Real Analysis

Lachlan Dufort-Kennett

December 12, 2022

Contents

Preface **3**

1 Introduction **4**
1.1 What is Analysis? 4

2 Notions from Set Theory **6**
2.1 Introduction to Sets 6
2.2 Algebra of Sets 7
2.3 Ordering & Bounding 8

3 The Real Numbers **10**
3.1 Why do we Need the Real Numbers? 10
3.2 Construction of the Reals 11
3.3 Subsets of the Real Line 11
3.4 Absolute Value 13
3.5 Inequalities 14

4 Sequences **15**
4.1 Sequences & Convergence 15
4.2 Properties of Limits 18
4.3 Monotone Sequences 21
4.4 Subsequences 23
4.5 Cauchy Sequences 25

5 Series **28**
5.1 Series 28
5.2 The Comparison Test 31
5.3 Series of Positive & Negative Terms 33
5.4 Power Series 35

6 Continuous Functions **38**
6.1 The Basic Definition of a Function 38
6.2 Operations on Functions 39
6.3 Classes of Functions 40

6.4 Limits 41
6.5 Continuity 41

7 Differentiation **48**
7.1 The Derivative 48
7.2 Important Theorems about Derivatives 52
7.3 Higher Derivatives 55

Bibliography **59**

Preface

Chapter 1

Introduction

1.1 What is Analysis?

Real analysis is the rigorous study of concepts that involve limits of infinite processes. Some examples of these that appear in all areas of mathematics are convergence of sequences or series, limits of functions, or derivatives and integrals. Until the 18th/19th century, our understanding of these concepts was based on intuition; such as the idea that a continuous function is one that can be drawn without removing the pen off the page, or that a differentiable function is one of which the graph has no sharp corners; and while these intuitions can take us very far, eventually the discovery of many 'paradoxes' which contradict these intuitions lead to the creation of a new foundation for mathematics based on rigour and mathematical logic.

An example of one of these contradictions comes from a Fourier series, which is an infinite summation of sine functions. It is pretty clear that the sine function is continuous, so therefore it would make sense that any superposition of sine functions is also continuous, right? It turns out that this is not the case.

Example 1.1. The function S defined by:

$$S(x) = \sum_{n=1}^{\infty} \frac{(-1)^{(n+1)}}{n} \sin nx \tag{1.1}$$

$$= \sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \cdots \tag{1.2}$$

is not continuous. We can see this very clearly if we look at the graph.

In order to reconcile these apparent contradictions, we have to use a number system which is 'complete' i.e. there are no gaps in the number line. It is a well-known fact that the rational numbers $\mathbb{Q}$ do not fulfill this requirement, as can be seen simply drawing a right-angled triangle with short sides of length 1. In this case the length of the hypotenuse, by the Pythagorean theorem, squares to 2, and it can be shown that this quantity cannot be expressed as a ratio of two integers.

Theorem 1.1 *There is no rational number x such that $x^2 = 2$.*

Proof. Suppose $x = \frac{p}{q}$, with $p, q \in \mathbb{N}$ and $\gcd(p, q) = 1$ (x is expressed as a rational number in lowest terms). Since $x^2 = 2$,

$$\frac{p^2}{q^2} = 2 \implies p^2 = 2q^2, \tag{1.3}$$

so p is an even number (a square can only be even if the number itself is even), thus $\exists k \in \mathbb{N}$ such that $p = 2k$. Hence,

$$(2k)^2 = 2q^2 \implies 2k^2 = q^2, \tag{1.4}$$

so q is also an even number, but this is a contradiction of our assumption that $\gcd(p, q) = 1$, and so the initial assumption that x can be expressed as a rational number must be false. ■

What this simple example shows is that to do analysis, we need to create a new *complete* number system, the **real numbers** $\mathbb{R}$, but before we do that, we need to lay down a few definitions about sets.

