MT2502 Analysis

Mike Todd

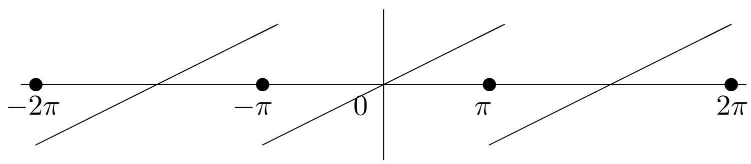
2015

0.1 What is Analysis?

Many mathematical objects, such as derivative and integral, and even the real numbers themselves, are best (most rigorously) defined as the limit of an infinite process. The problem is to understand what it means to find the limit of an infinite process, since we can never really go to the 'end' of an infinite process. For example, the sine function is a nice continuous (a notion we'll explore more later) function, so you might expect the function

$$S(x) = \frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots$$

to be continuous, but in fact it's not:



This particular example, now part of Fourier analysis (see MT2507) arose from the study of heat conduction (18th/19th century): at the time it wasn't even considered to be a genuine function. In fact many of the advances in what now would be thought of as applied maths, from calculus onwards, were originally made without the logical underpinning of analysis, which led to confusion and paradox in the 18th/19th century. Analysis now provides a framework for these problems and has since blossomed into a beautiful and useful theory.

Chapter 1

The rationals and the reals

The real number system is a very intuitive idea and has been used since at least the ancient Greek times. However, while ostensibly higher-powered notions like derivative and integral (1820s) use this intuitive notion, a rigorous formal definition of the real numbers came much later (1860s-early 20th century).

In this section, we'll discuss the construction of the reals using precise definitions, logic, set theory and the rational number system (although we'll have to leave out some of the construction in the interests of time).

First recall that the set of rational numbers $\mathbb{Q}$ is defined as the collection of all fractions of the form $\frac{p}{q}$ where $p \in \mathbb{Z}$ and $q \in \mathbb{N}$.

The first result is that adding/subtracting/multiplying two rational numbers yields a rational number: $\mathbb{Q}$ is closed under the usual algebraic operations.

Proposition 1.1. *If $a, b \in \mathbb{Q}$, then $a+b, a-b, ab \in \mathbb{Q}$. Moreover, if $b \neq 0$ then $\frac{a}{b} \in \mathbb{Q}$.*

The proof is an exercise.

While the rationals are a nice set from many points of view, they are insufficient for even the most basic mathematics. Eg a right-angled triangle with sides of length 1 and 1, by Pythagoras' Theorem has hypotenuse of length $\sqrt{2}$. But $\sqrt{2}$, which by definition is the positive solution x to $x^2 = 2$, isn't rational:

Proposition 1.2. *There is no rational number x such that $x^2 = 2$.*

Proof. Suppose that, on the contrary, there is some $x = \frac{p}{q}$ where $p, q \in \mathbb{N}$ such that $x^2 = 2$. First notice that if p and q have a common divisor (an integer which divides both) then we can factorise this out. Hence w.m.a. p and q have no common divisors.

Since $x^2 = 2$,

$$\frac{p^2}{q^2} = 2 \Rightarrow p^2 = 2q^2, \tag{1.0.1}$$

we know that p must be even (note that the only way a square of a number can be even is if the number itself is even). So p is even, so there must exist $n \in \mathbb{N}$ such that $p = 2n$. Hence $p^2 = (2n)^2 = 2q^2 \Rightarrow 2n^2 = q^2$, so similarly q is even. So p and q have a common divisor of 2, contradicting our assumption. Hence the proposition must be true. $\square$

This proposition implies that there are 'holes' in the rational numbers. In the same vein, the following example shows that while you can take a sequence of rational numbers which intuitively converge, they need not converge to a rational number.

Example 1.1. *Let $x_1 = 2$ and for $n \geqslant 1$, let*

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

(i.e., $x_1 = 2, x_2 = \frac{2}{2} + \frac{1}{2} = \frac{3}{2}, x_3 = \frac{\left(\frac{3}{2}\right)}{2} + \frac{1}{\left(\frac{3}{2}\right)} = \frac{3}{4} + \frac{2}{3} = \frac{17}{12}$.)

Claim. *The sequence is decreasing, i.e., $x_1 \geqslant x_2 \geqslant x_3 \geqslant \cdots$.*

Proof. We'll show that $x_n - x_{n+1} \geqslant 0$ for any n.

First, since $x_1 = 2 \geqslant \frac{3}{2} = x_2$, the claim is proved for $n = 1$. For $n \geqslant 2$:

$$\begin{aligned}
x_n - x_{n+1} &= x_n - \left(\frac{x_n}{2} + \frac{1}{x_n}\right) = \frac{x_n}{2} - \frac{1}{x_n} = \frac{1}{2x_n}(x_n^2 - 2) \\
&= \frac{1}{2x_n}\left(\left(\frac{x_{n-1}}{2} + \frac{1}{x_{n-1}}\right)^2 - 2\right) = \frac{1}{2x_n}\left(\frac{x_{n-1}^2}{4} + 1 + \frac{1}{x_{n-1}^2} - 2\right) \\
&= \frac{1}{2x_n}\left(\frac{x_{n-1}^4 + 4 - 4x_{n-1}^2}{4x_{n-1}^2}\right) = \frac{1}{2x_n}\left(\frac{(x_{n-1}^2 - 2)^2}{4x_{n-1}^2}\right) \geqslant 0.
\end{aligned}$$

□

The sequence is also bounded:

$$0 \leqslant x_n \leqslant x_1 = 2 \qquad \forall n \in \mathbb{N}.$$

Suppose now that, contrary to the statement of the proposition, $(x_n)_n$ does converge to a number $x \in \mathbb{Q}$. But since $x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$, any such limit must satisfy

$$x = \frac{x}{2} + \frac{1}{x} \Rightarrow \frac{x}{2} = \frac{1}{x} \Rightarrow x^2 = 2,$$

which by Proposition 1.2 is impossible if $x \in \mathbb{Q}$.

It is strange that a bounded decreasing sequence in $\mathbb{Q}$ doesn't converge in $\mathbb{Q}$. For such reasons we need to extend our set of numbers.

1.1 Ordering and bounding

Before defining the reals, we'll need some abstract, but widely applicable, definitions.

Definition 1.1. *An* ordered set $(X, <)$ *consists of a* set X *and a* relation $<$ *on X such that*

1. *The trichotomy law holds: exactly one of the following is true:*

$$x < y, \quad x = y, \quad y < x.$$

2. *The transitivity law holds: if $x, y, z \in X$ and $x < y$ and $y < z$ then $x < z$.*

We'll also use the notation $x \leqslant y$ which means that $x < y$ or $x = y$; $x > y$, which means $y < x$; and $x \geqslant y$ which means $y \leqslant x$.

Ordered sets can be very abstract, but a fairly concrete example is $(\mathbb{Q}, <)$, i.e., the rational numbers with the usual ordering $<$.

Definition 1.2. *Given an ordered set $(X, <)$ and $A \subseteq X$,*

1. *$u \in X$ is called an* upper bound *for A if*

$$\forall a \in A, \quad a \leqslant u.$$

   If there is an upper bound, we say that A is bounded above

2. *$\ell \in X$ is called an* lower bound *for A if*

$$\forall a \in A, \quad \ell \leqslant a.$$

   If there is a lower bound, we say that A is bounded below*.*

If A is bounded both above and below, we say that A is bounded*.*

N.B. There may be lots of upper/lower bounds for given sets A, which contrasts with the following notions.

Definition 1.3. *Let $(X, <)$ be an ordered set and $A \subseteq X$.*

1. *An element $M \in A$ is called a* maximum *for A if M is an upper bound for A, i.e., we require*

$$\forall a \in A, \quad a \leqslant M \text{ and } M \in A.$$

2. *An element $m \in A$ is called a* minimum *for A if m is an lower bound for A, i.e., we require*

$$\forall a \in A, \quad m \leqslant a \text{ and } m \in A.$$

Note that there can be at most one maximum for A: suppose we had two, say M_1 and M_2. Then by definition,

$$M_2 \leqslant M_1 \text{ and } M_1 \leqslant M_2, \text{ so } M_1 = M_2.$$

Similarly, there's at most one minimum.

Assuming it exists, we denote the unique maximum by

$$\max A, \quad \max_{x \in A} x, \text{ or } \max\{x \in A\},$$

and similarly the unique minimum by

$$\min A, \quad \min_{x \in A} x, \text{ or } \min\{x \in A\}.$$

Problem: Maxima/minima may not exist. Eg, set

$$I := \{q \in \mathbb{Q} : 0 < q \leqslant 1\}.$$

This set is bounded above (eg by 70) and below (eg by 0), but while it has a maximum (i.e., 1), it has no minimum. To cope with this kind of issue, define:

Definition 1.4. *Let $(X, <)$ be an ordered set and $A \subseteq X$ a non-empty subset of X.*

1. *Let $U(A)$ denote the set of upper bounds for A. Then an element $u \in U(A)$ is called a* least upper bound/supremum *for A if*

$$\forall v \in U(A), \quad u \leqslant v.$$

2. *Let $L(A)$ denote the set of lower bounds for A. Then an element $\ell \in L(A)$ is called a* greatest lower bound/infimum *for A if*

$$\forall m \in L(A), \quad m \leqslant \ell.$$

There can be at most one supremum (Exercise), so we denote this, if it exists, by

$$\sup A, \quad \sup_{x \in A} x, \quad \sup\{x \in A\}.$$

Similarly for infimum,

$$\inf A, \quad \inf_{x \in A} x, \quad \inf\{x \in A\}.$$

Lemma 1.3. *For A a non-empty subset of $\mathbb{Q}$:*

1. *if A has a maximum, then $\max A = \sup A$;*
2. *if A has a minimum, then $\min A = \inf A$.*

Proof. See tutorial sheet. $\square$

Example 1.2. *Recall the set $I := \{x \in \mathbb{Q} : 0 < x \leqslant 1\}$. Since I has a maximum, 1, then it shares the same supremum. However, while it doesn't have a minimum, $\inf I = 0$ (first 0 is clearly a lower bound; suppose that $\ell > 0$ is an infimum: then for large enough n, $\frac{1}{n} < \ell$, so since $\frac{1}{n} \in I$, ℓ isn't a lower bound for I).*

Example 1.3. *Given the ordered set $(\mathbb{Q}, <)$, define*

$$K := \{q \in \mathbb{Q} : 0 < q, q^2 < 2\}.$$

Clearly this set is bounded. However,

Proposition 1.4. *K has no supremum.*

Proof. Suppose that the proposition is false, so there exists a least upper bound $x \in \mathbb{Q}$ for K. We'll show that this is impossible.

Claim. $2 < x^2$.

Proof of claim. Suppose that the claim is false, so $x^2 \leqslant 2$. Since $x \in \mathbb{Q}$, Proposition 1.2 implies that $x^2 \neq 2$, so in fact $x^2 < 2$. Therefore, $\frac{2-x^2}{2x+1} > 0$. So choosing $n \in \mathbb{N}$ large enough, we can ensure that $\frac{2-x^2}{2x+1} > \frac{1}{n}$ (*). Then

$$\begin{aligned}\left(x + \frac{1}{n}\right)^2 &= x^2 + \frac{2x}{n} + \frac{1}{n^2} = x^2 + \frac{1}{n}\left(2x + \frac{1}{n}\right) \\ &\leqslant x^2 + \frac{1}{n}(2x+1) < x^2 + (2 - x^2) \text{ (by (*))} \\ &= 2\end{aligned}$$

So $x + \frac{1}{n}$ is a rational number whose square is < 2 and hence in K, so x can't be an upper bound on K, a contradiction. $\square$

The claim implies that $x^2 > 2$, so $\frac{x^2-2}{2x} > 0$ and so we can choose $m \in \mathbb{N}$ so that $\frac{x^2-2}{2x} > \frac{1}{m}$, which rearranges to

$$x^2 - \frac{2x}{m} > 2 \ (**).$$

Set $y := x - \frac{1}{m} < x$. Then

$$y^2 = x^2 - \frac{2x}{m} + \frac{1}{m^2} > x^2 - \frac{2x}{m} > 2$$

by (**). Hence $y < x$ is an upper bound for K, contradicting x being a *least* upper bound.

Adding this all together, the proposition is true. □

We want to work in a set of numbers which has no such 'holes', i.e., all bounded subsets have supremum/infimum.

Definition 1.5. *An ordered set* $(X, <)$ *is called* complete *if every non-empty set which is bounded above has a supremum.*

Note that $(\mathbb{Q}, <)$ is not complete by Proposition 1.4.

In the following result we'll use the notion of 'field' see eg Section 1.4 of Howie.

Theorem 1.5. *There exists an ordered field denoted* $(\mathbb{R}, <)$ *such that*

i) $\mathbb{Q} \subseteq \mathbb{R}$

ii) $(\mathbb{R}, <)$ *is complete.*

We omit the proof of this theorem, as well as the theorem which states that $(\mathbb{R}, <)$ is essentially unique.

The set above is called the *real numbers* $\mathbb{R}$. Note that elements of $\mathbb{R} \setminus \mathbb{Q}$ are called *irrational numbers.*

1.2 Absolute Value

(Note that this short section doesn't particularly fit into this chapter, but it'll be useful later.)

Definition 1.6. *Given* $x \in \mathbb{R}$*, the* absolute value *of* x*, denoted* $|x|$*, is defined as*

$$|x| := \begin{cases} -x & \text{if } x < 0, \\ x & \text{if } x \geqslant 0. \end{cases}$$

Theorem 1.6. *Given* $x, y \in \mathbb{R}$ *and* $a \geqslant 0$*,*

i) $|x| \geqslant 0$*;*

ii) $|-x| = |x|$*;*

iii) $|x| \leqslant a$ *iff* $-a \leqslant x \leqslant a$*;*

iv) $|xy| = |x||y|$*;*

v) (Triangle inequality) $|x + y| \leqslant |x| + |y|$*;*

vi) (Reverse triangle inequality) $\Big||x| - |y|\Big| \leqslant |x - y|$.

Proof. i)-iv): Exercise.

v): Since $-|x| \leqslant x \leqslant |x|$ and $-|y| \leqslant y \leqslant |y|$, summing:

$$-(|x| + |y|) \leqslant x + y \leqslant |x| + |y|,$$

so applying iii), we obtain v).

vi) Exercise. □

N.B. Absolute value is often used as a way of finding the distance between two real numbers $x, y \in \mathbb{R}$: i.e., this is $|x - y|$.

Chapter 2

Sequences

Now that we've laid a solid foundation for the real number system we can begin to address more 'analytic' issues; the first of these being sequences.

2.1 Sequences and convergence

Definition 2.1. *A sequence of real numbers is a function*

$$f : \mathbb{N} \to \mathbb{R}$$

going from the rationals to the reals. Usually we'll denote $f(n)$ by x_n and write the sequence as

$$(x_n)_n, \quad (x_n)_{n\in\mathbb{N}}, \quad (x_n)_{n=1}^{\infty}, \quad (x_1, x_2, \ldots).$$

Note that in contrast to set notation $\{\quad\}$, the order *is* important here, eg $(-1, 1, 1, -1, \ldots)$ means something different to $\{-1, 1, 1, -1, \ldots\} = \{-1, 1\}$. Moreover, $(1, 2, 3, 4, 5, \ldots) \neq (2, 1, 3, 4, 5, \ldots)$.

Sometimes we have a nice formula for the n-th term of a sequence, eg

$$(2, 4, 6, 8, \ldots) = (2n)_n \text{ i.e., } x_n = 2n; \quad \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right) = \left(\frac{1}{n}\right)_n \text{ i.e., } y_n = \frac{1}{n}.$$

Example 2.1.
- *If $b \in \mathbb{R}$ then the sequence $(x_n) = (b, b, b, \ldots)$ is called the* constant sequence b. *Eg the constant sequence 300 is* $(300, 300, 300, 300, \ldots)$.

- *Given the expression* $x_n = (-1)^n$ *for* $n \in \mathbb{N}$*, we obtain*

$$(x_n)_n = (-1, 1, -1, 1, \ldots).$$

- *For* $a_n = \frac{(-1)^n}{2^n}$ *for* $n \in \mathbb{N}$*, we obtain* $(a_n) = \left(\frac{-1}{2}, \frac{1}{4}, \frac{-1}{8}, \ldots\right)$.

