $k \in \mathbb{N} \cup \{0\}$ we extend the definition of the **binomial coefficient** as follows:

$$\binom{p}{k} := \frac{p(p-1)(p-2)\cdots(p-k+1)}{k!},$$

where we interpret the empty product as 1 when $k = 0$.

We now make sure that the key properties of binomial coefficients are still true in this more general setting.

Lemma 12.1. *For all $k \geq 1$ and all $p \in \mathbb{R}$*

$$\binom{p}{k} = \frac{p}{k}\binom{p-1}{k-1} = \frac{p-k+1}{k}\binom{p}{k-1} \qquad \text{and} \qquad \binom{p+1}{k} = \binom{p}{k} + \binom{p}{k-1}.$$

Proof. The first claim is clear by taking out a factor of $\frac{p}{k}$ or $\frac{p-k+1}{k}$ in the definition of $\binom{p}{k}$.

For the second we use the first claim (both parts) to show that

$$\binom{p}{k} + \binom{p}{k-1} = \frac{p-k+1}{k}\binom{p}{k-1} + \binom{p}{k-1} = \frac{p+1}{k}\binom{p}{k-1} = \binom{p+1}{k}. \qquad \square$$

Theorem 12.2 (Real Binomial Theorem)**.** *Let p be a real number. Then for all $|x| < 1$*

$$(1+x)^p = \sum_{k=0}^{\infty} \binom{p}{k} x^k.$$

Note that the coefficients are all non-zero provided p is not a natural number or zero; as we have a proof of the expansion in that case we may assume that $p \notin \mathbb{N} \cup \{0\}$.

Lemma 12.3. *The function f defined on $(-1,1)$ by $f(x) := (1+x)^p$ is differentiable, and satisfies $(1+x)f'(x) = pf(x)$. Also, $f(0) = 1$.*

Proof. The derivative is easily obtained by the chain rule from the definition of f; it is $f'(x) = p(1+x)^{p-1}$. Multiply by $(1+x)$ and get the required relationship. The value at 0 is clear. $\square$

Lemma 12.4. *The radius of convergence of $\sum \binom{p}{k} x^k$ is $R = 1$.*

Proof. Use the Ratio Test; we have that

$$\left|\frac{\binom{p}{k}x^k}{\binom{p}{k-1}x^{k-1}}\right| = \left|\frac{p-k+1}{k} \cdot x\right| \to |(-1) \cdot x| = |x|$$

as $k \to \infty$. This is < 1 iff $|x| < 1$. $\square$

Lemma 12.5. *The function g defined on $(-1,1)$ by $g(x) := \sum_{k=0}^{\infty} \binom{p}{k} x^k$ is differentiable, with derivative satisfying $(1+x)g'(x) = pg(x)$. Also, $g(0) = 1$.*

Proof. We have

$$\begin{aligned}
(1+x)g'(x) &= (1+x)\sum_{k=1}^{\infty}\binom{p}{k}kx^{k-1} && \text{Diff. of power series, } |x|<1\\
&= \sum_{k=1}^{\infty}\binom{p}{k}kx^{k-1} + \sum_{k=1}^{\infty}\binom{p}{k}kx^{k}\\
&= \sum_{k=0}^{\infty}\binom{p}{k+1}(k+1)x^{k} + \sum_{k=1}^{\infty}\binom{p}{k}kx^{k} && k\mapsto k+1 \text{ in } 1^{\text{st}} \text{ sum}\\
&= \sum_{k=0}^{\infty}\binom{p}{k}(p-k)x^{k} + \sum_{k=0}^{\infty}\binom{p}{k}kx^{k} && \tbinom{p}{k+1} = \tfrac{p-(k+1)+1}{k+1}\tbinom{p}{k}\\
&= p\sum_{k=0}^{\infty}\binom{p}{k}x^{k} = pg(x). && \square
\end{aligned}$$

Proof of the binomial theorem. Consider $F(x) = g(x)/f(x)$, which is well-defined on $(-1,1)$ as $f(x) > 0$. By the Quotient Rule we can calculate $F'(x)$, and then use the lemmas:

$$F'(x) = \frac{f(x)g'(x) - f'(x)g(x)}{f(x)^2} = \frac{p}{1+x}\,\frac{f(x)g(x) - f(x)g(x)}{f(x)^2} = 0.$$

Hence by the Constancy Theorem, $F(x)$ is constant, $F(x) = F(0) = 1$. This implies that $f(x) = g(x)$ on $(-1,1)$. □

Binomial Theorem at the end points

The existence of these functions and their equality at the end points $x = \pm 1$ requires more sophisticated argument. The following should be viewed as illustrations of the way various theorems can be exploited, rather than proofs to be learnt.

As we will be considering sums $\sum \binom{p}{n}x^n$ with $x = \pm 1$, it helps to first estimate how large the binomial coefficient $\binom{p}{n}$ is.