Chapter 2

Notions from Set Theory

2.1 Introduction to Sets

In modern mathematics, set theory is important as a foundation from which all of mathematics can be derived. All of the mathematical objects we deal with in analysis will be defined in terms of sets.

Definition 2.1. A **set** is an *unordered* collection of objects. If an object x is contained in a set A, we call x an **element** or **member** of A and write $x \in A$. Likewise if x is *not* contained in A we can say x **does not belong to** A and write $x \notin A$. Sets are notated with curly braces $\{\}$ around a comma-separated list of the elements.

The elements of a set can be any mathematical object but most sets we will encounter will be sets of numbers, such as $\mathbb{N} = \{1, 2, 3, ...\}$, the set of **natural numbers**[1]; or $\mathbb{Z} = \{..., -2 -1, 0, 1, 2, ...\}$, the set of **integers** ('...' means 'and so on forever').

Definition 2.2. There is a set which has no elements, and this set is unique. We call it the **empty set** and denote it $\emptyset$. By definition for every object x we have $x \notin \emptyset$. If we say a set A is **non-empty**, then $\exists$ some object x such that $x \in A$.

It is useful to have the ability to define a set in terms of some kind of predicate, for example let E be 'the set of all even integers'. We can write this in a more compact and unambiguous way as $E = \{2k : k \in \mathbb{Z}\}$ (The colon, sometimes replaced with $|$, is read as 'such that'). This is known as **set-builder notation**.

Definition 2.3. Two sets A and B are **equal** $\iff$

$$\forall x \in A, x \in B \text{ and } \forall y \in B, y \in A.$$

[1] Note that in this text, $0 \notin \mathbb{N}$. Generally there is no strong convention on whether 0 is included or not so it is important to check with every text.

Then we write $A = B$. In other words, both sets must have the same elements, but note that multiplicity and order do not matter, for example $\{1, 2, 3\} = \{3, 2, 1, 1\}$. Two sets defined using set-builder notation are equal if and only if their predicates are equivalent.

Definition 2.4. Let A be a set. Then a set B is a **subset** of $A \iff$

$$\forall x \in B, x \in A.$$

Then we write $B \subseteq A$. Note that every set has two trivial subsets, itself and $\emptyset$.

Definition 2.5. Let A be a set. Then a set B is a **proper subset** of $A \iff$

$$\forall x \in B, x \in A \text{ and } A \neq B.$$

Then we write $B \subset A$.

2.2 Algebra of Sets

There are operations that we can do on sets to form new sets. These operations can be interpreted as set-theoretic implementations of the Boolean operations and, or, & not. They also have satisfying parallels with operations and relations that we are familiar with for numbers. For the next three definitions, let A and B be sets.

Definition 2.6. The **union** of A and B, $A \cup B$, is defined as

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

Where 'or' is inclusive, so the union includes all elements of both A and B.

Definition 2.7. The **intersection** of A and B, $A \cap B$, is defined as

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

So the intersection includes only elements that are in both sets. A and B are said to be **disjoint** if and only if $A \cap B = \emptyset$, i.e. they have no elements in common.

Definition 2.8. The **difference** of A and B, $A \setminus B$, is defined as

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

We can also use this operation to define the complement of a set. For example, if A

is a subset of some set X, then the **complement** of A in X is given by $X \setminus A$, i.e. the elements of X that are *not* in A.

Definition 2.9. The **Cartesian product** of A and B, $A \times B$, is defined as

$$\{(a, b) : a \in A \text{ and } b \in B\}.$$

So in plain speech, the Cartesian product of two sets is the set of all ordered pairs where the first component is a member of the first set and the second component is a member of the second set. It can be seen as a generalisation of the notion of Cartesian coordinates in the plane. Sometimes the Cartesian product of a set X with itself is noted as X^2 instead of $X \times X$, for example $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$ denoting the set of all integer points in the plane.

2.3 Ordering & Bounding

Definition 2.10. A **partially ordered set** is an ordered pair $(X, \leq)$, where X is a set and $\leq$ is a binary relation which forms a **partial order** on X. For any two elements $x, y \in X$ we have either $x \leq y$, or $x \nleq y$.