Definition 2.2. *Let* $(x_n)_n$ *be a sequence of real numbers and* $x \in \mathbb{R}$*.*

1. *We say that* $(x_n)_n$ converges *to* x *(as* n *tends to infinity) if*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ s.t.\ \forall n \in \mathbb{N}, n \geqslant N \Rightarrow |x_n - x| \leqslant \varepsilon.$$

In this case we write

$$\lim_n x_n = x,\ \lim_{n\to\infty} x_n = x,\ x_n \to x\ as\ n \to \infty$$

A sequence $(x_n)_n$ *is called* convergent *if there exists* $x \in \mathbb{R}$ *such that* $\lim_{x\to\infty} x_n = x$*; otherwise the sequence is called* divergent*.*

2. *We say that* $(x_n)_n$ tends to infinity *(as* n *tends to infinity) if*

$$\forall K \in \mathbb{R}\ \exists N \in \mathbb{N}\ s.t.\ \forall n \in \mathbb{N}, n \geqslant N \Rightarrow x_n \geqslant K.$$

In this case we write

$$\lim_n x_n = \infty,\ \lim_{n\to\infty} x_n = \infty,\ x_n \to \infty\ as\ n \to \infty$$

3. *We say that* $(x_n)_n$ tends to minus infinity *(as* n *tends to infinity) if*

$$\forall K \in \mathbb{R}\ \exists N \in \mathbb{N}\ s.t.\ \forall n \in \mathbb{N}, n \geqslant N \Rightarrow x_n \leqslant K.$$

$$\lim_n x_n = -\infty,\ \lim_{n\to\infty} x_n = -\infty,\ x_n \to -\infty\ as\ n \to \infty$$

In other words, if $x_n \to x$ *then no matter how small* $\varepsilon > 0$ *is chosen, there will be some stage in the sequence (stage* N*) beyond which all elements* x_n *will lie in the interval* $[x - \varepsilon, x + \varepsilon]$*.*

Example 2.2. $\left(\frac{1}{n^2}\right)_n$ *is convergent:*

Proof. Let $\varepsilon > 0$. Given $N \in \mathbb{N}$, if $n \geqslant N$ then

$$\left|\frac{1}{n^2} - 0\right| = \frac{1}{n^2} \leqslant \frac{1}{N^2}.$$

So choosing N so that $\frac{1}{N^2} < \varepsilon$, i.e., $N > \frac{1}{\sqrt{\varepsilon}}$ completes the proof. □

Example 2.3. *Let x be some real number and let $(x_n)_n$ be the constant sequence $(x, x, x, \ldots)$. Then*

$$x_n \to x \text{ as } n \to \infty.$$

Proof. Let $\varepsilon > 0$. Then given $N \in \mathbb{N}$, $n \geqslant N$ implies

$$|x_n - x| = |x - x| = 0 \leqslant \varepsilon,$$

as required. □

Example 2.4. *The sequence $\left(\frac{1}{n}\right)_n$ converges to 0 as $n \to \infty$.*

Proof. Let $\varepsilon > 0$. Fix $N \in \mathbb{N}$ such that $N \geqslant \frac{1}{\varepsilon}$. Then

$$|x_n - 0| = \left|\frac{1}{n} - 0\right| = \frac{1}{n} \leqslant \frac{1}{N} \leqslant \varepsilon,$$

as required. □

Example 2.5. *The sequence $\left(\frac{4n+300}{5n+2}\right)_n$ is convergent: in fact*

$$\frac{4n+300}{5n+2} \to \frac{4}{5} \text{ as } n \to \infty.$$

Proof. Let $\varepsilon > 0$. For $N \in \mathbb{N}$, $n \geqslant N$ implies

$$\left|\frac{4n+300}{5n+2} - \frac{4}{5}\right| = \left|\frac{5(4n+300) - 4(5n+2)}{5(5n+2)}\right| = \frac{1492}{25n+10} < \frac{1492}{25N}.$$

So setting $N \geqslant \frac{1492}{25\varepsilon}$, we complete the proof. □

Example 2.6. *For $x_n = (-1)^n$, the sequence $(x_n)_n$ is divergent.*

(Idea: suppose there is a limit x and then show that there is some $\varepsilon > 0$ for which $|x_n - x| > \varepsilon$ even for some very large n.)

Proof. Suppose that there is a limit, call it $x \in \mathbb{R}$. Then set $\varepsilon = \frac{1}{2}$. Since x is a limit, there exists N such that $n \geqslant N$ implies $|x_n - x| \leqslant \frac{1}{2}$. Since there are arbitrarily large even numbers, there is always some $n \geqslant N$ (eg $n = 2N$) such that $x_n = 1$, so $|1 - x| \leqslant \frac{1}{2}$, in particular, $x \geqslant \frac{1}{2}$. On the other hand, since there are arbitrarily large odd numbers, there is always some $n \geqslant N$ (eg $n = 2N + 1$) such that $x_n = -1$, so $|-1 - x| \leqslant \frac{1}{2}$, hence $x \leqslant -\frac{1}{2}$. The inequality $\frac{1}{2} \leqslant x \leqslant -\frac{1}{2}$ is impossible, so there is no limit. □

Example 2.7. *$(\sqrt{n})_n$ tends to ∞ as $n \to \infty$.*

Proof. Let $K > 0$. Then set $N \in \mathbb{N}$ to be greater than K^2. Hence $n \geqslant N$ implies $x_n = \sqrt{n} \geqslant \sqrt{N} > K$, as required. □

Example 2.8. *$\left(\frac{1}{n} - n\right)_n$ tends to $-\infty$ as $n \to \infty$.*

Proof. Let $K \in \mathbb{R}$. Then set $N \in \mathbb{N}$ so that $N > 1 - K$. So $n \geqslant N$ implies

$$x_n = \frac{1}{n} - n \leqslant 1 - N < K,$$

as required. □

Example 2.9. *(Standard sequences) Let $a \in \mathbb{R}$.*

1. *If $|a| < 1$ then $(a^n)_n$ converges to 0.*
2. *If $a > 1$ then $(a^n)_n$ tends to ∞.*

Proof. We assume Bernoulli's Inequality: for $x \geqslant 0$ and $n \in \mathbb{N}$,

$$(1 + x)^n \geqslant 1 + nx \qquad (*).$$

(Proof is an easy exercise in induction.)

1) Using the fact that $|a| < 1$ to deduce that $\frac{1}{|a|} - 1 > 0$, for $N \in \mathbb{N}$ and $n \geqslant N$,

$$|a^n - 0| = |a|^n = \frac{1}{\left(1 + \left(\frac{1}{|a|} - 1\right)\right)^n} \leqslant \frac{1}{\left(1 + n\left(\frac{1}{|a|} - 1\right)\right)} \leqslant \frac{1}{\left(1 + N\left(\frac{1}{|a|} - 1\right)\right)},$$

so choosing $N \in \mathbb{N}$ such that $N \geqslant \frac{1}{\varepsilon\left(\frac{1}{|a|} - 1\right)}$, we are finished.

2) Let $K \in \mathbb{R}$. Then for $n \in \mathbb{N}$ and $n \geqslant N$,

$$a^n = (1 + (a - 1))^n \geqslant 1 + n(a - 1) \geqslant N(a - 1).$$

So if $N \geqslant \frac{K}{a-1}$, we are finished. □

2.2 Limit Theorems

In this section we'll consider uniqueness and algebraic properties of limits.

Theorem 2.1. *If a sequence is convergent, its limit is unique.*

Proof. Let $(x_n)_n$ be a sequence that converges to both s and t. Let $\varepsilon > 0$. Then since $x_n \to s$, there exists $N_1 \in \mathbb{N}$ such that $n \geqslant N_1$ implies

$$|x_n - s| \leqslant \varepsilon.$$

Similarly, since $x_n \to t$, there exists $N_2 \in \mathbb{N}$ such that $n \geqslant N_2$ implies

$$|x_n - t| \leqslant \varepsilon.$$

Therefore taking $n \geqslant \max\{N_1, N_2\}$,

$$|s - t| = |s - x_n + x_n - t| \leqslant |s - x_n| + |x_n - t| \leqslant \varepsilon + \varepsilon = 2\varepsilon.$$

Since this holds for *any* $\varepsilon > 0$, this means $s = t$. $\square$

If a sequence can have very large values for large n, this can cause problems for the algebraic properties of that sequence (eg see next theorem). The following definition deals with that.

Definition 2.3. *A sequence $(x_n)_n$ is called* bounded *if there exists $M \geqslant 0$ such that $|x_n| \leqslant M$ for all $n \in \mathbb{N}$.*

Theorem 2.2. *Every convergent sequence is bounded.*

Proof. Suppose that $(x_n)_n$ is a convergent sequence and denote its limit by $x \in \mathbb{R}$. Then there exists $N \in \mathbb{N}$ such that $|x_n - x| \leqslant 1$ for all $n \geqslant N$. So $n \geqslant N$ implies

$$|x_n| = |x_n - x + x| \leqslant |x_n - x| + |x|.$$

So defining $M := \max\{|x_1|, |x_2|, \ldots, |x_N|, |x| + 1\}$, we deduce that $|x_n| \leqslant M$ for all $n \in \mathbb{N}$. $\square$

The next theorem simplifies many questions involving combinations of more than one sequence.

Theorem 2.3. *Let $a \in \mathbb{R}$ and $(x_n)_n, (y_n)_n$ be convergent sequences with $x_n \to x$ and $y_n \to y$ as $n \to \infty$. Then*

1. $x_n + y_n \to x + y$;
2. $x_n y_n \to xy$;
3. *If $y_n \neq 0$ for all $n \in \mathbb{N}$ and $y \neq 0$ then* $\frac{1}{y_n} \to \frac{1}{y}$;
4. *If $y_n \neq 0$ for all $n \in \mathbb{N}$ and $y \neq 0$ then* $\frac{x_n}{y_n} \to \frac{x}{y}$;
5. $ax_n \to ax$ *and* $a + x_n \to a + x$.

Proof. 1) Let $\varepsilon > 0$. Since $x_n \to x$, there exists $N_1 \in \mathbb{N}$ such that

$$|x_n - x| \leqslant \frac{\varepsilon}{2} \quad \forall n \geqslant N_1.$$

Similarly, since $y_n \to y$, there exists $N_2 \in \mathbb{N}$ such that

$$|y_n - y| \leqslant \frac{\varepsilon}{2} \quad \forall n \geqslant N_2.$$

Set $N := \max\{N_1, N_2\}$. Then $n \geqslant N$ implies

$$|(x_n + y_n) - (x + y)| = |(x_n - x) + (y_n - y)| \leqslant |x_n - x| + |y_n - y| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required.

2) Let $\varepsilon > 0$. First note that

$$\begin{aligned} |x_n y_n - xy| &= |x_n y_n - x_n y + x_n y - xy| \\ &\leqslant |x_n y_n - x_n y| + |x_n y - xy| \\ &\leqslant |x_n||y_n - y| + |y||x_n - x| \qquad (*) \end{aligned}$$

Since $x_n \to x$, there exists $N_1 \in \mathbb{N}$ such that $|x_n - x| \leqslant \frac{\varepsilon}{2(|y|+1)}$ for all $n \geqslant N_1$.

By Theorem 2.2, $(x_n)_n$ is bounded, i.e., there exists $M \geqslant 0$ such that $|x_n| \leqslant M$ for all $n \in N$. Since also $y_n \to y$, there exists $N_2 \in \mathbb{N}$ such that $|y_n - y| \leqslant \frac{\varepsilon}{2(M+1)}$ for all $n \geqslant N_2$. Let $N := \max\{N_1, N_2\}$. Then by $(*)$, $n \geqslant N$ implies

$$|x_n y_n - xy| \leqslant |x_n||y_n - y| + |y||x_n - x| \leqslant M \cdot \frac{\varepsilon}{2(M+1)} + |y| \cdot \frac{\varepsilon}{2(|y|+1)} \leqslant \varepsilon,$$

as required.

3) Let $\varepsilon > 0$. First observe that

$$\left|\frac{1}{y_n} - \frac{1}{y}\right| = \left|\frac{y - y_n}{y_n y}\right| = \frac{|y - y_n|}{|y||y_n|} \quad (**).$$

Since $y_n \to y$, there exists $N_1 \in \mathbb{N}$ such that $n \geqslant N_1$ implies

$$|y - y_n| \leqslant \frac{|y|}{2} \quad \forall n \geqslant N_1.$$

Therefore,

$$|y| = |y_n + (y - y_n)| \leqslant |y_n| + |y - y_n| \leqslant |y_n| + \frac{|y|}{2}$$

which implies $|y_n| \geqslant \frac{|y|}{2}$ for all $n \geqslant N_1$.

Moreover, there exists $N_2 \in \mathbb{N}$ such that $n \geqslant N_2$implies

$$|y_n - y| \leqslant \frac{\varepsilon |y|^2}{2} \quad \forall n \geqslant N_2.$$

So setting $N := \max\{N_1, N_2\}$, $(**)$ implies

$$\left|\frac{1}{y_n} - \frac{1}{y}\right| \leqslant \frac{\varepsilon |y|^2}{2|y|\left(\frac{|y|}{2}\right)} = \varepsilon,$$

as required.

4) Since $\frac{x_n}{y_n} = x_n \cdot \frac{1}{y_n}$, this follows by 2) and 3).

5) Exercise. $\square$

Example 2.10. *Let $p \in \mathbb{N}$. Then $(x_n)_n = \left(2 + \frac{1}{n^p}\right)_n$ converges to 2.*

Proof. Let $(a_n)_n$ be the constant sequence 2 and let $(b_n)_n$ be the sequence given by $b_n = \frac{1}{n^p}$ for all $n \in \mathbb{N}$. Then $x_n = a_n + b_n$.

Further, let $c_n = \frac{1}{n}$ for all $n \in \mathbb{N}$. We know that $\frac{1}{n} \to 0$ as $n \to \infty$. So by Theorem 2.3, $c_n c_n \cdots c_n = c_n^p \to 0$ as $n \to \infty$. Hence $a_n \to 2$ and $b_n \to 0$ as $n \to \infty$, so by Theorem 2.3, $x_n \to 2 + 0 = 2$ as $n \to \infty$. $\square$

(Above we used the constant sequence 2 and the sequence $(1/n)_n$ as building blocks for which we knew the limiting behaviour. We'll be able to use this type of idea from here on, unless we are asked to prove 'from first principles' (or a similar phrase) that a sequence converges.)

Example 2.11. *The sequence* $\left(\frac{n^2+4n}{n^3-3}\right)_n$ *is convergent, in fact*

$$\left(\frac{n^2+4n}{n^3-3}\right)_n \to 0 \text{ as } n \to \infty.$$

Proof.

$$\frac{n^2+4n}{n^3-3} = \frac{\frac{1}{n}+\frac{4}{n^2}}{1-\frac{3}{n^3}} = \frac{\frac{1}{n}+4\left(\frac{1}{n}\right)^2}{1-3\left(\frac{1}{n}\right)^3}.$$

Since $\frac{1}{n} \to 0$, by Theorem 2.3,

$$\frac{n^2+4n}{n^3-3} = \frac{\frac{1}{n}+4\left(\frac{1}{n}\right)^2}{1-3\left(\frac{1}{n}\right)^3} \to \frac{0+4\cdot 0^2}{1-3\cdot 0^3} = 0.$$

□

2.3 Monotone Sequences and Subsequences

In most of our examples of convergent sequences so far, proving convergence has involved guessing a limit before taking any further steps. In this section we'll develop tools which can overcome this problem: even when we don't have a candidate for a limit, in some cases we can use the completeness of the reals to guarantee that one exists.

Definition 2.4. *Let* $(x_n)_n$ *be a sequence of real numbers.*

- *We say that a sequence is* increasing *if it satisfies* $x_1 \leqslant x_2 \leqslant x_3 \leqslant \cdots$.
- *We say that a sequence is* decreasing *if it satisfies* $x_1 \geqslant x_2 \geqslant x_3 \geqslant \cdots$.
- *We say that a sequence is* monotone *if it is either increasing or decreasing.*

Example 2.12. *a)* $(a_n) = (n)_n$ *is increasing;*

b) $(b_n) = (4^n)_n$ *is increasing;*

c) $(c_n) = \left(3 - \frac{1}{n}\right)_n$ *is increasing;*

d) $(d_n) = (1, 1, 2, 2, 3, 3, 4, 4, \ldots)_n$ *is increasing (despite some adjacent terms being equal);*

e) $(e_n) = \left(\frac{3}{n}\right)_n$ *is decreasing;*

f) $(f_n) = (-2n)_n$ *is decreasing;*

g) $(g_n) = (k, k, k, k, \ldots)$ *for some* $k \in \mathbb{R}$ *is both increasing and decreasing;*

h) $(h_n) = ((-1)^n)_n$ *is not monotone;*

i) $(i_n) = \left(\frac{(-1)^n}{n}\right)_n$ *is not monotone.*

Note that $(c_n)_n, (e_n)_n, (g_n)_n, (h_n)_n, (i_n)_n$ *are all bounded, while* $(a_n)_n, (b_n)_n, (d_n)_n$ *are unbounded. Also note that the bounded monotone sequences* $((c_n)_n, (e_n)_n, (g_n)_n)$ *are convergent, while the unbounded sequences are not (as in Theorem 2.2). These are examples of a broader phenomenon:*

Theorem 2.4 (Monotone Convergence Theorem)**.** *Let* $(x_n)_n$ *be a monotone sequence. Then the following are equivalent:*

1. $(x_n)_n$ *is convergent;*
2. $(x_n)_n$ *is bounded.*

Proof. (1$\Rightarrow$ 2): This is true by Theorem 2.2.

(2$\Rightarrow$ 1): Assume first that $(x_n)_n$ is bounded and increasing. Let $A := \{x_n : n \in \mathbb{N}\}$. Since we have assumed that A is bounded, A has a supremum $x = \sup A \in \mathbb{R}$ by the completeness of the reals.

Claim. $x_n \to x$ *as* $n \to \infty$.

Proof of Claim. Let $\varepsilon > 0$. Since $x = \sup A$, $x - \varepsilon$ is not an upper bound for A. Hence there exists $N \in \mathbb{N}$ such that $x_N \geqslant x - \varepsilon$. Since $(x_n)_n$ is increasing, $n \geqslant N$ implies

$$x - \varepsilon \leqslant x_N \leqslant x_n \leqslant x, \text{ i.e., } |x_n - x| \leqslant \varepsilon \ \forall n \geqslant N,$$

so the claim is proved. □

The proof where $(x_n)_n$ is decreasing is similar. □

Corollary 2.5. *Suppose that $(x_n)_n$ is a bounded sequence.*

1. *If $(x_n)_n$ is increasing then $\lim_{n\to\infty} x_n = \sup_n x_n$.*
2. *If $(x_n)_n$ is decreasing then $\lim_{n\to\infty} x_n = \inf_n x_n$.*

Proof. This follows immediately from the proof of Theorem 2.4. □

Example 2.13. *If $0 < a < 1$ then $a^n \to 0$ as $n \to \infty$. (Note that we've already proved this, but here's an alternative proof.)*

Proof. Let $x_n = a^n$. Since $a^{n+1} < a^n$ for all $n \in \mathbb{N}$, $(x_n)_n$ is a decreasing sequence bounded above by a and below by 0. So MCT implies there is a limit $x \in \mathbb{R}$. Therefore (see eg TS2 Q11) $x_{n+1} \to x$ also. So $x_{n+1} = a^{n+1} = ax_n \to x$. But since $x_n \to x$, Theorem 2.3 implies that $ax_n \to ax$. Hence $ax = x$. Since $a \neq 1$, the only solution is $x = 0$. □

Exercise: extend this to $|a| < 1$.