Lemma 12.6. *For any $p \in \mathbb{R}$ we have $\binom{p}{n} = O(n^{-(p+1)})$ as $n \to \infty$.*

Proof. We first note that

$$\binom{p}{n} = \frac{p(p-1)\cdots(p-n+1)}{n(n-1)\cdots 1} = \pm\Big(1 - \frac{p+1}{1}\Big)\cdots\Big(1 - \frac{p+1}{n}\Big).$$

Now for $x \in [0,1]$ we have $1 - x \le e^{-x}$ as $e^{-x} + x - 1$ has positive derivative for $x > 0$ and is 0 at $x = 0$. Let $s \in \mathbb{N}$ be fixed so that $s > p+1$, say $s = \max(\lfloor p+2 \rfloor, 1)$. Then

$$\Big|\binom{p}{n}\Big| \le \prod_{k=1}^{s-1}\Big|1 - \frac{p+1}{k}\Big| \cdot \prod_{k=s}^{n} e^{-(p+1)/k} = C\exp\Big(-(p+1)\sum_{k=s}^{n}\frac{1}{k}\Big),$$

where C is a constant just depending on p (and s). But from *Analysis I* we know that $\sum_{k=1}^{n}\frac{1}{n} - \log n \to \gamma$ as $n \to \infty$, so in particular $|\sum_{k=s}^{n}\frac{1}{k} - \log n|$ is bounded as $n \to \infty$ (with s fixed). Thus as $\exp x$ is increasing in x we can bound

$$\left|\binom{p}{n}\right| \leq C\exp\big(-(p+1)\log n + C'\big) = C''n^{-(p+1)},$$

for suitable constants C' and C''. □

The case when $x = 1$

Theorem 12.7. *For any $p > -1$ the series $\sum_{n=0}^{\infty}\binom{p}{n}$ is convergent with sum 2^p.*

Remark. It is easy to see that for $p \leq -1$, $\binom{p}{n} \not\to 0$, so the series is divergent.

Proof. We apply Taylor's Theorem to $(1+x)^p$ on the interval $[0,1]$ (with n replaced by $n-1$ for convenience). We have $\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}(1+x)^p = \binom{p}{k}(1+x)^{p-k}$, so for each $n \geq 1$, there is a point $\xi_n \in (0,1)$ such that

$$2^p = \sum_{k=0}^{n-1}\binom{p}{k} + E_n, \qquad \text{where} \qquad E_n = \binom{p}{n}(1+\xi_n)^{p-n}.$$

Hence for $n > p$ we have $|E_n| \leq |\binom{p}{n}|$. But then by Lemma 12.6, $|E_n| = O(n^{-(p+1)})$, and so $E_n \to 0$ as $n \to \infty$ since $p + 1 > 0$. □

Remark. In the above proof we could not make use of the $(1+\xi_n)^{p-n}$ factor to show E_n is small as we could have ξ_n tending very rapidly to 0 as $n \to \infty$.

The case when $x = -1$

For $x = -1$ we have not yet defined $(1+x)^p = 0^p$.

For $p \in \mathbb{N}$ we have the usual algebraic definition, so $0^p = 0$ for $p \geq 1$. Can we define 0^p sensibly for any other values of p?

For $p > 0$: when $x > 0$ we defined $x^p := \exp(p\log x)$. As $\log x \to -\infty$ as $x \to 0^+$, we have $\exp(p\log x) \to 0$ as $x \to 0^+$. Thus to make x^p continuous at $x = 0$ we should define $0^p = 0$ for all $p > 0$. This we now do.

If $p = 0$: one normally defines $0^0 = 1$, although some authors prefer 0^0 to be left undefined. Having $0^0 = 1$ certainly makes sense for polynomials and power series $\sum c_k x^k$ when $x = 0$, and is always interpreted this way in this context. The disadvantage of defining 0^0 is that x^y cannot be made continuous at $(x,y) = (0,0)$, so one cannot assume $f(x)^{g(x)}$ converges when $f(x), g(x) \to 0$.

Theorem 12.8. *For any $p > 0$ the series $\sum_{n=0}^{\infty}\binom{p}{n}(-1)^n$ is convergent with sum 0.*

Remark. For $p = 0$ the sum is 1, and for $p < 0$ it is easy to show that the sum diverges.

Proof. In this case, Taylor's theorem does not help. But we can get the result by showing the binomial series is uniformly convergent, and hence continuous, on $[-1, 1]$.