Example 2.1. Let S be a set, then let $P(S)$ be the set of subsets of S. Then $(P(S), \subseteq)$ is a partially ordered set. For any two elements $A, B \in P(S)$ we have either $A \subseteq B$ or $A \nsubseteq B$.

Definition 2.11. A **totally ordered set** is an ordered pair $(X, \leq)$, where X is a set and $\leq$ is a binary relation which forms a **total order** on X. So for any two elements $x, y \in X$ we have $x \leq y$, or $y \leq x$, or both (never neither)[a]. It is possible for a subset of a partially ordered set to be totally ordered (under the same relation).

[a]This is known as the **trichotomy law**

Example 2.2. $(\mathbb{N}, \leq)$ — the set of natural numbers with the greater than or equal to relation — forms a totally ordered set. For any $m, n \in \mathbb{N}$ we have $m \leq n$, or $n \leq m$, or both $(m = n)$. Any subset of $\mathbb{N}$ is also totally ordered under $\leq$.

The key difference between partial orders and total orders is that partial orders are only useful for comparing some elements with each other whereas total orders can compare all elements with each other.

There are several different ways we can create definitions for different kinds of 'extreme' elements in a set, all with subtle differences between them. We will go through some of these now.

Definition 2.12. Let $(P, \leq)$ be a partially ordered set and let $S \subseteq P$.

- An element $g \in S$ is the **greatest element** of S if $\forall s \in S$, $s \leq g$. If a greatest element exists, clearly it is unique (due to the trichotomy law). In other words, a greatest element is an element that is greater than *all* other elements. However, note that greatest elements may not necessarily exist.

- An element $m \in S$ is a **maximal element** of S if there does *not* exist any $s \in S$ with $(s \neq m)$ such that $m \leq s$. A set may have multiple maximal elements, and they may exist without there being a greatest element. Maximal elements also may not necessarily exist. In simple terms, a maximal element is one that is not smaller than any other element. Notice how this definition is subtly different to the definition of the greatest element.

- An element $u \in P$ is an **upper bound** for S in $P \iff \forall s \in S$, $s \leq u$. Note that if $P = S$, then the definition of a greatest element of S becomes equivalent to the definition of an upper bound of S in S, therefore an element $g \in S$ is the greatest element of S if and only if g is an upper bound for S and $g \in S$. It is also plausible that S may *not* have a greatest element while also having an upper bound in P (that's a tricky one to get your head around).

- Let T be a *totally ordered* subset of P. Then the greatest element of T and the maximal element of T are the same, in which case this element is called the **maximum** of T.

- Let $U(S) \subseteq P$ be the set of all upper bounds for S in P. Then $l \in U(S)$ is the **supremum** (or **least upper bound**) of S if $\forall u \in U(S)$, $l \leq u$. Basically the supremum of S is the least element of $U(S)$. Thus if suprema exist, they are unique. If a greatest element or maximum exists, then it is the supremum, and this is the only case where the supremum of a set will lie in the set itself.

- If we switch the elements around the $\leq$ sign in all the above definitions, we obtain the definitions for the dual notions of all those just defined: the **least element**, **minimal element**, **lower bound**, **minimum**, and **infimum** (or **greatest lower bound**).

Chapter 3

The Real Numbers

3.1 Why do we Need the Real Numbers?

The rational numbers have many desirable properties by themselves which allow us to do a lot of mathematics with them. One of these is that they are closed under regular arithmetic operations of addition and multiplication. In fact, $\mathbb{Q}$ forms an **ordered field** when paired with the total order $\leq$. The rational numbers, along with the integers and natural numbers, satisfy the **Archimedean property**,

$$\forall q \in \mathbb{Q},\ \exists n \in \mathbb{N} \text{ s.t. } n > q,$$

and they also possess the property of **density**, which states that between any two rational numbers there exists another rational number.