Example 2.14. *Define $(x_n)_n$ by $x_1 = 1$ and $x_{n+1} = \sqrt{1 + x_n}$ for $n \in \mathbb{N}$. The sequence is bounded and increasing, thus convergent: indeed*

$$x_n \to \frac{1+\sqrt{5}}{2}.$$

Proof. **Claim 1.** $|x_n| \leqslant 2$ for all $n \in \mathbb{N}$.

Proof of Claim 1. We prove this by induction. For $n = 1$, $x_1 = 1 \leqslant 2$. Suppose that the claim is true for $n \in \mathbb{N}$. Then

$$x_{n+1} = \sqrt{1 + x_n} \leqslant \sqrt{1+2} = \sqrt{3} \leqslant 2,$$

so the claim is true for $n + 1$. Hence by induction, the claim holds. □

Claim 2. $(x_n)_n$ is increasing, i.e., for all $n \in \mathbb{N}$, $x_n \leqslant x_{n+1}$.

Proof. For $n = 1$ we have $x_1 = 1 \leqslant \sqrt{1+1} = x_2$. Assume the claim is true for $n \in \mathbb{N}$. Then $x_{n+1} = \sqrt{1 + x_n} \leqslant \sqrt{1 + x_{n+1}} = x_{n+2}$, so the claim is true for $n + 1$. Hence by induction the claim holds. □

Combining these claims and the MCT there exists $x \in \mathbb{R}$ such that $x_n \to x$.

By Theorem 2.3, $1 + x_n \to 1 + x$, and by TS2 Q5, setting $y_n = 1 + x_n$ and $y = 1 + x$, we have $\sqrt{y_n} \to \sqrt{y}$, i.e., $\sqrt{1 + x_n} \to \sqrt{1 + x}$. Since $x_{n+1} = \sqrt{1 + x_n} \to x$, this means that $x = \sqrt{1 + x}$ and hence $x^2 = 1 + x$. Therefore

$$x = \frac{1 + \sqrt{5}}{2} \text{ or } x = \frac{1 - \sqrt{5}}{2}.$$

But since $x_n \geqslant 0$ for all $n \in \mathbb{N}$, only the positive solution for x is possible, i.e., $x = \frac{1+\sqrt{5}}{2}$. □

Example 2.15. *Let $(x_n)_n$ be given by*

$$x_n = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2}.$$

The sequence is bounded and increasing, and hence convergent (in fact the limit is $\frac{\pi^2}{6}$, but we won't show that here).

Proof.

Claim. $0 < x_n \leqslant 2$ *for all* $n \in \mathbb{N}$.

Proof of Claim. We write

$$\begin{aligned} 0 &\leqslant \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots + \frac{1}{n^2} \leqslant 1 + \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{(n-1)n} \\ &= 1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 + 1 - \frac{1}{n} \leqslant 2. \end{aligned}$$

□

So $(x_n)_n$ is bounded. Clearly $(x_n)_n$ is increasing since

$$x_n = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} \leqslant \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \frac{1}{(n+1)^2} = x_{n+1},$$

so the proof is complete. □

Definition 2.5. *Let $(x_n)_n$ be a sequence and $(m_k)_k$ be a sequence of natural numbers $m_1 < m_2 < m_3 < \cdots$. then the sequence $(x_{m_k})_k$ is called a* subsequence *of $(x_n)_n$.*

Note that in this definition, we must have $m_k \geqslant k$ for all $k \in \mathbb{N}$.

Example 2.16. *The sequence $(x_n)_n = \left(\frac{1}{n}\right)_n$ has various subsequences, eg*

- $(s_k)_k = (x_{2k})_k = \left(\frac{1}{2k}\right)_k = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots\right)$;
- $(t_k)_k = (x_{k+3})_k = \left(\frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \ldots\right)$;
- *the sequence $(x_n)_n$ is a subsequence of itself.*

Note that eg $\left(\frac{1}{3}, \frac{1}{2}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \ldots\right)$ is not *a subsequence of $(x_n)_n$: the order of terms must be preserved.*

Theorem 2.6. *Suppose that $(x_n)_n$ converges to $x \in \mathbb{R}$. Then any subsequence also converges to $x \in \mathbb{R}$.*

Proof. Let $\varepsilon > 0$. Since $x_n \to x$ then there exists $N \in \mathbb{N}$ such that $n \geqslant N$ implies $|x_n - x| \leqslant \varepsilon$. Since $(x_{m_k})_k$ is a subsequence of $(x_n)_n$. Let $K \in \mathbb{N}$ be such that $k \geqslant N$ implies $m_k \geqslant N$. Then

$$|x_{m_k} - x| \leqslant \varepsilon \text{ for all } k \geqslant K,$$

as required. $\square$

N.B. that we could have made the statement of the theorem into an 'if and only if' since if any subsequence of $(x_n)_n$ converges to x, then certainly $(x_n)_n$ converges to x - since $(x_n)_n$ is a subsequence of itself.

Theorem 2.7 (Monotone Subsequence Theorem)**.** *Any sequence of real numbers $(x_n)_n$ has a monotone subsequence.*

Proof. In this proof, we say that $p \in \mathbb{N}$ is a 'peak' if

$$x_p \geqslant x_n \text{ for all } n \geqslant p.$$

Case 1: $(x_n)_n$ has infinitely many peaks. We list the peaks in the order in which they occur: $p_1 < p_2 < p_3 < \cdots$. From the definition of peaks, we have

$$x_{p_1} \geqslant x_{p_2} \geqslant x_{p_3} \geqslant \cdots,$$

so $(x_{p_k})_k$ is the monotone (decreasing) subsequence we require.

Case 2: $(x_n)_n$ has finitely many peaks. Again we list all the peaks in the order in which they occur: $p_1 < p_2 < \cdots < p_N$. Let $t_1 > p_N$. Since t_1 is not a peak, there exists $t_2 > t_1$ such that $x_{t_2} > x_{t_1}$. Since t_2 is also not a peak, there exists $t_3 > t_2$ such that $x_{t_3} > x_{t_2}$. Proceeding in this way, we obtain an infinite sequence $t_1 < t_2 < t_3 < t_4 < \cdots$ such that $x_{t_1} < x_{t_2} < x_{t_3} < x_{t_4} < \cdots$. So $(x_{t_k})_k$ is the monotone (increasing) subsequence we require. $\square$

Theorem 2.8 (Bolzano-Weierstrass Theorem)**.** *Let $(x_n)_n$ be a bounded sequence of real numbers. Then there exists a subsequence $(x_{m_k})_k$ and a real number x such that $x_{m_k} \to x$ as $n \to \infty$.*

Proof. The Monotone Subsequence Theorem guarantees the existence of a monotone subsequence $(x_{m_k})_k$. Then since $(x_{m_k})_k$ is also bounded, the Monotone Convergence Theorem implies that $(x_{m_k})_k$ converges to some limit x. $\square$

Example 2.17. *Let $(x_n)_n = ((-1)^n)_n$. Then this is a bounded sequence, eg bounded above by 2 and below by -2. So the Bolzano-Weierstrass Theorem implies there exists a convergent subsequence. In fact in this case we can check this by hand: eg the subsequence $(x_{4n})_n = (1, 1, 1, \ldots)$ converges to 1. On the other hand, the subsequence $(x_{6n+1})_n = (-1, -1, -1, \ldots)$ converges to -1.*

Example 2.18. *Consider $(\sin n)_n$. Note that this sequence is bounded above by 1 and below by -1. Hence the Bolzano-Weierstrass Theorem implies that there exists $(m_k)_k$ such that $(x_{m_k})_k$ is a convergent subsequence. (In fact, while we won't show this here, for any $x \in [-1, 1]$ there exists a subsequence which converges to x.)*

Chapter 3

Series

Outside maths, 'sequence' and 'series' are often understood to mean the same thing. However, in maths they are distinct notions.

Notation: Given a sequence $(x_n)_n$, for $n \in \mathbb{N}$ and $m \in \mathbb{N}$ with $n \leqslant m$, we write

$$x_n + x_{n+1} + x_{n+2} + \cdots + x_m = \sum_{k=n}^{m} x_k.$$

Definition 3.1.

- *Given a sequence $(x_n)_n$, we can formally write $x_1 + x_2 + x_3 + \cdots$ as $\sum_{n=1}^{\infty} x_n$. This is called the (infinite)* series *generated by $(x_n)_n$.*

- *For each $n \in \mathbb{N}$, we write*

$$s_n = x_1 + x_2 + \cdots + x_n = \sum_{k=1}^{n} x_k,$$

  *the n*th partial sum *of our series.*

- *We say that the series $\sum_{n=1}^{\infty} x_n$* converges *if the partial sums converge, i.e., the limit $\lim_{n\to\infty} s_n$ exists. In this case, $\sum_{n=1}^{\infty} x_n$ is the (infinite) sum of our sequence $(x_n)_n$.*

- *If $(s_n)_n$ does not converge, we say that the series* diverges*. If $s_n \to +\infty$ or $s_n \to -\infty$, we write $\sum_{n=1}^{\infty} x_n = +\infty$, $\sum_{n=1}^{\infty} x_n = +\infty$ respectively.*

N.B. If $(s_n)_n$ doesn't converge or tend to $+\infty$ or $-\infty$, then we don't think of $\sum_{n=1}^{\infty} x_n$ as a sum.

Example 3.1. *Consider the sequence $(x_n)_n$ where $x_n = \frac{1}{n(n+1)}$ for all $n \in \mathbb{N}$. The associated series is*

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

The nth partial sum is

$$\begin{aligned} s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} &= \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)} \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1}. \end{aligned}$$

Since $\left(\frac{1}{n+1}\right)_n$ converges to 0 and $(1)_n$ converges to 1, then $(s_n)_n = \left(1 - \frac{1}{n+1}\right)_n$ converges to 1. Hence $\sum_{n=1}^{\infty} x_n$ is a convergent series with sum equal to 1.

Example 3.2. *Let $(x_n)_n = ((-1)^n)_n$. Then the series this generates is $\sum_{n=1}^{\infty}(-1)^n$. Here the nth partial sum is*

$$s_n = \sum_{k=1}^{n}(-1)^k = \begin{cases} -1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

Hence $(s_n)_n$ is a divergent sequence (check), so $(x_n)_n = ((-1)^n)_n$ is divergent.

In the first example, but not the second, the sequence which generated the sequence converged to zero. In fact this is necessary for a sequence to converge:

Theorem 3.1. *Let $\sum_{n=1}^{\infty} x_n$ be a convergent series. Then $x_n \to 0$ as $n \to \infty$.*

Proof. $\sum_{n=1}^{\infty} x_n$ being convergent means that the partial sums $(s_n)_n$ converge to some limit s. Note that

$$x_n = s_n - s_{n-1} \text{ for all } n \geqslant 2.$$

Since $s_n \to s$ and $s_{n-1} \to s$, this means that $x_n \to s - s = 0$, as required. $\square$

It might be nice if the converse of this theorem were true (i.e., $x_n \to 0$ implies $\sum_{n=1}^{\infty} x_n$ converges), but as in the next example, this isn't true, which leads to some fundamentally important theory.

Example 3.3. *The* harmonic series *is defined as*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots.$$

Theorem 3.2. $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

Proof. Letting $s_n = \sum_{k=1}^{\infty} \frac{1}{k}$ be the nth partial sum, we'll show $s_n \to \infty$ by grouping summands in an appropriate way. For $k \in \mathbb{N}$, we write

$$s_{2^k} = 1 + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \cdots + \frac{1}{16}\right)$$
$$+ \cdots + \left(\frac{1}{2^{k-1}+1} + \frac{1}{2^{k-1}+2} + \cdots + \frac{1}{2^k}\right).$$

So the jth bracketed expression is

$$t_j = \frac{1}{2^{j-1}+1} + \frac{1}{2^{j-1}+2} + \cdots + \frac{1}{2^j}.$$

There are $2^j - 2^{j-1} = 2^{j-1}(2-1) = 2^{j-1}$ terms in this sum, each of which is $\geqslant \frac{1}{2^j}$. So $t_j \geqslant 2^{j-1}\frac{1}{2^j} = \frac{1}{2}$ for any $j \in \mathbb{N}$. Hence

$$s_{2^k} = 1 + t_1 + t_2 + \cdots + t_k \geqslant 1 + \frac{k}{2}.$$

So for any $M \in \mathbb{R}$, let $M' \geqslant M$ be an integer. Then $n \geqslant 2^{2M'}$ implies that

$$s_n \geqslant s_{2^{2M'}} \geqslant 1 + \frac{2M'}{2} = 1 + M' > M.$$

So $s_n \to \infty$, as required. $\square$

Example 3.4. *Given $a \in \mathbb{R}$ and $r \in \mathbb{R}\backslash\{0\}$, the sequence $(ar^{n-1})_n$ generates the* geometric series $\sum_{n=1}^{\infty} ar^{n-1}$.

Theorem 3.3. *The geometric series $\sum_{n=1}^{\infty} ar^{n-1}$ converges if and only if $|r| < 1$. When $|r| < 1$ then $\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}$.*

Proof. We start by noting that the nth partial sum is

$$s_n = a + ar + ar^2 + \cdots + ar^{n-1} = \frac{a(1-r^n)}{1-r} \qquad \text{(Exercise)}.$$

So if $|r| < 1$ then $r^n \to 0$ (see Example 2.9), so $s_n \to \frac{a}{1-r}$ by Theorem 2.3.

If $r = 1$ then $s_n = na$ which tends to $+\infty$ if $a > 0$ and $-\infty$ if $a < 0$.

If $r = -1$, then s_n is zero if n is even and a if n is odd, so the series is divergent.

If $|r| > 1$ then $(s_n)_n$ diverges (Exercise). $\square$

So a particular case is the sequence $\left(\left(\frac{1}{2}\right)^{n-1}\right)_n$ generating the series

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n-1} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = \frac{1}{1-\frac{1}{2}} = 2.$$

3.1 The Comparison Test

As we've seen before, its common to try to understand complicated mathematical examples in terms of some basic building blocks. The Comparison Test is another example of this approach: given a new series $\sum_{n=1}^{\infty} x_n$, we can try to investigate its convergence in terms of some known series $\sum_{n=1}^{\infty} a_n$ which we've studied before.

Theorem 3.4 (The Comparison Test)**.** *Suppose that $(x_n)_n$ and $(a_n)_n$ are sequences with no negative terms.*

1. *If $\sum_{n=1}^{\infty} a_n$ is convergent and $x_n \leqslant a_n$ for all $n \in \mathbb{N}$, then $\sum_{n=1}^{\infty} x_n$ is convergent.*
2. *If $\sum_{n=1}^{\infty} a_n$ is divergent and $x_n \geqslant a_n$ for all $n \in \mathbb{N}$, then $\sum_{n=1}^{\infty} x_n$ is divergent.*

Proof. 1) Denote the partial sums by $A_n = \sum_{k=1}^{n} a_n$ and $X_n = \sum_{k=1}^{n} x_n$. By assumption $(A_n)_n$ is convergent: let A denote the limit. Since all summands a_n are positive, $(A_n)_n$ is monotone increasing with $A_n \leqslant A$. Since $x_k \leqslant a_k$ for all $k \in \mathbb{N}$, we have $X_n \leqslant A_n \leqslant A$ for all $n \in \mathbb{N}$, so $(X_n)_n$ is bounded.

Hence the Monotone Convergence Theorem implies that $(X_n)_n$ is convergent (i.e., $\sum_{n=1}^{\infty} x_n$ is convergent).

2) Suppose that $\sum_{n=1}^{\infty} x_n$ is convergent. Then, swapping the roles of $(a_n)_n$ and $(x_n)_n$ in 1), we deduce that $\sum_{n=1}^{\infty} a_n$ converges. This contradiction shows that $\sum_{n=1}^{\infty} x_n$ divverges. □

With minor changes, the proof implies a stronger result which we will also refer to as the Comparison Test:

Corollary 3.5. *Suppose that $(x_n)_n$ and $(a_n)_n$ are sequences with no negative terms.*

1. *If $\sum_{n=1}^{\infty} a_n$ is convergent and there exists $N \in \mathbb{N}$ and $c > 0$ such that $x_n \leqslant ca_n$ for all $n \geqslant N$, then $\sum_{n=1}^{\infty} x_n$ is convergent.*
2. *If $\sum_{n=1}^{\infty} a_n$ is divergent and there exists $N \in \mathbb{N}$ and $c > 0$ such that $x_n \geqslant ca_n$ for all $n \geqslant \mathbb{N}$, then $\sum_{n=1}^{\infty} x_n$ is divergent.*

Example 3.5. *Show that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent.*

Proof. For $n \geqslant 1$, $n^2 \geqslant \frac{n(n+1)}{2}$, so

$$\frac{1}{n^2} \leqslant \frac{2}{n(n+1)}.$$

Now since $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ is convergent by Example 3.1, the Comparison Test (in particular Corollary 3.5 part 1 with $c = 2$) implies that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent. □

More generally:

Theorem 3.6. *Let $\alpha \in \mathbb{R}$. Then $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ is convergent if and only if $\alpha > 1$.*

Proof. Suppose that $\alpha \geqslant 2$. Then $\frac{1}{n^\alpha} \leqslant \frac{1}{n^2}$ for all $n \in \mathbb{N}$, so $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ is convergent using Example 3.5 and the Comparison Test.

Now suppose that $\alpha \leqslant 1$. Then $\frac{1}{n^\alpha} \geqslant \frac{1}{n}$ for all $n \in \mathbb{N}$. So using the fact that the harmonic series is divergent, the comparison test implies that $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ is divergent.

We omit the proof of convergence in the case $\alpha \in (1,2)$ (the standard proof uses the 'Integral Test'). $\square$

Example 3.6. *Given $x_n = \frac{n+1}{5n^3-n-1}$, is $\sum_{n=1}^{\infty} x_n$ convergent?*

Yes: Since $n+1 \leqslant 2n \leqslant 2n^3$ for all $n \in \mathbb{N}$,

$$\frac{n+1}{5n^3-n-1} \leqslant \frac{2n}{5n^3-n-1} < \frac{2n}{5n^3-2n^3} = \frac{2n}{3n^3} = \frac{2}{3}\frac{1}{n^2}.$$

So using the Comparison Test, with comparison series $\sum_{n=1}^{\infty} \frac{1}{n^2}$, we see that $\sum_{n=1}^{\infty} x_n$ is convergent.