We have $|\binom{p}{n}x^n| \le M_n := |\binom{p}{n}|$, for all $x \in [-1, 1]$. But by Lemma 12.6, $M_n = O(n^{-(p+1)})$ and $\sum n^{-(p+1)}$ converges for $p > 0$ by the Integral Test. Thus by the Comparison Test, $\sum M_n$ converges and so we have uniform convergence of the series $\sum \binom{p}{n}x^n$ on $[-1, 1]$ by the M-test. As each of the terms $\binom{p}{n}x^n$ is continuous in x, this implies the infinite sum is continuous for all $x \in [-1, 1]$. Continuity at $x = -1$ then implies that

$$\sum_{n=0}^{\infty} \binom{p}{n}(-1)^n = \lim_{x \to (-1)^+} (1 + x)^p = \lim_{x \to (-1)^+} e^{p \log(1+x)} = \lim_{y \to -\infty} e^{py} = 0$$

as $y = \log(1 + x) \to -\infty$ as $x \to (-1)^+$, $p > 0$, and $\exp(z) \to 0$ as $z \to -\infty$. □

Continuity of a real power series at the endpoints

You seen in *Analysis I* that a real power series $\sum c_k x^k$ with finite non-zero radius of convergence R converges absolutely for any x for which $|x| < R$. You also saw examples which show that the series may converge absolutely, may convergence non-absolutely, or may diverge, at each of the points $x = R$ and $x = -R$.

We showed in Section 7 that $f(x) := \sum_{k=0}^{\infty} c_k x^k$ defines a continuous function f on $(-R, R)$, irrespective of how the series behaves at $\pm R$. But what if the series does converge at $\pm R$? Can we deduce that the value is what one would expect assuming f is continuous there? In the examples we have seen it did, and indeed, the answer turns out to be Yes!

By replacing $f(x)$ with $f(\pm x/R)$ we may assume without loss of generality that $R = 1$ and we are interested in the series at $x = R = 1$. The following is then the result that we want.

Theorem 12.9 (Abel's Continuity Theorem)**.** *Assume that* $\sum c_k$ *converges. Then* $\sum c_k x^k$ *converges uniformly on* $[0, 1]$*. In particular* $\sum_{k=0}^{\infty} c_k x^k$ *is continuous on* $[0, 1]$ *and*

$$\lim_{x \to 1^-} \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{\infty} c_k.$$

Remark. We note that uniform convergence of $\sum c_k x^k$ follows immediately from the M-test when $\sum |c_k|$ converges, so the interesting case is when $\sum c_k$ is not absolutely convergent.

Proof. Fix $\varepsilon > 0$. Then by the Cauchy Convergence Criterion for series, there is an N such that for $n \ge m > N$,

$$\Big|\sum_{k=m}^{n} c_k\Big| < \varepsilon.$$

Now fix $m > N$ and define $S_n = \sum_{k=m}^{n} c_k$ for $n \geq m-1$ with the convention that $S_{m-1} = 0$. We note that $c_n = S_n - S_{n-1}$ for all $n \geq m$. Thus[37]

$$\begin{aligned}
\sum_{k=m}^{n} c_k x^k &= \sum_{k=m}^{n} S_k x^k - \sum_{k=m}^{n} S_{k-1} x^k && c_k = S_k - S_{k-1} \\
&= \sum_{k=m}^{n} S_k x^k - \sum_{k=m-1}^{n-1} S_k x^{k+1} && k \mapsto k+1 \text{ in 2nd sum} \\
&= \sum_{k=m}^{n-1} S_k (x^k - x^{k+1}) + S_n x^n && \text{combine terms noting } S_{m-1} = 0
\end{aligned}$$

Hence by the Triangle inequality, and noting that $|S_n| < \varepsilon$ for $n \geq m$ and $x^k - x^{k+1} \geq 0$,

$$\Big|\sum_{k=m}^{n} c_k x^k\Big| \leq \sum_{k=m}^{n-1} \varepsilon(x^k - x^{k+1}) + \varepsilon x^n = \varepsilon x^m \leq \varepsilon$$

for any $x \in [0,1]$. Thus by Cauchy's Criterion for uniform convergence of series, Corollary 7.10, we have that $\sum c_k x^k$ is uniformly convergent on $[0,1]$.

Continuity of $\sum c_k x^k$ and the limit as $x \to 1^-$ now follow from Theorem 7.2. □

Example 12.10. (Recall Example 10.3.) We have (by the Differentiation and Constancy theorems) that for $x \in (-1,1)$,

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k}.$$

As $\log(1+x)$ is continuous at $x = 1$ and $\sum \frac{(-1)^{k-1}}{k}$ converges by the Alternating Series Test, we deduce that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

Warning. Abel's Theorem only applies in situations where the sum is a genuine power series of the form $\sum c_k x^k$. For example, recall that for $x \in (-1,1)$

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{n}.$$

Now consider for $x \in [0,1]$ the series

$$f(x) := \sum_{k=1}^{\infty} \frac{x^k - x^{2k}}{n}.$$

Then clearly for $x \in (-1,1)$,

$$f(x) = -\log(1-x) + \log(1-x^2) = \log \tfrac{1-x^2}{1-x} = \log(1+x).$$

But the series for $f(x)$ converges at $x = 1$ and $f(1) = \sum 0 = 0$. But $f(1^-) = \lim_{x\to 1^-} \log(1+x) = \log 2 \neq 0$.

[37]This is called Abel's summation formula – think 'integration by parts'.