Theorem 3.1 *$\mathbb{Q}$ is dense.*

Proof. Let $p, q \in \mathbb{Q}$ with $p < q$. Then let $r = \frac{p+q}{2}$, so $r \in \mathbb{Q}$ and $p < r < q$. ∎

This process can be repeated indefinitely, which shows that there are actually *infinitely many* rational numbers between each rational number. This being the case, there are still many numbers that we need that are not in $\mathbb{Q}$. Specifically, since limits are one of the central topics of study in analysis, it would be nice if we could guarantee the existence of limits for convergent sequences. It turns out that this is equivalent to the completeness property, which we will now define.

Definition 3.1. A partially ordered set X is **complete**[a] if every non-empty subset of X which is bounded above (has an upper bound) has a least upper bound (supremum) in X.

[a] This property is sometimes called **Dedekind completeness**, or simply the **least-upper-bound property**.

It can be shown wih a counterexample that the rational numbers are not complete.

Theorem 3.2 *Let $S = \{q \in \mathbb{Q} : q^2 < 2\}$. Then there does not exist a rational number r such that $r = \sup(S)$.*

Proof. S is clearly bounded from above, for example $\frac{3}{2}$ is an upper bound $\left(\left(\frac{3}{2}\right)^2 = \frac{9}{4} > 2\right)$, so by the completeness property we would expect a supremum to exist. Let us assume by way of contradiction that $r = \sup(S)$ and that $r \in \mathbb{Q}$. Then, by the trichotomy law, we have $r^2 < 2$, $r^2 > 2$, or $r^2 = 2$.
If $r^2 < 2$, let $\varepsilon > 0$, then

$$\begin{aligned}(r+\varepsilon)^2 &= r^2 + \varepsilon(2r+\varepsilon) \\ &< r^2 + \varepsilon(2r+1) && (\varepsilon < 1) \\ &< r^2 + 2 - r^2 && (\varepsilon < \tfrac{2-r^2}{2r+1}) \\ &= 2.\end{aligned}$$

So there exists a rational number greater than r which is an element of S, which means that r is not an upper bound for S, contradicting our assumption that $r = \sup(S)$.
If $r^2 < 2$, let $\varepsilon > 0$, then

$$\begin{aligned}(r-\varepsilon)^2 &= r^2 - \varepsilon(2r+\varepsilon) \\ &> r^2 - \varepsilon(2r+1) && (\varepsilon < 1) \\ &> r^2 - r^2 + 2 && (\varepsilon < \tfrac{r^2-2}{2r+1}) \\ &= 2.\end{aligned}$$

So there exists a rational number less than r which is an upper bound for S, which contradicts our assumption that $r = \sup(S)$.

Thus $r^2 = 2$, which we have already shown is impossible if $r \in \mathbb{Q}$, meaning that no supremum exists in $\mathbb{Q}$ for S. ■

3.2 Construction of the Reals

We are giving this real number system a lot of praise, but how can we actually show that it exists?

3.3 Subsets of the Real Line

It is helpful to note at this point that the completeness property from before also guarantees the existence of infimums for non-empty subsets that are bounded below.

Theorem 3.3 *Every non-empty subset of $\mathbb{R}$ that is bounded below has an infimum in $\mathbb{R}$.*

Proof. Let S be a non-empty subset of $\mathbb{R}$ that is bounded above. Then by the completeness property there exists $\ell \in \mathbb{R}$ such that $\ell = \sup(S)$. Now consider the set $S' = \{-s : s \in S\}$. It is clear that S' is a non-empty subset of $\mathbb{R}$ which is bounded below, and that $-\ell = \inf(S)$. ■

Quite often we will need some compact notation for denoting a continuous subset of $\mathbb{R}$, such as the positive real numbers, all real numbers between -1 and 4, etc. These subsets are called *intervals*, and we will define them now.