Note that given a sequence $(x_n)_n$, for any $m \in \mathbb{N}$, we can ask the same question about the convergence of $\sum_{n=m}^{\infty} x_n$ (actually by Corollary 3.5, convergence/divergence of this series is equivalent to that of $\sum_{n=1}^{\infty} x_n$).

Example 3.7. *Investigate the convergence of $\sum_{n=1}^{\infty} x_n$ where $x_n = \frac{n-1}{n^2+3n+4}$.*

For all $n \in \mathbb{N}$, $n^3 + 3n + 4 \leqslant n^2 + 3n^2 + 4n^2 = 8n^2$. Moreover, $n - 1 \geqslant \frac{n}{2}$ for all $n \geqslant 3$. So for $n \geqslant 3$,

$$\frac{n-1}{n^2+3n+4} \geqslant \frac{\frac{n}{2}}{8n^2} = \frac{1}{16n}$$

So we can use the Comparison Test, comparing our series with the harmonic series ($x_n \geqslant c\frac{1}{n}$ for $n \geqslant N$ in Corollary 3.5 with $N = 3$ and $c = \frac{1}{16}$) to see that $\sum_{n=1}^{\infty} x_n = \infty$.

Example 3.8. *Investigate the convergence of $\sum_{n=1}^{\infty} x_n$ where $x_n = \frac{\sqrt{2n^3+2}}{n^3+3}$.*

The dominant term on the top is $\sqrt{n^3} = n^{\frac{3}{2}}$ and on the bottom is n^3. So we will try to compare with $\frac{n^{\frac{3}{2}}}{n^3} = \frac{1}{n^{\frac{3}{2}}}$.

$$x_n = \frac{\frac{1}{n^{\frac{3}{2}}}\sqrt{2n^3+2}}{\frac{1}{n^{\frac{3}{2}}}(n^3+3)} = \frac{\frac{1}{\sqrt{n^3}}\sqrt{2n^3+2}}{\frac{n^3}{n^{\frac{3}{2}}} + \frac{3}{n^{\frac{3}{2}}}} = \frac{\sqrt{2+\frac{2}{n^3}}}{n^{\frac{3}{2}} + \frac{3}{n^{\frac{3}{2}}}} \leqslant \frac{\sqrt{4}}{n^{\frac{3}{2}}} = \frac{2}{n^{\frac{3}{2}}}.$$

Since $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ is convergent, the Comparison Test implies that $\sum_{n=1}^{\infty} x_n$ converges.

Theorem 3.7 (The Ratio Test)**.** *Let $(x_n)_n$ be a sequence of positive terms.*

i) If there exists $N \in \mathbb{N}$ and $r < 1$ such that $n \geqslant N$ implies that $\frac{x_{n+1}}{x_n} \leqslant r$ then $\sum_{n=1}^{\infty} x_n$ is convergent.

ii) If there exists $N \in \mathbb{N}$ and $r > 1$ such that $n \geqslant N$ implies that $\frac{x_{n+1}}{x_n} \geqslant r$ then $\sum_{n=1}^{\infty} x_n = \infty$.

Note that for the geometric series $\sum_{n=1}^{\infty} ar^{n-1}$, the $(n+1)$st term over the nth term is $\frac{ar^n}{ar^{n-1}} = r$, so the ratio test agrees with Theorem 3.3:

- $0 < r < 1$ implies $\sum_{n=1}^{\infty} ar^{n-1}$ is convergent.
- $r > 1$ implies $\sum_{n=1}^{\infty} ar^{n-1}$ is divergent.

Proof of the Ratio Test. By Corollary 3.5, we only need to apply the comparison test to terms x_n for $n \geqslant N$.

i) $x_{N+1} \leqslant x_N$ and $x_{N+2} \leqslant rx_{N+1} \leqslant r^2 x_N$ and so on, so $x_{N+k} \leqslant r^k x_N$. So in the Comparison Test, we use $(a_n)_n = (x_N r^{-N} \cdot r^n)_n$ as our sequence to compare $(x_n)_n$ with. Since $(a_n)_n$ is a convergent geometric series and for $n \geqslant N$ we have $x_n = x_{N+(n-N)} \leqslant r^{n-N} x_N = a_n$, the comparison test implies that $\sum_{n=1}^{\infty} x_n$ converges.

ii) Exercise. $\square$

Note that if we know that $\lim_{n\to\infty} \frac{x_{n+1}}{x_n}$ exists, then we can take this as our value of r in the comparison test, so we just check if this limit is > 1 or < 1.

It's important to note that in the case that a series has $\frac{x_{n+1}}{x_n} \to 1$, the Ratio Test gives us no information at all. For example, $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent, but $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent, but in the former case, $\frac{x_{n+1}}{x_n} = \frac{\frac{1}{n+1}}{\frac{1}{n}} = \frac{n}{n+1} \to 1$, while in the latter case, $\frac{x_{n+1}}{x_n} = \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \frac{n^2}{(n+1)^2} \to 1$.

Example 3.9. *Given a fixed $x > 0$, does $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converge?*

Solution: The sequence of terms here is $x_n = \frac{x^n}{n!}$. Then

$$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{x^{n+1}}{(n+1)!}\right)}{\left(\frac{x^n}{n!}\right)} = \frac{x^{n+1} n!}{x^n (n+1)!} = \frac{x}{n+1} \to 0 \text{ as } n \to \infty.$$

Hence the ratio test shows that $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ *converges, regardless of the value of* x *we had in the beginning.*

Note that one consequence of this is that by Theorem 3.1, $\frac{x^n}{n!} \to 0$ *as* $n \to \infty$*, independently of the value of* x *(this can be extended to negative* x *also, as we'll see later).*

Example 3.10. *Does the series* $\sum_{n=1}^{\infty} \frac{n}{3^n}$ *converge?*

- *Solution 1: Letting* $x_n = \frac{n}{3^n}$ *for all* $n \in \mathbb{N}$*, we consider*

  $$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{n+1}{3^{n+1}}\right)}{\left(\frac{n}{3^n}\right)} = \frac{3^n(n+1)}{3^{n+1}n} = \frac{1}{3}\left(\frac{n}{n+1}\right) \to \frac{1}{3} \text{ as } n \to \infty.$$

  So by the Ratio Test, the sequence is convergent.

- *Solution 2: Since* $x_n \leqslant \frac{1}{2^n}$ *for all* $n \in \mathbb{N}$*, and we know that the series* $\sum_{n=1}^{\infty} \frac{1}{2^n}$ *is convergent, the Comparison Test implies that* $\sum_{n=1}^{\infty} \frac{n}{3^n}$ *converges.*

Example 3.11. *Does the series* $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$ *converge?*

Solution: To apply the Ratio Test, let $x_n = \frac{1}{n^2+1}$*. Then*

$$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{1}{(n+1)^2+1}\right)}{\left(\frac{1}{n^2+1}\right)} = \frac{n^2+1}{(n+1)^2+1} = \frac{n^2+1}{n^2+2n+1} = \frac{1+\frac{1}{n^2}}{1+\frac{2}{n}+\frac{1}{n^2}} \to 1 \text{ as } n \to \infty,$$

so the Ratio Test tells us nothing in this case.

On the other, hand $\frac{1}{n^2+1} \leqslant \frac{1}{n^2}$ *for all* $n \in \mathbb{N}$ *and since* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *converges, the Comparison Test implies that* $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$ *is convergent.*

Example 3.12. *Does the series* $\sum_{n=1}^{\infty} \frac{n!}{n^4+3}$ *converge?*

Solution: Letting $x_n = \frac{n!}{n^4+3}$*,*

$$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{(n+1)!}{(n+1)^4+3}\right)}{\left(\frac{n!}{n^4+3}\right)} = \frac{(n+1)!}{n!}\frac{n^4+3}{(n+1)^4+3} = (n+1)\left(\frac{n^4+3}{(n+1)^4+3}\right) \to \infty \text{ as } n \to \infty,$$

so the Ratio Test (with r *any number* > 1*) implies that the series diverges.*

3.2 Sums of positive and negative terms

So far, we've mostly restricted ourselves to series $\sum_{n=1}^{\infty} x_n$ where all $x_n \geqslant 0$. But consider, for example $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$. Does it converge?

Definition 3.2. *A series $\sum_{n=1}^{\infty} x_n$ is called* absolutely convergent *if $\sum_{n=1}^{\infty} |x_n|$ is convergent.*

Theorem 3.8. *Every absolutely convergent series converges.*

Proof is omitted.

Example 3.13. $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ *is convergent since it is absolutely convergent:* $\sum_{n=1}^{\infty} \left| \frac{(-1)^n}{n^2} \right| = \sum_{n=1}^{\infty} \frac{1}{n^2}$*, which is convergent.*

Geometric series $\sum_{n=1}^{\infty} ar^{n-1}$ give examples of series with positive and negative terms. For example for $a = 1$ and $r = -\frac{1}{2}$, we consider $\sum_{n=1}^{\infty} \left(-\frac{1}{2}\right)^{n-1}$. We already know that this is convergent since $|r| = \frac{1}{2} < 1$. Moreover note that it's also absolutely convergent since $\sum_{n=1}^{\infty} \left| \left(-\frac{1}{2}\right)^{n-1} \right| = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n-1}$, which is also convergent.

Example 3.14. *The series $\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \cdots$ is actually convergent (which we won't prove), but it is not absolutely convergent, since $\sum_{n=1}^{\infty} \left| \frac{(-1)^n}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n}$ which is the harmonic series, which is divergent (when a series converges, but doesn't absolutely converge, it's called* conditionally convergent*).*

Another test for convergence:

Theorem 3.9 (The Root Test)**.** *Let $\sum_{n=1}^{\infty} x_n$ be a series.*

1. *If there exists $N \in \mathbb{N}$ and $r < 1$ such that $n \geqslant N$ implies $|x_n|^{\frac{1}{n}} \leqslant r$, then $\sum_{n=1}^{\infty} x_n$ converges.*

2. *If there exists $N \in \mathbb{N}$ and $r > 1$ such that $n \geqslant N$ implies $|x_n|^{\frac{1}{n}} \geqslant r$, then $\sum_{n=1}^{\infty} x_n$ diverges.*

Proof. 1) $n \geqslant N$ implies $|x_n|^{\frac{1}{n}} \leqslant r$ so $|x_n| \leqslant r^n$. Since $a_n = r^n$ generates a convergent geometric series, by the Comparison Test $\sum_{n=1}^{\infty} |x_n|$ converges so $\sum_{n=1}^{\infty} x_n$ converges by Theorem 3.8.

2) $n \geqslant N$ implies $|x_n| \geqslant r^n$. So in particular, x_n does not converge to zero as $n \to \infty$, so by Theorem 3.1, $\sum_{n=1}^{\infty} x_n$ diverges. $\square$

N.B. If $\lim_{n\to\infty} |x_n|^{\frac{1}{n}}$ exists, then we can take this limit as the value of r in the Root Test (so we get convergence if this limit is < 1 and divergence if it's > 1). Note that if this limit is equal to zero, then we can take any $r < 1$ in case 1) of the Root Test (eg $r = 1/2$), so the series is convergent.

N.B. Similarly to the Ratio Test, $\lim_{n\to\infty} |x_n|^{\frac{1}{n}} = 1$ tells us nothing about convergence.

Useful tools here:

Theorem 3.10. *1.* $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$.

2. If $r > 0$ *then* $\lim_{n\to\infty} r^{\frac{1}{n}} = 1$.

3. Suppose that $(a_n)_n$ *is a sequence of positive numbers and* $\frac{a_{n+1}}{a_n} \to r$. *Then* $|a_n|^{\frac{1}{n}} \to r$.

Proof. 1) Let $b_n = n^{\frac{1}{n}} - 1$ for all $n \in \mathbb{N}$. By Theorem 2.3, it suffices to show that $b_n \to 0$. First note that $b_n \geqslant 0$ for all $n \in \mathbb{N}$. Moreover $1 + b_n = n^{\frac{1}{n}}$ implies $n = (1 + b_n)^n$. Using the first three terms from the binomial expansion of $(1 + b_n)^n$,

$$n = (1 + b_n)^n \geqslant 1 + nb_n + \frac{1}{2}n(n-1)b_n^2 > \frac{1}{2}n(n-1)b_n^2$$

Simplifying and rearranging $n > \frac{1}{2}n(n-1)b_n^2$ to $b_n^2 < 2/(n-1)$ whenever $n \geqslant 2$, we obtain $b_n \leqslant \sqrt{\frac{2}{n-1}} \to 0$, as required.

2) Suppose first that $r \geqslant 1$. Then for $n \geqslant r$ we have $1 \leqslant r^{\frac{1}{n}} \leqslant n^{\frac{1}{n}}$. By 1), this gives $\lim_{n\to\infty} r^{\frac{1}{n}} = 1$.

3) Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $n \geqslant N$ implies

$$\left|\frac{a_{n+1}}{a_n} - r\right| \leqslant \frac{\varepsilon}{2} \quad \text{i.e., } r - \frac{\varepsilon}{2} \leqslant \frac{a_{n+1}}{a_n} \leqslant r + \frac{\varepsilon}{2}.$$

Since $n \geqslant N$ implies

$$(a_n)^{\frac{1}{n}} = \left(\frac{a_n}{a_{n-1}}\frac{a_{n-1}}{a_{n-2}}\cdots\frac{a_{N+1}}{a_N}a_N\right)^{\frac{1}{n}},$$

we obtain

$$\left(\left(r-\frac{\varepsilon}{2}\right)^{n-N}a_N\right)^{\frac{1}{n}} \leqslant (a_n)^{\frac{1}{n}} = \left(\frac{a_n}{a_{n-1}}\frac{a_{n-1}}{a_{n-2}}\cdots\frac{a_{N+1}}{a_N}a_N\right)^{\frac{1}{n}} \leqslant \left(\left(r+\frac{\varepsilon}{2}\right)^{n-N}a_N\right)^{\frac{1}{n}},$$

so

$$\left(r-\frac{\varepsilon}{2}\right)^{1-\frac{N}{n}}(a_N)^{\frac{1}{n}} \leqslant (a_n)^{\frac{1}{n}} \leqslant \left(r+\frac{\varepsilon}{2}\right)^{1-\frac{N}{n}}(a_N)^{\frac{1}{n}}.$$

Since $\frac{N}{n} \to 0$ and $(a_N)^{\frac{1}{n}} \to 1$ as $n \to \infty$, there exists $M \geqslant N$ such that $n \geqslant M$ implies

$$r-\varepsilon \leqslant (a_n)^{\frac{1}{n}} \leqslant r+\varepsilon, \text{ i.e., } |a_n - r| \leqslant \varepsilon,$$

as required. □

Example 3.15. *Consider the series*

$$\sum_{n=0}^{\infty} 2^{(-1)^n-n} = 2+\frac{1}{4}+\frac{1}{2}+\frac{1}{16}+\frac{1}{8}+\frac{1}{64}+\cdots$$

Setting $x_n = 2^{(-1)^n-n}$, *the Ratio Test gives*

$$\frac{x_{n+1}}{x_n} = \frac{2^{(-1)^{n+1}-(n+1)}}{2^{(-1)^n-n}} = \frac{1}{2}2^{(-1)^{n+1}-(-1)^n} = \begin{cases} 2 & \textit{if } n \textit{ is odd,} \\ \frac{1}{8} & \textit{if } n \textit{ is even.} \end{cases}$$

So the Ratio Test tells us nothing. However, for the Root Test,

$$|x_n|^{\frac{1}{n}} = (2^{(-1)^n-n})^{\frac{1}{n}} = 2^{\frac{(-1)^n}{n}-1} = \frac{1}{2}\cdot 2^{\frac{(-1)^n}{n}}$$

Since for n *even, we have* $2^{\frac{1}{n}} \to 1$ *(see Theorem 3.10) and for* n *odd* $2^{-\frac{1}{n}} = \left(\frac{1}{2}\right)^{\frac{1}{n}} \to 1$ *(again, see Theorem 3.10), Theorem 2.3 implies that* $\lim_{n\to\infty}|x_n|^{\frac{1}{n}} = \frac{1}{2}$, *so the Root Test implies that the series converges.*

3.3 Power Series

Definition 3.3. *Given a sequence* $(a_n)_n$, *the series* $\sum_{n=1}^{\infty} a_n x^n$ *is called a* power series

In a power series there is a variable x. So if, for a particular x, $\sum_{n=1}^{\infty} a_n x^n$ converges, then the series can be interpreted as a sum and so the power series is a function of x. These come up in different areas of maths and physics, for example probability and combinatorics, as well as lots of areas of applied maths.

Determining convergence can be a difficult problem, but what always happens is either:

i) the power series converges for all $x \in \mathbb{R}$;

ii) the power series converges only at $x = 0$;

iii) the power series converges in some interval of x-values with centre zero.

Theorem 3.11. *Let $\sum_{n=1}^{\infty} a_n x^n$ be a power series and suppose $\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \beta$ and $R = \frac{1}{\beta}$ (if $\beta = 0$, then set $R = \infty$; if $\beta = \infty$, then set $R = 0$).*

1. *The power series converges for $|x| < R$.*
2. *The power series diverges for $|x| > R$.*

R is called the *radius of convergence* of the power series. Note that 1) is vacuous if $R = 0$ and 2) is vacuous if $R = \infty$.

Proof. Let $b_n = a_n x^n$ for all $n \in \mathbb{N}$, so our series is that generated by $(b_n)_n$. Then in order to apply the Root Test, we compute

$$|b_n|^{\frac{1}{n}} = |a_n x^n|^{\frac{1}{n}} = |a_n|^{\frac{1}{n}} |x| \to \beta |x| \text{ as } n \to \infty.$$

We will input the value $\beta|x|$ into the Root Test.

Case 1: a) $\beta = 0$ (so $R = \infty$, and necessarily $|x| < R$). Then the Root Test implies that $\sum_{n=1}^{\infty} a_n x^n$ is convergent.

b) $\beta > 0$ and $\beta|x| < 1$ (so $|x| < R$). Again the Root Test implies that $\sum_{n=1}^{\infty} a_n x^n$ is convergent.

Case 2: a) $\beta = \infty$ (so $R = 0$). Then the Root Test implies that $\sum_{n=1}^{\infty} a_n x^n$ is divergent.

b) β is finite and $\beta|x| > 1$ (so $|x| > R$). Again the Root Test implies that $\sum_{n=1}^{\infty} a_n x^n$ is divergent. $\square$

Note that if $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ exists, then it equals β and $\lim_{n\to\infty} |a_n|^{\frac{1}{n}} = \beta$ (Exercise), so it's often easier to check this ratio, rather than finding the root directly.

Example 3.16. *Consider* $\sum_{n=1}^{\infty} \frac{1}{n!} x^n$.

Then $a_n = \frac{1}{n!}$. *Hence*

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

Hence $\beta = 0$ *and* $R = \infty$, *so the radius of convergence is* ∞, *i.e., the series converges for all* $x \in \mathbb{R}$ *(actually to* e^x*).*

Example 3.17. *Consider* $\sum_{n=1}^{\infty} \frac{1}{n} x^n$.

Letting $a_n = \frac{1}{n}$,

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\frac{1}{n+1}}{\frac{1}{n}} = \frac{n}{n+1} \to 1,$$

so $\beta = 1$ *and* $R = 1$ *- the radius of convergence is 1. This means that the power series converges for all* $x \in (-1, 1)$, *and diverges for* $x \in (-\infty, -1) \cup (1, \infty)$. *In fact, Theorem 3.2 and Example 3.13 imply that the series converges if and only if* $x \in [-1, 1)$.

Example 3.18. *For* $\sum_{n=1}^{\infty} \frac{2^n}{n^2} x^n$, $a_n = \frac{2^n}{n^2}$. *We calculate*

$$\left(\frac{2^n}{n^2}\right)^{\frac{1}{n}} = \frac{2}{n^{\frac{2}{n}}} = \frac{2}{(n^{\frac{1}{n}})^2} \to 2$$

(Note that we use $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$ *from Theorem 3.10 here.) So* $\beta = 2$ *and the radius of convergence is* $\frac{1}{2}$.

Chapter 4

Cauchy Sequences

It's important to know if a sequence has a limit or not. The Monotone Convergence Theorem is an example of a result which guarantees the existence of a limit, but its main drawback is that the sequence has to be monotone. We will show that any 'Cauchy sequence' has a limit.

Definition 4.1. *A sequence* $(x_n)_n$ *of real numbers is called a* Cauchy sequence *if for all* $\varepsilon > 0$*, there exists* $N \in \mathbb{N}$ *such that* $\forall n, m \in \mathbb{N}$*,* $n, m \geqslant N$ *implies*

$$|x_n - x_m| \leqslant \varepsilon.$$

Example 4.1. $(x_n)_n = \left(\frac{1}{n}\right)_n$ *is Cauchy.*

Proof. Let $\varepsilon > 0$. Then setting $N \in \mathbb{N}$ such that $N \geqslant \frac{2}{\varepsilon}$, $n, m \geqslant N$ implies that

$$|x_m - x_n| = \left|\frac{1}{n} - \frac{1}{m}\right| \leqslant \frac{1}{n} + \frac{1}{m} \leqslant \frac{1}{N} + \frac{1}{N} = \frac{2}{N} \leqslant \varepsilon,$$

so $(x_n)_n$ is Cauchy. □

Example 4.2. $(x_n)_n = \left(\frac{n}{1+n}\right)_n$ *is Cauchy.*

Proof. Let $\varepsilon > 0$. Then set $N \in \mathbb{N}$ such that $N \geqslant \frac{2}{\varepsilon}$. Then for $n, m \geqslant N$, assuming $n \leqslant m$ implies that

$$\begin{aligned}|x_m - x_n| &= \left|\frac{n}{1+n} - \frac{m}{1+m}\right| = \left|\frac{n(1+m) - m(1+n)}{(1+n)(1+m)}\right| \\ &= \left|\frac{n-m}{(1+n)(1+m)}\right| \leqslant \frac{|n-m|}{nm} = \frac{m-n}{nm} = \frac{1}{n} - \frac{1}{m} \\ &\leqslant \frac{1}{n} + \frac{1}{m} \leqslant \frac{1}{N} + \frac{1}{N} = \frac{2}{N} \leqslant \varepsilon.\end{aligned}$$

The case where $n > m$ follows similarly. So $\left(\frac{n}{1+n}\right)_n$ is Cauchy. □

Example 4.3. $(x_n)_n = \left(1 + \frac{1}{2} + \cdots \frac{1}{n}\right)_n$ *is not Cauchy.*

Proof. We'll show that $(x_n)_n$ satisfies the negation of the definition of Cauchy sequence. That is, we must show that

$$\exists \varepsilon > 0 : \ \forall N \in \mathbb{N}, \ \exists m, n \in \mathbb{N} \text{ s.t. } n, m \geqslant N \text{ and } |x_n - x_m| > \varepsilon. \quad (*)$$

We start by fixing some $N \in \mathbb{N}$. Put $n = N$ and $m = 2N$, so clearly $n, m \geqslant N$. Then

$$\begin{aligned}|x_m - x_n| &= |x_N - x_{2N}| \\ &= \left|\left(1 + \frac{1}{2} + \cdots + \frac{1}{N}\right) - \left(1 + \frac{1}{2} + \cdots + \frac{1}{N} + \frac{1}{N+1} + \cdots + \frac{1}{2N}\right)\right| \\ &= \frac{1}{N+1} + \cdots + \frac{1}{2N} \geqslant \frac{1}{2N} + \cdots + \frac{1}{2N} = N\frac{1}{2N} = \frac{1}{2} > \frac{1}{4}.\end{aligned}$$

Since N was arbitrary here, this calculation implies that we can put $\varepsilon = \frac{1}{4}$ into $(*)$ to derive that $\left(1 + \frac{1}{2} + \cdots \frac{1}{n}\right)_n$ is not Cauchy. □

Note that in the first two examples here the sequence was convergent as well as being Cauchy, while in the third example, the sequence was neither convergent nor Cauchy. This suggests that being convergent and Cauchy are related, an idea which we develop below.

Proposition 4.1. *Every convergent sequence is Cauchy.*

Proof. Suppose that $(x_n)_n$ is a convergent sequence with limit x. Let $\varepsilon > 0$. Since $x_n \to x$ there exists $N \in \mathbb{N}$ such that $n \geqslant N$ implies $|x_n - x| \leqslant \frac{\varepsilon}{2}$. So for $n, m \in \mathbb{N}$, $n, m \geqslant N$ implies

$$|x_n - x_m| = |x_n - x + x - x_m| \leqslant |x_n - x| + |x - x_m| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required. $\square$

Lemma 4.2. *Any Cauchy sequence is bounded.*

Proof. Let $(x_n)_n$ be a Cauchy sequence. Then there exists $N \in \mathbb{N}$ such that $n, m \geqslant N$ implies $|x_n - x_m| \leqslant 1$. Hence $n \geqslant N$ implies

$$|x_n| = |x_n - x_N + x_N| \leqslant |x_n - x_N| + |x_N| \leqslant 1 + |x_N|.$$

Therefore, for

$$M := \max\{|x_1|, |x_2|, \ldots, |x_N|, |x_N| + 1\},$$

$|x_n| \leqslant M$ for all $n \in \mathbb{N}$. $\square$

Lemma 4.3. *Let $(x_n)_n$ be a Cauchy sequence. If $(x_n)_n$ has a subsequence $(x_{m_k})_k$ that converges to some real number $x \in \mathbb{R}$, i.e., $x_{m_k} \to x$ as $k \to \infty$, then the original sequence $(x_n)_n$ also converges to x, i.e., $x_n \to x$ as $n \to \infty$.*

Proof. Let $\varepsilon > 0$. Since $(x_n)_n$ is Cauchy, there exists $N_1 \in \mathbb{N}$ such that $n \geqslant N_1$ implies

$$|x_n - x_m| \leqslant \frac{\varepsilon}{2}.$$

Since, moreover, $x_{m_k} \to x$, there exists $N_2 \in \mathbb{N}$ such that $k \geqslant N_2$ implies

$$|x_{m_k} - x| \leqslant \frac{\varepsilon}{2}.$$

Let $N := \max\{N_1, N_2\}$. Then for $n \in \mathbb{N}$, $n \geqslant N$ implies

$$|x_n - x| = |x_n - x_{m_n} + x_{m_n} - x| \leqslant |x_n - x_{m_n}| + |x_{m_n} - x| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(Here we used that $m_n \geqslant n \geqslant N$ to estimate $|x_n - x_{m_n}|$.) $\square$

Theorem 4.4 (The General Principle of Convergence)**.** *Let $(x_n)_n$ be a sequence of real numbers. The following are equivalent.*

1. *$(x_n)_n$ is convergent.*
2. *$(x_n)_n$ is Cauchy.*

Proof. $(1 \Rightarrow 2)$: This is Proposition 4.1.

$(2 \Rightarrow 1)$: Let $(x_n)_n$ be a Cauchy sequence. Lemma 4.2 implies that $(x_n)_n$ is bounded. So the Bolzano-Weierstrass Theorem implies that there exists a subsequence $(x_{m_k})_k$ which converges to some limit x. Finally Lemma 4.3 implies that $(x_n)_n$ itself is also convergent (converging to x). □

N.B. Note that this theorem is also known as the Cauchy Convergence Criterion.

Example 4.4. *Let $(x_n)_n$ be a sequence and assume there is $A > 0$ such that*

$$|x_{n+1} - x_n| \leqslant \frac{A}{n^2} \text{ for all } n \in \mathbb{N}.$$

Then $(x_n)_n$ is convergent.

Proof. By the General Principle of Convergence, it suffices to show that $(x_n)_n$ is a Cauchy sequence. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\frac{1}{N-1} \leqslant \frac{\varepsilon}{A}$. Then $n \geqslant m \geqslant N$ implies

$$\begin{aligned}
|x_n - x_m| &\leqslant |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \cdots + |x_{m+2} - x_{m+1}| + |x_{m+1} - x_m| \\
&\leqslant A\left(\frac{1}{n^2} + \frac{1}{(n+1)^2} + \cdots + \frac{1}{(m+1)^2} + \frac{1}{m^2}\right) \\
&\leqslant A\left(\frac{1}{(n-1)n} + \frac{1}{n(n+1)} + \cdots + \frac{1}{m(m+1)} + \frac{1}{(m-1)m}\right) \\
&\leqslant A\left(\left(\frac{1}{n-2} - \frac{1}{n-1}\right) + \left(\frac{1}{n-3} - \frac{1}{n-2}\right) + \cdots \right. \\
&\qquad\qquad \left. \cdots + \left(\frac{1}{m} - \frac{1}{m+1}\right) + \left(\frac{1}{m-1} - \frac{1}{m}\right)\right) \\
&= A\left(\frac{1}{m-1} - \frac{1}{n-1}\right) \leqslant \frac{A}{n-1} \leqslant \frac{A}{N-1} \leqslant \varepsilon.
\end{aligned}$$

Hence $(x_n)_n$ is a Cauchy, so $(x_n)_n$ is convergent. □

Example 4.5. *The sequence $(x_n)_n$ where*

$$x_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n$$

is convergent. The number $\gamma = \lim_{n\to\infty} 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n$ is called Euler's constant. *Whether it's rational or not is a long-standing question in maths.*

Proof. By Example 4.4, convergence of $(x_n)_n$ will follow if $|x_{n+1} - x_n| \leqslant \frac{2}{n^2}$ for all $n \in \mathbb{N}$. We compute this using the fact that $\log x = \int_1^x \frac{1}{t}\ dt$:

$$
\begin{aligned}
|x_{n+1} - x_n| &= \left|\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \frac{1}{n+1} - \log(n+1)\right) - \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right)\right| \\
&= \left|\frac{1}{n+1} - (\log(n+1) - \log n)\right| = \left|\frac{1}{n+1} - \log\left(\frac{n+1}{n}\right)\right| \\
&= \left|\frac{1}{n+1} - \frac{1}{n} + \frac{1}{n} - \log\left(1 + \frac{1}{n}\right)\right| \leqslant \left|\frac{1}{n+1} - \frac{1}{n}\right| + \left|\frac{1}{n} - \log\left(1 + \frac{1}{n}\right)\right| \\
&= \frac{1}{n(n+1)} + \left|\int_1^{1+\frac{1}{n}} 1\ dt - \int_1^{1+\frac{1}{n}} \frac{1}{t}\ dt\right| \leqslant \frac{1}{n^2} + \left|\int_1^{1+\frac{1}{n}} 1 - \frac{1}{t}\ dt\right| \\
&= \frac{1}{n^2} + \left|\int_1^{1+\frac{1}{n}} \frac{t-1}{t}\ dt\right| \leqslant \frac{1}{n^2} + \left|\int_1^{1+\frac{1}{n}} t - 1\ dt\right| \\
&\leqslant \frac{1}{n^2} + \left|\int_1^{1+\frac{1}{n}} \frac{1}{n}\ dt\right| \leqslant \frac{1}{n^2} + \frac{1}{n^2} = \frac{2}{n^2},
\end{aligned}
$$

as required. □

N.B. We could also have used the theory of Cauchy sequences to get further information on series via Cauchy properties of partial sums (this leads to proofs of Theorem 3.8 and the statement in Example 3.14).

Chapter 5

Continuous Functions

Recall that given sets A and B, $f : A \to B$ is a *function* if for each $x \in A$ there exists a unique value $y' inB$ such that $f(x) = y$ (A is the *domain* of f and $f(A) = \{y \in B : \exists x \in A \text{ s.t. } f(x) = y\}$ is the *range* of f).

Given an interval $I \subseteq \mathbb{R}$, a rough description of a function $f : I \to \mathbb{R}$ being continuous is that it can be graphed without removing the pen from the paper. Here we'll give a more rigorous treatment.

Definition 5.1. *Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$ be a function on A. Then f is said to be* continuous at a point $x_0 \in A$ *if*

$$\forall \varepsilon > 0 \exists \delta > 0 \; s.t. \; \forall x \in A, |x - x_0| \leqslant \delta \Rightarrow |f(x) - f(x_0)| \leqslant \varepsilon.$$

- We say that f is *discontinuous at* $x_0 \in A$ if f is not continuous at x_0.
- We say that f is *continuous* if it is continuous at all $x_0 \in A$.
- We say that f is *discontinuous* if it is not continuous.

Theorem 5.1. *Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$ be a function. Then the following are equivalent:*

1. *f is continuous at x_0.*
2. *If $(x_n)_n$ is any sequence in A with $x_n \to x_0$, then $f(x_n) \to f(x_0)$.*

Proof. **(**1) $\Rightarrow$ 2)**):** Suppose that $(x_n)_n \subseteq A$ has $x_n \to x_0$. To show that $f(x_n) \to f(x_0)$, let $\varepsilon > 0$. Since f is continuous at x_0, there exists $\delta > 0$ such that $|x - x_0| \leqslant \delta$ implies $|f(x) - f(x_0)| \leqslant \varepsilon$. Since $x_n \to x_0$, there exists $N \in \mathbb{N}$ such that $|x_n - x_0| \leqslant \delta$ for all $n \geqslant N$. So if $n \in \mathbb{N}$ has $n \geqslant N$, then $|f(x_n) - f(x_0)| \leqslant \varepsilon$. Therefore 2) follows.

(2) $\Rightarrow$ 1)**):** To obtain a contradiction, we assume 2), but assume that 1) is false, i.e., [negating $\forall \varepsilon > 0\ \exists \delta > 0$ s.t. $\forall x \in A$, $|x - x_0| \leqslant \delta \Rightarrow |f(x) - f(x_0)| \leqslant \varepsilon$],

$$\exists \varepsilon > 0 \text{ s.t. } \forall \delta > 0\ \exists x \in A \text{ s.t. } |x - x_0| \leqslant \delta \text{ and } |f(x) - f(x_0)| > \varepsilon. \quad (*)$$

Suppose that $\varepsilon > 0$ is such that $(*)$ holds, i.e.,

$$\forall \delta > 0\ \exists x \in A \text{ s.t. } |x - x_0| \leqslant \delta \text{ and } |f(x) - f(x_0)| > \varepsilon.$$

In particular, for each $n \in \mathbb{N}$, setting $\delta = \frac{1}{n}$ gives some $x_n \in A$ such that

$$|x_n - x_0| \leqslant \frac{1}{n} \text{ and } |f(x_n) - f(x_0)| > \varepsilon.$$

But this implies that we have a sequence $(x_n)_n$ such that $x_n \to x_0$, but $(f(x_n))_n$ doesn't converge to $f(x_0)$, which contradicts 2). So f must in fact be continuous at x_0. $\square$

This theorem means that we now have two equivalent definitions of continuity:

- the original one, Definition 5.1, which is called the *ε-δ definition of continuity*;
- the condition "for all sequences $(x_n)_n \subset A$ such that $x_n \to x_0$, we have $f(x_n) \to f(x_0)$", which is called the *sequential definition of continuity.*

Example 5.1. *Let $c \in \mathbb{R}$ and define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = c$ for all $x \in \mathbb{R}$. Then f is continuous.*

Proof using the ε-δ definition of continuity. Let $x_0 \in \mathbb{R}$. Let $\varepsilon > 0$ and set $\delta = 300$. Then $x \in \mathbb{R}$ and $|x - x_0| \leqslant \delta$ implies

$$|f(x) - f(x_0)| = |c - c| = 0 < \varepsilon,$$

so f is continuous at x_0. Since $x_0 \in \mathbb{R}$ was arbitrary, f is continuous. $\square$

Example 5.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be the identity map, i.e., $f(x) = x$ for all $x \in \mathbb{R}$. Then f is continuous.*

Proof using the ε-δ definition of continuity. Let $x_0 \in \mathbb{R}$. Let $\varepsilon > 0$. Then for $\delta = \epsilon$, for all $x \in \mathbb{R}$, $|x - x_0| \leqslant \delta$ implies

$$|f(x) - f(x_0)| = |x - x_0| \leqslant \delta = \varepsilon,$$

so f is continuous at x_0. Since $x_0 \in \mathbb{R}$ was arbitrary, f is continuous. □

Example 5.3. *Define $f : \mathbb{R} \to \mathbb{R}$ by*

$$f(x) = 2x^2 + 1.$$

Then f is continuous.

Proof using the ε-δ definition of continuity. Let $x_0 \in \mathbb{R}$. Let $\varepsilon > 0$. Note that if $|x-x_0| \leqslant 1$ implies $|x| \leqslant |x_0|+1$. Now choose $\delta = \min\left\{1, \frac{\varepsilon}{2(2|x_0|+1)}\right\}$. Then $x \in \mathbb{R}$ and $|x - x_0| \leqslant \delta$ implies

$$\begin{aligned} |f(x) - f(x_0)| &= |(2x^2 + 1) - (2x_0^2 + 1)| = 2|x^2 - x_0^2| = 2|(x - x_0)(x + x_0)| \\ &= 2|x - x_0||x + x_0| \leqslant |x - x_0| 2\left(|x| + |x_0|\right) \\ &\leqslant |x - x_0| 2\left(|x_0| + 1 + |x_0|\right) \leqslant \delta 2\left(2|x_0| + 1\right) \leqslant \varepsilon, \end{aligned}$$

so f is continuous at x_0. Since $x_0 \in \mathbb{R}$ was arbitrary, f is continuous. □

Proof using the sequential definition of continuity. Let $x_0 \in \mathbb{R}$. Let $(x_n)_n$ be a sequence in $\mathbb{R}$ such that $x_n \to x_0$. Then by Theorem 2.3,

$$f(x_n) = 2x_n^2 + 1 \to 2x_0^2 + 1,$$

so f is continuous. □

Example 5.4. *Let $f : (0, \infty) \to \mathbb{R}$ be given by*

$$f(x) = \frac{1}{x^2}.$$

Then f is continuous.

Proof using the ε-δ definition of continuity. Let $x_0 \in (0,\infty)$. Let $\varepsilon > 0$. First note that

$$|f(x)-f(x_0)| = \left|\frac{1}{x^2} - \frac{1}{x_0^2}\right| = \left|\frac{x^2 - x_0^2}{x^2x_0^2}\right| = \frac{|x-x_0||x+x_0|}{x^2x_0^2} \leqslant \frac{|x-x_0|\,(|x|+|x_0|)}{x^2x_0^2}. \quad (*)$$

$\delta = \min\left\{\frac{x_0}{2}, \frac{x_0^3\varepsilon}{10}\right\}$. Then $x \in (0,\infty)$ and $|x-x_0| \leqslant \delta$ implies $-\delta \leqslant x - x_0 \leqslant \delta$, so

$$x \geqslant x_0 - \delta \geqslant x_0 - \frac{x_0}{2} = \frac{x_0}{2}$$

and

$$x \leqslant x_0 + \delta \leqslant x_0 + \frac{x_0}{2} = \frac{3x_0}{2}.$$

So for all $x \in (0,\infty)$, by $(*)$, $|x-x_0| \leqslant \delta$ implies

$$|f(x)-f(x_0)| \leqslant \frac{|x-x_0|\,(x+x_0)}{x^2x_0^2} \leqslant \frac{\delta\left(\frac{3x_0}{2}+x_0\right)}{\left(\frac{x_0^2}{2}\right)x_0^2} = \frac{\delta\left(\frac{5x_0}{2}\right)}{\left(\frac{x_0^4}{4}\right)} = \delta\frac{10}{x_0^3} \leqslant \varepsilon.$$

so f is continuous at each $x_0 \in (0,\infty)$ and hence continuous. □

N.B. The sequential definition can also be used in conjunction with Theorem 2.3 here (Exercise).

Example 5.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

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

Then f is continuous at any $x_0 \in \mathbb{R} \setminus \{0\}$, but discontinuous at 0.

Proof. We'll first show that f is discontinuous at 0. Let $(x_n)_n = \left(-\frac{1}{n}\right)_n$. Then $x_n \to 0$. Also since $x_n < 0$, $f(x_n) = -1$ for all $n \in \mathbb{N}$: but this means that $f(x_n)$ does not tend to $1 = f(x_0)$, so f is not continuous at 0.

Now let $x_0 \in \mathbb{R} \setminus \{0\}$. Let $\varepsilon > 0$. Then since $x_0 \neq 0$, $\delta := \frac{|x_0|}{2} > 0$. So for all $x \in \mathbb{R}$, if $|x-x_0| \leqslant \delta = \frac{|x_0|}{2}$ then $-\frac{|x_0|}{2} \leqslant x - x_0 \leqslant \frac{|x_0|}{2}$, which implies

$$x_0 - \frac{|x_0|}{2} \leqslant x \leqslant \frac{|x_0|}{2}.$$

Hence

$$\begin{cases} x_0 > 0 \Rightarrow |x_0| = x_0 \text{ and } x \geqslant \frac{x_0}{2} > 0, \text{ so } f(x) = 1 = f(x_0), \\ x_0 < 0 \Rightarrow |x_0| = -x_0 \text{ and } x \leqslant -\frac{x_0}{2} < 0, \text{ so } f(x) = -1 = f(x_0), \end{cases}$$

so in either case $|f(x) - f(x_0)| = 0 < \varepsilon$. □

Example 5.6. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

$$f(x) = \begin{cases} x & \textit{if } x \in \mathbb{Q}, \\ 0 & \textit{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Then f is continuous only at 0.

Proof. We'll first show that f is continuous at 0. Let $\varepsilon > 0$. Then for $\delta = \varepsilon$, for all $x \in \mathbb{R}$ with $|x - x_0| \leqslant \delta$, we have

$$|f(x) - f(x_0)| = |f(x) - 0| = |f(x)| \leqslant |x| \leqslant \delta \leqslant \varepsilon,$$

so f is continuous at 0.

We next show that f is discontinuous at $x_0 \neq 0$.

Case 1: $x_0 \in \mathbb{R} \setminus \mathbb{Q}$. Then $f(x_0) = 0$. Since the rationals are dense in $\mathbb{R}$ there exists $x_n \to x_0$ such that $x_n \in \mathbb{Q}$ for all $n \in \mathbb{N}$. Hence $f(x_n) = x_n$, but $\lim_{n\to\infty} f(x_n) = x_0 \neq f(x_0) = 0$. Hence f is not continuous at x_0.

Case 2: $x_0 \in \mathbb{Q}$. Then $f(x_0) = x_0$. Since for $x_n = x + \frac{\sqrt{2}}{n}$, x_n is irrational and $f(x_n) = 0$ for all $n \in \mathbb{N}$. But since $x_n \to x_0$ and $(f(x_n))_n$ does not converge to $f(x_0) = x_0$, f is not continuous at x_0. □

As we've seen before, it can be useful to break complicated mathematical objects down into more basic building blocks and thus understand the more complicated object. For example, taking

$$f(x) = x^2 \qquad g(x) = x^3$$

as our building blocks (which we can fairly easily prove are continuous), we can make

$$f+g : x \mapsto x^2+x^3, \quad f-g : x \mapsto x^2-x^3, \quad f\cdot g : x \mapsto f(x)g(x) = x^2x^3 = x^5,$$

$$f \circ g : x \mapsto f(g(x)) = f(x^3) = (x^3)^2 = x^6.$$

As one might expect, continuity is preserved by the operations above (namely addition, subtraction, multiplication, composition), which is the content of the next two results.

Theorem 5.2. *Let $A \subseteq \mathbb{R}$ and $x_0 \in A$. Suppose that $f, g : A \to \mathbb{R}$ are functions which are continuous at x_0 and let $\lambda \in \mathbb{R}$. Define*

- $f+g$ *by* $(f+g)(x) = f(x) + g(x)$,
- fg *by* $(fg)(x) = f(x)g(x)$,
- λf *by* $(\lambda f)(x) = \lambda f(x)$,
- $\min(f,g)$ *by* $\min(f,g)(x) = \min\{f(x), g(x)\}$,
- $\max(f,g)$ *by* $\max(f,g)(x) = \max\{f(x), g(x)\}$,
- $|f|$ *by* $|f|(x) = |f(x)|$,
- *If* $g(x) \neq 0$ *for all* $x \in A$ *then* $\frac{f}{g}$ *is given by* $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$.

Then

1. $f+g$ *is continuous at* x_0,
2. fg *is continuous at* x_0,
3. λf *is continuous at* x_0,
4. $\min(f,g)$ *is continuous at* x_0,
5. $\max(f,g)$ *is continuous at* x_0,
6. $|f|$ *is continuous at* x_0,
7. *If* $g(x) \neq 0$ *for all* $x \in A$ *then* $\frac{f}{g}$ *is continuous at* x_0.

Proof. 1) Let $(x_n)_n \subseteq A$ have $x_n \to x_0$. Then since f and g are continuous at x_0, $f(x_n) \to f(x_0)$ and $g(x_n) \to g(x_0)$. So by Theorem 2.3,

$$(f+g)(x_n) = f(x_n) + g(x_n) \to f(x_0) + g(x_0) = (f+g)(x_0),$$

so $f+g$ is continuous at x_0.

2), 3), 6) and 7) follow similarly.

5) Since
$$\max(f,g) = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|,$$
the result follows by 1), 3) and 6).

4) Since
$$\min(f,g) = \frac{1}{2}(f+g) - \frac{1}{2}|f-g|,$$
the result follows by 1), 3) and 6). □

One consequence of this theorem is that since $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ is continuous (Exercise), any polynomial $p : \mathbb{R} \to \mathbb{R}$ is continuous.

Theorem 5.3. *Let $A, B \subseteq \mathbb{R}$ and $x_0 \in A$. Suppose that $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be functions with $f(x_0) \in B$. If f is continuous at x_0 and g is continuous at $f(x_0)$, then $g \circ f$ is continuous at x_0.*

Proof. Let $(x_n)_n \subseteq A$ be a sequence with $x_n \to x_0$. Since f is continuous at x_0, the sequence $(f(x_n))_n$ has $f(x_n) \to f(x_0)$. Since g is continuous at $f(x_0)$, this means
$$(g \circ f)(x_n) = g(f(x_n)) \to g(f(x_0)) = (g \circ f)(x_0),$$
i.e., $(g \circ f)(x_n) \to (g \circ f)(x_0)$, so $g \circ f$ is continuous at x_0. □

We'll next derive some major theorems on continuous functions which are intuitively obvious, but require quite sophisticated proofs. We start with a definition.

Definition 5.2. *A function $f : A \to \mathbb{R}$ is called* bounded *if there exists $M > 0$ such that*
$$|f(x)| \leqslant M \text{ for all } x \in A.$$

Theorem 5.4 (The Extreme Value Theorem)**.** *Given $a < b$, let $f : [a,b] \to \mathbb{R}$ be a continuous function. Then*

1. *f is bounded;*
2. *f attains its maximum, i.e., there exists $x_0 \in [a,b]$ such that*
$$f(x) \leqslant f(x_0) \text{ for all } x \in [a,b];$$

3. f attains its minimum, i.e., there exists $y_0 \in [a,b]$ such that

$$f(x) \geqslant f(y_0) \text{ for all } x \in [a,b].$$

Proof. 1) Assume, for a contradiction, that f is unbounded above. So for each $n \in \mathbb{N}$ there exists $x_n \in [a,b]$ such that $f(x_n) > n$. Since $(x_n)_n \subseteq [a,b]$, it is a bounded sequence, so the Bolzano-Weierstrass Theorem implies that there is a subsequence $(x_{m_k})_k$ and $c \in \mathbb{R}$ such that $x_{m_k} \to c$ as $k \to \infty$. Since $x_{m_k} \in [a,b]$ for all $k \in \mathbb{N}$, this means that $c \in [a,b]$. Since f is continuous in $[a,b]$, in particular it's continuous at c, so $x_{m_k} \to c$ implies $f(x_{m_k}) \to f(c)$. Hence $(f(x_{m_k}))_k$ is bounded sequence by Theorem 2.2. But this contradicts the fact that $f(x_{m_k}) > m_k$ for all k, so f is bounded above. The proof that f is bounded below follows similarly.

2) Let $M := \sup\{f(x) : x \in [a,b]\}$. We'll show that there exists $x_0 \in [a,b]$ such that $f(x_0) = M$. By definition of sup, for each $n \in \mathbb{N}$ there exists $x_n \in [a,b]$ such that

$$M - \frac{1}{n} \leqslant f(x_n) \leqslant M.$$

(see eg TS1 Q9). This implies that

$$f(x_n) \to M \text{ as } n \to \infty \quad (*).$$

The Bolzano-Weierstrass Theorem implies that for the sequence $(x_n)_n$ there exists a convergent subsequence $(x_{m_k})_k$ with limit, say $x_0 \in [a.b]$. Continuity of f means that $x_{m_k} \to x_0$ implies that

$$f(x_{m_k}) \to f(x_0) \text{ as } k \to \infty \quad (**).$$

Combining $(*)$ and $(**)$ shows that $f(x_0) = M$, so 2) is proved.

3) Apply 2) to the function $-f$ (Exercise). □

N.B. If we allowed f to be defined on an open interval (a,b) (or half-open), then the conclusions of the Extreme Value Theorem may fail, as the next examples show.

Example 5.7. *Define $f : (0,1) \to \mathbb{R}$ by*

$$f(x) = \frac{1}{x}.$$

Then f is continuous (Exercise), but not bounded (eg given any $M \geqslant 0$, there exists $n \in \mathbb{N}$ such that $n > M$, so $f\left(\frac{1}{n}\right) = n > M$.).

Example 5.8. *Define* $f : (0,1) \to \mathbb{R}$ *by* $f(x) = x$ *for all* $x \in (0,1)$. *Then* f *is bounded and continuous, but it does not have a maximum since* $\{f(x) : x \in (0,1)\} = (0,1)$, *so it certainly can't attain its maximum.*

The following is the main result of this chapter.

Theorem 5.5 (The Intermediate Value Theorem)**.** *Let* $I \subseteq \mathbb{R}$ *be an interval and let* $f : I \to \mathbb{R}$ *be a continuous function. If* $a, b \in I$ *with* $a < b$ *and* y *lies between* $f(a)$ *and* $f(b)$ *(i.e., either* $f(a) < y < f(b)$ *or* $f(b) < y < f(a)$*), then there exists* $x \in (a,b)$ *such that*

$$f(x) = y.$$

Proof. Assume that $f(a) < y < f(b)$ (the other case follows similarly). Let

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

Since $a \in S$, $S \neq \emptyset$. Set $x_0 := \sup S$. The idea is that $f(x_0$ should be y.

- Proof that $f(x_0) \leqslant y$: for each $n \in \mathbb{N}$, since $x_0 - \frac{1}{n} < x_0$, we know that $x_0 - \frac{1}{n}$ is not an upper bound on S, so there exists $s_n \in S$ such that

  $$x_0 - \frac{1}{n} < s_n \leqslant x_0.$$

  This means that $s_n \to x_0$ and the continuity of f implies that

  $$f(s_n) \to f(x_0).$$

  Since $s_n \in S$, $f(s_n) < y$ for each n, so any limit of $(f(s_n))_n$ must be $\leqslant y$, i.e., $f(x_0) \leqslant y$.

- Proof that $f(x_0) \geqslant y$: For each $n \in \mathbb{N}$, let $t_n := \min\left\{b, x_0 + \frac{1}{n}\right\}$. So $t_n \in [a,b]$, $t_n \to x_0$ and by continuity,

  $$f(t_n) \to f(x_0). \quad (*).$$

  By definition for each $n \in \mathbb{N}$, $t_n \notin S$, so $f(t_n) \geqslant y$. Hence this inequality passes to the limit, i.e., $(*)$ implies $f(x_0) \geqslant y$.

In conclusion, $f(x_0) = y$, as required. Note that $x_0 \in [a,b]$, but $f(a) < f(x_0) < f(b)$, so in fact $x_0 \in (a,b)$. $\square$

Corollary 5.6. *Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a continuous function. Then the range of f: i.e., $f(I) = \{f(x) : x \in I\}$ is either an interval or a single point.*

Proof. If there are two distinct points in $f(I)$ (say $f(a) \neq f(b)$), the IVT guarantees that all points between these are also in $f(I)$ (eg $f(a), f(b) \in f(I)$ and $f(a) > f(b)$ implies all $y \in (f(b), f(a))$, $y \in f(I)$). $\square$

Note that if I is closed then $f(I) = [\inf f(I), \sup f(I)]$.

Example 5.9. *Suppose that $f : [0,1] \to [0,1]$ is a continuous function. Then there exists a 'fixed point', i.e., there exists $x_0 \in [0,1]$ such that $f(x_0) = x_0$.*

Proof. Let $g(x) = f(x) - x$. By Theorem 5.2, g is also continuous on $[0,1]$. Now notice that

$$\left.\begin{array}{l} g(0) = f(0) - 0 = f(0) \geqslant 0 \\ g(1) = f(1) - 1 \leqslant 1 - 1 = 0 \end{array}\right\} \text{ so } 0 \in [g(1), g(0)].$$

If $g(0) = 0$ or $g(1) = 0$ then 0 or 1 are fixed points, respectively. Assume that $g(0) > 0 > g(1)$. Then the IVT implies that there exists $x_0 \in (0,1)$ such that $g(x_0) = 0$, i.e., $f(x_0) - x_0 = 0$, so $f(x_0) = x_0$, as required. $\square$

Example 5.10. *If $y > 0$ and $m \in \mathbb{N}$ then y has a positive mth root.*

Proof. Suppose $y \neq 1$, since then the required root is 1. By Theorem 5.2, $f(x) = x^m$ is continuous. Now let

$$b = \begin{cases} 1 & \text{if } y < 1, \\ y & \text{if } y > 1. \end{cases}$$

Then $0 < y < b^m$, i.e., $f(0) < y < f(b)$. So the IVT implies that there exists $x \in (0,b)$ such that $f(x) = y$, i.e., $x^m = y$, i.e., $x^m = y$, so x is an mth root of y. $\square$

5.1 Limits

Suppose that $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$ is a function. It will be convenient later to use the notation, for $a \in \mathbb{R}$,

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

This means that there exists $\ell \in \mathbb{R}$ such that as x becomes arbitrarily close to, but not equal to, a, $f(x)$ becomes arbitrarily close to ℓ. More formally:

Definition 5.3. $\lim_{x \to a} f(x)$ *exists and equals* $\ell \in \mathbb{R}$ *if for all* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that* $x \in A \setminus \{a\}$ *with* $|x - a| \leqslant \delta$ *implies*

$$|f(x) - \ell| \leqslant \varepsilon.$$

Example 5.11. *Let* $f : [0, \infty) \setminus \{1\} \to \mathbb{R}$ *be defined by*

$$f(x) = \frac{x^2 - 1}{x^2 + x - 2}.$$

Then f *is not defined at 1, but* $\lim_{x \to 1} f(x)$ *exists.*

[Note that a standard technique here would be to write

$$\frac{x^2 - 1}{x^2 + x - 2} = \frac{(x+1)(x-1)}{(x+2)(x-1)} = \frac{x+1}{x+2},$$

but strictly speaking, the function on the LHS is not equal to that on the RHS since LHS is not defined at 1.]

Proof. We make a guess that $\ell = \frac{2}{3}$. Let $\varepsilon > 0$. Since $x + 2 \geqslant 2$, for all $x \geqslant 0$, for all $x \in [0, \infty) \setminus \{1\}$,

$$\begin{aligned} \left| f(x) - \frac{2}{3} \right| &= \left| \frac{x^2 - 1}{x^2 + x - 2} - \frac{2}{3} \right| = \left| \frac{x+1}{x+2} - \frac{2}{3} \right| \\ &= \left| \frac{3(x+1) - 2(x+2)}{3(x+2)} \right| = \frac{|x-1|}{3(x+2)} \leqslant \frac{|x-1|}{6}. \end{aligned}$$

So setting $\delta = 6\varepsilon$, $x \in [0, \infty) \setminus \{1\}$ and $|x - 1| \leqslant \delta$ implies $\left| f(x) - \frac{2}{3} \right| \leqslant \varepsilon$. $\square$

Theorem 5.7. *Suppose that* $f : A \to \mathbb{R}$ *is a function. Given a point* $x_0 \in A$, *then* $\lim_{x \to x_0} f(x)$ *exists and equals* $f(x_0)$ *if and only if* f *is continuous at* $f(x_0)$.

Proof. Exercise. □

For general functions, it's possible for $\lim_{x \to x_0} f(x)$ to exist, but not equal $f(x_0)$:

Example 5.12. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

$$f(x) = \begin{cases} x^2 & \text{if } x \neq 3, \\ 0 & \text{if } x = 3. \end{cases}$$

Then $\lim_{x \to 3} f(x) = 9$ ($\neq f(3)$).

Proof. Let $\varepsilon > 0$. First note that $|x - 3| \leqslant 1$ implies $2 \leqslant x \leqslant 4$, so also if $x \neq 3$ then

$$|f(x) - 9| = |x^2 - 9| = |x - 3||x + 3| \leqslant 7|x - 3|.$$

So if $\delta := \min\left\{1, \frac{\varepsilon}{7}\right\}$, then $x \in \mathbb{R} \setminus \{3\}$ and $|x - 3| \leqslant \delta$ implies

$$|f(x) - 9| \leqslant 7|x - 3| \leqslant \varepsilon.$$

□

Theorem 5.8. *Suppose that $f, g : \mathbb{R} \to \mathbb{R}$ are functions and $a \in \mathbb{R}$ has $\lim_{x \to a} f(x) = \ell$, $\lim_{x \to a} g(x) = m$. Then*

1. $\lim_{x \to a}(f + g)(x) = \ell + m$;
2. $\lim_{x \to a}(fg)(x) = \ell m$;
3. $\lim_{x \to a} \left(\frac{f}{g}\right)(x) = \frac{\ell}{m}$ *so long as* $m \neq 0$;
4. $\lim_{x \to a} |f|(x) = |\ell|$

Proof. 1) Let $\varepsilon > 0$. Then there exists $\delta_1 > 0$ such that $x \in \mathbb{R} \setminus \{a\}$, $|x - a| \leqslant \delta_1$ implies

$$|f(x) - \ell| \leqslant \frac{\varepsilon}{2}.$$

Similarly, there exists $\delta_2 > 0$ such that $x \in \mathbb{R} \setminus \{a\}$, $|x - a| \leqslant \delta_2$ implies

$$|g(x) - \ell| \leqslant \frac{\varepsilon}{2}.$$

So setting $\delta := \min\{\delta_1, \delta_2\}$, $x \in \mathbb{R} \setminus \{a\}$, $|x - a| \leqslant \delta$ implies

$$\begin{aligned} |(f+g)(x) - (\ell + m)| &= |(f(x) - \ell) + (g(x) - m)| \\ &\leqslant |f(x) - \ell| + |g(x) - m| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

The proofs of the 2), 3), 4) are exercises. $\square$

It's sometimes useful to be able to distinguish between the limit of a function as we approach from the left, and the limit when we approach from the right.

Definition 5.4. *Given $A \subset \mathbb{R}$ and $f : A \to \mathbb{R}$, for $a \in A$,*

1. *We say that the* left-limit *of $f(x)$ as $x \to a$ exists and equals ℓ and write*
$$\lim_{x \to a^-} f(x) = \ell,$$
if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $x \in A$ has $x < a$ and $|x - a| \leqslant \delta$ then
$$|f(x) - \ell| \leqslant \varepsilon.$$

2. *We say that the* right-limit *of $f(x)$ as $x \to a$ exists and equals ℓ and write*
$$\lim_{x \to a^+} f(x) = \ell,$$
if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $x \in A$ has $x > a$ and $|x - a| \leqslant \delta$ then
$$|f(x) - \ell| \leqslant \varepsilon.$$

Example 5.13. *Define $f : \mathbb{R} \to \mathbb{R}$ by*

$$f(x) = \begin{cases} |x| + \frac{x}{|x|} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then $\lim_{x \to 0^-} f(x) = -1$ and $\lim_{x \to 0^+} f(x) = 1$.

Theorem 5.9. *Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and $a \in A$. Then $\lim_{x \to a} f(x)$ exists if and only if $\lim_{x \to a^-} f(x)$ and $\lim_{x \to a^+} f(x)$ both exist and are equal. Moreover, if $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x) = \ell$ then $\lim_{x \to a} f(x) = \ell$.*

Proof. Suppose that $\lim_{x\to a} f(x) = \ell$. Then for all $\varepsilon > 0$ there exists $\delta > 0$ such that $x \in A \setminus \{a\}$, $|x - a| \leqslant \delta$ implies $|f(x) - \ell| \leqslant \varepsilon$. So in particular if $x \in A$ and $x < a$ then $|x - a| \leqslant \delta$ implies $|f(x) - \ell| \leqslant \varepsilon$, so $\lim_{x\to a^-} f(x) = \ell$; and if $x \in A$ and $x > a$ then $|x - a| \leqslant \delta$ implies $|f(x) - \ell| \leqslant \varepsilon$, so $\lim_{x\to a^+} f(x) = \ell$.

Now suppose that $\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x)$ and denote this value by ℓ. Set $\varepsilon > 0$. Then there exists $\delta_1 > 0$ such that if $x \in A$ and $x < a$ then $|x - a| \leqslant \delta_1$ implies $|f(x) - \ell| \leqslant \varepsilon$. Also there exists $\delta_2 > 0$ such that if $x \in A$ and $x > a$ then $|x - a| \leqslant \delta_2$ implies $|f(x) - \ell| \leqslant \varepsilon$. So setting $\delta := \min\{\delta_1, \delta_2\}$, if $x \in A \setminus \{a\}$ then $|x - a| \leqslant \delta$ implies $|f(x) - \ell| \leqslant \varepsilon$, i.e., $\lim_{x\to a} f(x) = \ell$. □

Note that in Example 5.13, $\lim_{x\to 0} f(x)$ does not exist since $\lim_{x\to 0^-} f(x) \neq \lim_{x\to 0^+} f(x)$.

Definition 5.5. *Given $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and $a \in A$,*

1. *we say that $\lim_{x\to a} f(x) = \infty$ if for all $M \in \mathbb{R}$, there exists $\delta > 0$ such that if $x \in A \setminus \{a\}$ then $|x - a| \leqslant \delta$ implies $f(x) \geqslant M$.*

2. *we say that $\lim_{x\to a} f(x) = -\infty$ if for all $M \in \mathbb{R}$, there exists $\delta > 0$ such that if $x \in A \setminus \{a\}$ then $|x - a| \leqslant \delta$ implies $f(x) \leqslant M$.*

Chapter 6

Differentiation

Differentiation is a crucial topic in any area of science where a system evolves in time. In this chapter we'll give the fundamental ideas of differentiation a rigorous treatment, aided by our knowledge of limits.

Definition 6.1. *Let $I \subseteq \mathbb{R}$ be an open interval, $f : I \to \mathbb{R}$ a function and $c \in I$. Then f is* differentiable *at c if*

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \qquad \textit{exists,}$$

in which case we denote this limit by $f'(c)$, the derivative *of f at c.*

Letting c vary over the whole of I, if f is differentiable at every such c, then $f' : I \to \mathbb{R}$ can be thought of as a function $f'(x)$. This can also be denoted

$$\frac{d}{dx} f(x), \frac{df}{dx}, D_x f(x).$$

Note that letting $x = c + h$, $x \to c$ if and only if $h \to 0$, so we can also use

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h}$$

existing as the definition of differentiability and the value of $f'(c)$.

Example 6.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Then $f'(2) = 4$, since $x \neq 2$ implies*

$$\frac{f(x) - f(c)}{x - 2} = \frac{x^2 - 2^2}{x - 2} = \frac{x^2 - 4}{x - 2} = \frac{(x-2)(x+2)}{x - 2} = x + 2,$$

so since $g(x) = x$ is continuous at 2, Theorems 5.7 and 5.8 imply

$$\lim_{x\to 2} \frac{f(x)-f(c)}{x-c} = \lim_{x\to 2} x+2 = 2+2 = 4.$$

Similarly, for $c \in \mathbb{R}$, for $x \neq c$,

$$\frac{f(x)-f(c)}{x-c} = \frac{x^2-c^2}{x-c} = \frac{x^2-c}{x-c} = \frac{(x-c)(x+c)}{x-c} = x+c.$$

Hence Theorems 5.7 and 5.8 imply

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{x\to c} x+c = c+c = 2c.$$

So f is differentiable at every $x \in \mathbb{R}$, and $f'(x) = 2x$.

Example 6.2. *Set $f : [0,\infty) \to \mathbb{R}$ to be $f(x) = \sqrt{x}$. Then for $c > 0$, $x \neq c$,*

$$\begin{aligned}\frac{f(x)-f(c)}{x-c} &= \frac{\sqrt{x}-\sqrt{c}}{x-c} = \frac{\sqrt{x}-\sqrt{c}}{x-c}\left(\frac{\sqrt{x}+\sqrt{c}}{\sqrt{x}+\sqrt{c}}\right)\\ &= \frac{x-c}{(x-c)(\sqrt{x}+\sqrt{c})} = \frac{1}{\sqrt{x}+\sqrt{c}}.\end{aligned}$$

So

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{x\to c} \frac{1}{\sqrt{x}+\sqrt{c}} = \frac{1}{2\sqrt{c}}$$

by Theorems 5.7 and 5.8 and TS6 Q3. So for $x > 0$, $f'(x) = \frac{1}{2\sqrt{x}} = \frac{1}{2}x^{-\frac{1}{2}}$.

Example 6.3. *Given $n \in \mathbb{N}$, define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^n$. Then $f'(x) = nx^{n-1}$.*

Proof. Let $c \in \mathbb{R}$. Then for $x \in \mathbb{R}$ we observe that

$$f(x)-f(c) = x^n - c^n = (x-c)(x^{n-1}+cx^{n-2}+c^2x^{n-3}+\cdots+c^{n-2}x+c^{n-1}).$$

So if $x \neq c$,

$$\frac{f(x)-f(c)}{x-c} = x^{n-1}+cx^{n-2}+c^2x^{n-3}+\cdots+c^{n-2}x+c^{n-1}$$

and

$$\begin{aligned}f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} &= \lim_{x\to c} x^{n-1}+cx^{n-2}+c^2x^{n-3}+\cdots+c^{n-2}x+c^{n-1}\\ &= c^{n-1}+cc^{n-2}+c^2c^{n-3}+\cdots+c^{n-2}c+c^{n-1} = nc^{n-1}.\end{aligned}$$

$\square$

Our first main theorem of this chapter shows that differentiability is stronger than continuity.

Theorem 6.1. *Let $I \subseteq \mathbb{R}$ be an open interval. Then for $c \in I$, f being differentiable at c implies f is continuous at c.*

Proof. Since $\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = f'(c)$ and $\lim_{x\to c} x - c = 0$, Theorem 5.8 implies that

$$\lim_{x\to c} f(x) - f(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}(x-c) = \left(\lim_{x\to c} \frac{f(x)-f(c)}{x-c}\right)\left(\lim_{x\to c}(x-c)\right)$$
$$= f'(c) \cdot 0 = 0,$$

i.e., $\lim_{x\to c} f(x) = f(c)$. So by Theorem 5.7, f is continuous at c. □

It is not true that f being continuous at c implies f is differentiable at c, as the following example shows.

Example 6.4. *Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = |x|$. Then f is continuous (see Theorem 5.2 plus the fact that $g(x) = x$ is continuous), but f is not differentiable at 0 since the left limit is*

$$\lim_{x\to 0^-} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^-} \frac{-x}{x} = -1,$$

but the right limit is

$$\lim_{x\to 0^+} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^-} \frac{x}{x} = 1,$$

so $\lim_{x\to 0} \frac{f(x)-f(0)}{x-0}$ doesn't exist by Theorem 5.9.

We've seen that, intuitively, a function is continuous if it has 'no gaps' and the above example suggests that intuitively a function is differentiable if it has 'no corners'. However, a function like $f(x) = x^2 \sin \frac{1}{x}$ (see TS10) eludes such intuition.

The next theorem gives the standard rules of differential calculus.

Theorem 6.2. *Suppose that $I \subseteq \mathbb{R}$ is an open interval, $f, g : I \to \mathbb{R}$ and f and g are differentiable at $c \in I$.*

1. *Given a constant* $\lambda \in \mathbb{R}$*, the function* λf *is differentiable at* c *and* $(\lambda f)' = \lambda f'$.
2. $f + g$ *is differentiable at* c *and* $(f+g)' = f' + g'$.
3. $f \cdot g$ *is differentiable at* c *and* $(f \cdot g)' = f \cdot g' + f' \cdot g$.
4. *If* $g(c) \neq 0$*, then* $\frac{1}{g}$ *is differentiable at* c *and* $\left(\frac{1}{g}\right)' = \frac{-g'}{g^2}$.

Proof. 1)

$$\lim_{x \to c} \frac{(\lambda f)(x) - (\lambda f)(c)}{x - c} = \lim_{x \to c} \lambda \left(\frac{f(x) - f(c)}{x - c} \right),$$

so as in Theorem 5.8, this limit exist and equals $\lambda f'(c)$.

2) Exercise.

3) In this case

$$\begin{aligned} \lim_{x \to c} \frac{(fg)(x) - (fg)(c)}{x - c} &= \lim_{x \to c} \left(\frac{f(x)g(x) - f(x)g(c) + f(x)g(c) - f(c)g(c)}{x - c} \right) \\ &= \lim_{x \to c} \left[f(x) \left(\frac{g(x) - g(c)}{x - c} \right) + g(c) \left(\frac{f(x) - f(c)}{x - c} \right) \right]. \end{aligned}$$

Since f is differentiable at c, it is continuous at c (Theorem 6.1), so Theorems 5.7 and 5.8 imply that this limit exists and equals $f(c)g'(c) + g(c)f'(c)$.

4) First note that $g(c) \neq 0$ and g being continuous at c (Theorem 6.1) means that there exists $\delta > 0$ such that $x \in (c - \delta, c + \delta)$ implies $g(x) \neq 0$. Then

$$\begin{aligned} \lim_{x \to c} \frac{\left(\frac{1}{g}\right)(x) - \left(\frac{1}{g}\right)(c)}{x - c} &= \lim_{x \to c} \frac{1}{x - c} \left(\frac{1}{g(x)} - \frac{1}{g(c)} \right) \\ &= \lim_{x \to c} \frac{1}{x - c} \left(\frac{g(c) - g(x)}{g(x)g(c)} \right) \\ &= \lim_{x \to c} - \left(\frac{g(x) - g(c)}{x - c} \right) \frac{1}{g(x)g(c)}. \end{aligned}$$

So since g is continuous at c and $g(c) \neq 0$, Theorems 5.7 and 5.8 imply that this limit exists and is equal to $-\frac{g'(c)}{g(c)^2}$. ☐

Corollary 6.3. *For an open interval* $I \subseteq \mathbb{R}$*,* $f, g : I \to \mathbb{R}$ *and* $c \in I$ *such that* f *and* g *are differentiable at* c*,*

1. *$f - g$ is differentiable at c and $(f-g)' = f' - g'$*

2. *if also $g(c) \neq 0$, then $\frac{f}{g}$ is differentiable at c and $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.*

Proof. 1) $f - g = f + (-1)g$, so the proof is immediate by parts 1) and 2) of Theorem 6.2.

2) Since $\frac{f}{g} = f \cdot \left(\frac{1}{g}\right)$, part 4) of Theorem 6.2 implies that $\frac{1}{g}$ is differentiable at c, so part 3) of Theorem 6.2 implies

$$\left(\frac{f}{g}\right)' = f\left(\frac{1}{g}\right)' + f'\left(\frac{1}{g}\right) = -\frac{fg'}{g^2} + \frac{f'}{g} = \frac{f'g - fg'}{g^2},$$

as required. □

One consequence of the last two results is that if $p, q : \mathbb{R} \to \mathbb{R}$ are polynomials, then p and q are differentiable. Moreover, $\frac{p}{q}$ (which is known as a *rational function*) is differentiable on the set $\{x \in \mathbb{R} : q(x) \neq 0\}$.

Example 6.5. *Let $n \in \mathbb{N}$ and let $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be given by*

$$f(x) = \frac{1}{x^n} \text{ for } x \in \mathbb{R} \setminus \{0\}.$$

Then $f'(x) = \frac{-n}{x^{n+1}}$.

Proof. Let $g(x) = x^n$. Then Theorem 6.2 implies

$$f'(x) = \left(\frac{1}{g}\right)' = \frac{-g'}{(g)^2} = \frac{-nx^{n-1}}{x^{2n}} = \frac{-n}{x^{n+1}}.$$

□

Note we could also have phrased this as $f(x) = x^{-n}$ and $f'(x) = -nx^{-n-1}$.

The final rule of differentiability we'll give is about composition of functions.

Theorem 6.4 (Chain Rule)**.** *Let $I, J \subseteq \mathbb{R}$ be open intervals and $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ be functions. Suppose that $c \in I$, $f(c) \in J$, f is differentiable at c and g is differentiable at $f(c)$. Then $g \circ f$ is differentiable at c and*

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

'Proof': We wish to show

$$\frac{(g \circ f)(x) - (g \circ f)(c)}{x - c} \to (g'(f(c))f'(c),$$

so we could write

$$\frac{(g \circ f)(x) - (g \circ f)(c)}{x - c} = \frac{g(f(x)) - g(f(c))}{x - c} = \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \frac{f(x) - f(c)}{x - c}$$

and then argue that as $x \to c$, $f(x) \to f(c)$, so

$$\lim_{x \to c} \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} = \lim_{u \to f(c)} \frac{g(u) - g(f(c))}{u - f(c)} = g'(f(c)).$$

The problem here is that $f(x) - f(c)$ could be zero for many values of x, so the above argument is invalid.

We omit a genuine proof of the Chain Rule.

Example 6.6. *Given the function $h : \mathbb{R} \to \mathbb{R}$ where $h(x) = (x^3 + x^2 + 3)^6$, compute $h'(x)$.*

Solution: We can write $h(x) = g \circ f(x)$ where $f(x) = x^3 + x^2 + 3$ and $g(u) = u^6$. So since $f'(x) = 3x^2 + 2x$ and $g'(u) = 6u^5$, the Chain Rule implies that

$$\begin{aligned} h'(x) = (g \circ f)'(x) = g'(f(x))f'(x) &= 6(f(x))^5(3x^2 + 2x) \\ &= 6(x^3 + x^2 + 3)^5(3x^2 + 2x). \end{aligned}$$

Example 6.7. *Suppose that $h : \mathbb{R} \to \mathbb{R}$ is defined by $h(x) = \sin(x^3 + 7x)$. Then*

$$h'(x) = (3x^2 + 7)\cos(x^3 + 7x).$$

Proof. We can write $h(x) = g \circ f(x)$ where $f(x) = x^3 + 7x$ and $g(u) = \sin u$. Then since $f'(x) = 3x^2 + 7$ and $g'(u) = \cos u$, the Chain Rule implies that

$$h'(x) = (g \circ f)'(x) = g'(f(x))f'(x) = (\cos(x^3 + 7x))(3x^2 + 7).$$

□

Our next results extract information about maxima/minima and average slopes using the definition of derivative.

Theorem 6.5. *Let $I \subseteq \mathbb{R}$ be an open interval and $f : I \to \mathbb{R}$ a differentiable function.*

1. *If f attains its maximum at $c \in I$ then $f'(c) = 0$.*
2. *If f attains its minimum at $c \in I$ then $f'(c) = 0$.*

Proof. 1): By assumption,

$$f(x) \leqslant f(c) \text{ for all } x \in I.$$

Hence

$$f(x) - f(c) \leqslant 0 \text{ for all } x \in I.$$

If $x \in I$ has $x < c$, then $x - c < 0$, so

$$\frac{f(x) - f(c)}{x - c} \geqslant 0, \text{ hence } \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \geqslant 0.$$

On the other hand, if $x \in I$ has $x > c$, then $x - c > 0$, so

$$\frac{f(x) - f(c)}{x - c} \leqslant 0, \text{ hence } \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leqslant 0.$$

Since f is differentiable at c, the left and right limits must exist and be equal, which is only possible if

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = 0$$

as in Theorem 5.9.

2) follows similarly. □

The main Theorem of this chapter is the Mean Value Theorem. This gives information on the derivative of a function between two given points. The following result is a simpler version of this, and roughly states that if a graph is differentiable and $f(a) = f(b)$ then there is a point between a and b where the graph is flat.

Theorem 6.6 (Rolle's Theorem)**.** *Suppose that $a, b \in \mathbb{R}$ with $a < b$. Suppose further that $f : [a, b] \to \mathbb{R}$ is continuous and is differentiable on (a, b) with $f(a) = f(b)$. Then there exists $\theta \in (a, b)$ such that*

$$f'(\theta) = 0.$$

Proof.

- **Case 1:** $f(x') = f(a)$ for all $x' \in [a, b]$. Then $f'(x) = 0$ for all $x \in (a, b)$, so we are finished.

- **Case 2:** there exists $y \in (a, b)$ such that $f(y) > f(a)$. Since $f : [a, b] \to \mathbb{R}$ is continuous, the Extreme Value Theorem implies that there exists $\theta \in [a, b]$ such that

  $$f(x) \leqslant f(\theta) \text{ for all } x \in [a, b].$$

  In particular, since there exists $y \in (a, b)$ such that $f(y) > f(a)$, this means that $\theta \neq a, b$, so $\theta \in (a, b)$. So f has a maximum at $\theta \in (a, b)$. Since f is differentiable on (a, b), Theorem 6.5 implies that $f'(\theta) = 0$.

- **Case 3:** there exists $y \in (a, b)$ such that $f(y) < f(a)$. This follows similarly to case 2, with minimum replacing maximum.

□

One immediate consequence of Rolle's Theorem is the following.

Theorem 6.7 (Mean Value Theorem)**.** *Suppose that $a, b \in \mathbb{R}$ with $a < b$, $f : [a, b] \to \mathbb{R}$ is continuous and f is differentiable on (a, b). Then there exists $\theta \in (a, b)$ such that*

$$f'(\theta) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Set $g(x) = f(x) - \lambda x$ where we choose λ such that $g(a) = g(b)$. Thus

$$f(a) - \lambda a = f(b) - \lambda b,$$

so

$$\lambda = \frac{(f(b) - f(a)}{b - a}.$$

We can then apply Rolle's Theorem to g to obtain $\theta \in (a, b)$ such that $g'(\theta) = 0$, which means that

$$f'(\theta) = \lambda = \frac{f(b) - f(a)}{b - a},$$

as required. □

Note that one way of phrasing the Mean Value Theorem is that there exists $\theta \in (a, b)$ such that

$$f(b) = f(a) + (b - a)f'(\theta).$$

(This is useful in itself, but also tees up Taylor's Theorem.)

Theorem 6.8. *Let $f : [a, b] \to \mathbb{R}$ be continuous and f be differentiable on (a, b). If $f'(x) = 0$ for all $x \in (a, b)$ then there exists $k \in \mathbb{R}$ such that $f(x) = k$ for all $x \in [a, b]$.*

Proof. Let $k = f(a)$ and $c \in (a, b]$. Then MVT implies there exists $\theta \in (a, c)$ such that $f(c) = f(a) + (c - a)f'(\theta)$. Since $f'(\theta) = 0$ then $f(c) = f(a) = k$, as required. $\square$

Example 6.8. *Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Prove (without actually differentiating!)*

1. *there exists $\theta \in \mathbb{R}$ such that $f'(\theta) = 1$;*
2. *there exists $\theta \in \mathbb{R}$ such that $f'(\theta) = -2$;*

Proof. 1) Since $f(0) = 0$ and $f(1) = 1$, MVT implies there exists $\theta \in (0, 1)$ such that

$$f'(\theta) = \frac{f(1) - f(0)}{1 - 0} = 1.$$

2) Since $f(0) = 0$ and $f(-2) = 4$, MVT implies there exists $\theta \in (-2, 0)$ such that

$$f'(0) = \frac{f(0) - f(-2)}{0 - -2} = \frac{0 - 4}{2} = -2.$$

$\square$

6.1 Higher Derivatives

Differentiating a function once yields another function f'. If this function is differentiable, differentiating again yields yet another function f'', the *second derivative* of f. Similarly we can often find the third derivative etc.

Repeating this process n times, if this is possible, yields the *nth derivative* $f^{(n)}$. This can also be denoted

$$(D_x^n f)(x), \frac{d^n}{dx^n} f(x).$$

Supposing that f and g are $n \geqslant 2$ times differentiable,

1. Since we know $(f+g)' = f' + g'$, then

   $$(f+g)'' = (f'+g')' = f'' + g''.$$

   (Clearly $(f+g)^{(n)} = f^{(n)} + g^{(n)}$.)

2. If $\lambda \in R$, then $(\lambda f)^{(n)} = \lambda f^{(n)}$.

3. $(fg)''$ is more difficult. By Theorem 6.2, $(fg)' = f'g + fg'$, so

   $$\begin{aligned}(fg)'' &= (f'g + fg')' = (f'g)' + (fg')' = (f''g + f'g') + (f'g' + fg'') \\ &= f''g + 2f'g' + fg''.\end{aligned}$$

   This generalises further for $(fg)^{(n)}$ (Leibniz's Theorem), but we omit that here.

The following theorem, a fundamental theorem used throughout mathematics, is a generalisation of the MVT

Theorem 6.9 (Taylor's Theorem)**.** *Let $n \in \mathbb{N}$ and suppose that $f : [a,b] \to \mathbb{R}$ is a continuous function which is n times differentiable on (a,b). Then there exists $\theta \in (a,b)$ such that*

$$\begin{aligned}f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f''(a) + \cdots \\ \cdots + \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + \frac{(b-a)^n}{n!}f^{(n)}(\theta).\end{aligned}$$

Proof. We define the function $F_n : [a,b] \to \mathbb{R}$ by

$$F_n(x) = f(b) - f(x) - (b-x)f'(x) - \frac{(b-x)^2}{2!}f''(x) - \cdots - \frac{(b-x)^{n-1}}{(n-1)!}f^{(n-1)}(x).$$

Then

$$\begin{aligned} F_n'(x) = &- f'(x) + [f'(x) - (b-x)f''(x)] \\ &+ \left[(b-x)f''(x) - \frac{1}{2!}(b-x)^2 f'''(x)\right] + \cdots \\ &\quad \cdots + \left[\frac{(b-x)^{n-2}}{(n-2)!} f^{(n-1)}(x) - \frac{(b-x)^{n-1}}{(n-1)!} f^{(n)}(x)\right] \qquad (*) \\ = &-\frac{(b-x)^{n-1}}{(n-1)!} f^{(n)}(x). \end{aligned}$$

Now set $G_n : [a,b] \to \mathbb{R}$ to be

$$G_n(x) = F_n(x) - \left(\frac{b-x}{b-a}\right)^n F_n(a).$$

Then $G_n(a) = F_n(a) - F_n(a) = 0$ and $G_n(b) = F_n(b) - 0 \cdot F_n(a) = 0$. So since G_n is differentiable, Rolle's Theorem implies that there exists $\theta \in (a,b)$ such that $G_n'(\theta) = 0$. I.e., using $(*)$,

$$\begin{aligned} 0 = G_n(\theta) &= F_n'(\theta) - \left(\frac{-n(b-\theta)^{n-1}}{(b-a)^n} F_n(a)\right) \\ &= -\frac{(b-\theta)^{n-1}}{(n-1)!} f^{(n)}(\theta) + \frac{n(b-\theta)^{n-1}}{(b-a)^n} F_n(a) \\ &= n\frac{(b-\theta)^{n-1}}{(b-a)^n}\left(F_n(a) - \frac{(b-\theta)^{n-1}}{(n-1)!} f^{(n)}(\theta)\right) \\ &= n\frac{(b-\theta)^{n-1}}{(b-a)^n}\Big(f(b) - f(a) - (b-a)f'(a) - \cdots \\ &\qquad \cdots - \frac{(b-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) - \frac{(b-a)^n}{n!} f^{(n)}(\theta)\Big). \end{aligned}$$

Hence

$$f(b) = f(a) + (b-a)f'(a) + \cdots + \frac{(b-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) + \frac{(b-a)^n}{n!} f^{(n)}(\theta),$$

as required. □

A more Scottish version of this result comes from setting $a = 0$ and $b = x$ to obtain Maclaurin's Theorem:

$$f(x) = f(0) + xf'(x) + x^2 \frac{f''(0)}{2!} + \cdots \frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0) + R_n$$

where $R_n = \frac{x^n}{n!} f^{(n)}(y)$ for some $y \in (0, x)$. (In fact, there is evidence that Gregory `http://www-history.mcs.st-andrews.ac.uk/history/Biographies/Gregory.html` 'discovered Taylor's Theorem' before Taylor, while working at St Andrews.)

Example 6.9. *Let f"$\mathbb{R} \to \mathbb{R}$ be given by $f(x) = e^x$. Then $f'(x) = e^x$, $f''(x) = e^x, \cdots, f^{(n)}(x) = e^x$. So since $e^0 = 1$, the expression from Maclaurin's Theorem gives*

$$f'(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!} e^y$$

for some $y \in (0, x)$.

Example 6.10. *If $f(x) = \sin x$, then $f(0) = 0$; $f'(x) = \cos x$, so $f'(0) = 1$; $f''(x) = -\sin x$, so $f''(0) = 0$; $f'''(x) = -\cos x$, so $f'''(0) = -1$; $f''''(x) = \sin x$, so there is $y \in (0, x)$ such that*

$$f(x) = 0 + 1 \cdot x + 0 \cdot \frac{x^2}{2!} - 1\frac{x^3}{3!} + \frac{x^4}{4!} \sin y = x - \frac{x^3}{3!} + \frac{x^4}{4!} \sin y.$$

Example 6.11. *If $f(x) = \log(1 + x)$ (here we mean $\log$ to base e), then $f(0) = \log 1 = 0$; $f'(x) = \frac{1}{1+x}$, so $f'(0) = 1$; $f''(x) = \frac{-1}{(1+x)^2}$, so $f''(0) = -1$; $f'''(x) = \frac{2}{(1+x)^3}$, so $f'''(0) = 2$. In fact we can show by induction that $f^{(n)}(x) = \frac{(-1)^{n+1}(n-1)!}{(1+x)^n}$, so*

$$f(x) = x - \frac{x^2}{2!} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots.$$

We can actually show that this is a convergent series (this goes slightly beyond this course). Note that in particular,

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots.$$

Appendix A

Some notation for MT2502 and beyond

Important note: notation varies from mathematician to mathematician (within reason, usually), so a definitive list is impossible.

A.1 Logic

$\forall$	$\exists$	s.t.	$\Rightarrow$	$\Leftarrow$	$\iff$	iff
for all	there exists	such that	implies	is implied by	if and only if	if and only if

A.2 Numbers

- $\mathbb{N}$ denotes the set of natural numbers. In MT2502 this means $1, 2, 3, \ldots$ ($\ldots$ means 'and so on forever'). (Note that for many other mathematicians, and courses at St Andrews, $\mathbb{N}$ is $0, 1, 2, \ldots$: there is no universally, or even locally, agreed convention.)
- $\mathbb{Z}$ denotes the integers: the set of whole numbers, both positive and negative, and 0, i.e., $\ldots, -2, -1, 0, 1, 2, \ldots$
- $\mathbb{Q}$ denotes the rational numbers, i.e, numbers which can be expressed $\frac{p}{q}$ where p, q are integers (it's common to assume that one of these is

not-negative since this doesn't change the set in question).

- $\mathbb{R}$ denotes the real numbers, discussed in MT2502.

A.3 Sets

- $\{A, B, C\}$ means the set of objects A, B, C - these can be numbers, general algebraic objects, sets, sets of sets... anything. Eg $\mathbb{N} = \{1, 2, 3, \ldots\}$. Note that the order is not important, so $\{C, A, B\} = \{A, B, C\}$. Also repetitions don't change anything, i.e, $\{A, B, C, B\} = \{A, B, C\}$.

- Given a set X, $x \in X$ means that x is in the set X. Conversely, $x \notin X$ means that x is not in the set X.

- Given a set X and a property P, the notation $\{x \in X : x \text{ satisfies property } P\}$ means the set of x in X which satisfy property P. (Note that $\{x \in X \mid x \text{ satisfies property } P\}$ has the same meaning.) Eg $\{x \in \mathbb{N} : x \text{ is divisible by } 2\}$ is the set of positive even numbers.

- In MT2502, $\subseteq$ means 'subset'. So if $A \subseteq B$, then any point in A is also in B. Since B is a subset of itself, $B \subseteq B$. (Some mathematicians write $\subset$ to mean the same thing, while for others this means a 'proper' subset - i.e., excluding the possibility that the subset is equal to the whole set.)

- The symbol $\cup$ is translated as 'or', while the symbol $\cap$ is 'and'. So given two sets A and B, $A \cup B = \{x : x \in A \text{ or } x \in B\}$ (note that the point x could be in both A and B here). Similarly $A \cap B$ is the set of points in both A and B.

- The symbol $\setminus$ is used for subtracting sets, so that $A \setminus B = \{x \in A : x \notin B\}$. (Some mathematicians simply use a minus sign instead.)

- It's useful to have an idea of the 'empty set', which is denoted by $\emptyset$. So, for example, one way of writing that a set A is not empty is that $A \neq \emptyset$.