Definition 3.2. Let $a, b \in \mathbb{R}$ with $a < b$. Then the **closed interval** from a to b is given by

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\},$$

and the **open interval** from a to b is given by

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

Likewise the **half-open intervals** may be defined similarly as

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

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

To denote a half-unbounded or fully unbounded interval, we can use a ∞ symbol. For example:

$$[a, \infty) = \{x \in \mathbb{R} : a \leq x\},$$

or

$$(-\infty, 0) = \{x \in \mathbb{R} : x < 0\},$$

and we could also say

$$(-\infty, \infty) = \mathbb{R}.$$

It is also useful to note that

$$\sup((a, b)) = \sup([a, b]) = b, \quad \inf((a, b)) = \inf([a, b]) = a$$

.

3.4 Absolute Value

Definition 3.3. Let $x \in \mathbb{R}$. Then the **absolute value of** x is defined as

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

The absolute value has several useful properties.

Theorem 3.4 (Properties of Absolute Value) *Let* $x, y \in \mathbb{R}$*, then*

(i) $|x| \geq 0$,

(ii) $|x| = 0 \iff x = 0 \ \big(|x-y| = 0 \iff x = y\big)$,

(iii) $|-x| = |x| = \max(x, -x)$,

(iv) $x \leq |x|$,

(v) $|x| \leq a \iff -a \leq x \leq a,$ *for* $a \in \mathbb{R}$,

(vi) *Similarly,* $|x| \geq a \iff x \leq -a$ *or* $x \geq a$,

(vii) $|xy| = |x||y|$.

Proof. (i), (ii), (iii), and (iv) all follow immediately from the definition.

(v) Since $|x| \geq 0, a \geq 0$. Therefore clearly $x \geq -a$ and $x \leq |x| \leq a$.

(vi) Follows similarly.

(vii) The result follows immediately if x or y is 0, or if x and y are both either positive negative. Lets assume without loss of generality that $x < 0$, then, noting that $xy = -|xy|$ and $x = -|x|$ we find

$$|xy| = -xy = (-x)(y) = |x||y|.$$

■

These basic properties allow us to prove a basic theorem about inequalities with absolute values which will be absolutely indispensable in more advanced proofs.

Theorem 3.5 (Triangle Inequalities) *Let* $x, y \in \mathbb{R}$*, then we have*

(i) *(Triangle Inequality)* $|x+y| \leq |x| + |y|$,

(ii) *(Reverse Triangle Inequality)* $\big||x| - |y|\big| \leq |x-y|$.

Proof. The proof of these statements follow directly from the basic properties of the absolute value, with the reverse triangle inequality following directly from the normal triangle inequality.

(i) From (iv) we have $-|x| \leq x \leq |x|$ and similarly for y. Adding these two inequalities we get

$$\begin{aligned} -|x| - |y| &\leq x + y \leq |x| + |y| \\ -\big(|x| + |y|\big) &\leq x + y \leq |x| + |y|, \end{aligned}$$

and thus from (v) we have $|x+y| \leq |y|$.

(ii) Following from the triangle equality, note that

$$\begin{aligned} |x| &= |(x-y)+y| \leq |x-y| + |y| \\ |y| &= |(y-x)+x| \leq |y-x| + |x|. \end{aligned}$$

Rearranging each equation in turn, we get

$$\begin{aligned} |x| - |y| &\leq |x-y| \\ |y| - |x| = -\big(|x| - |y|\big) &\leq |y-x| = |x-y|, \end{aligned}$$

and so using (vi) we obtain $\big||x| - |y|\big| \leq |x-y|$.

■

3.5 Inequalities

Manipulating inequalities is an extremely useful trick in analysis because they appear everywhere. We have already discovered and proved the triangle inequality, and now we will discuss a few more.

Theorem 3.6 (General Triangle Inequality)

Theorem 3.7 (AM-GM Inequality)

Theorem 3.8 (Cauchy-Swartz Inequality)

Chapter 4

Sequences

4.1 Sequences & Convergence

A sequence is an infinitely long ordered list of real numbers. The precise definition is as follows.

Definition 4.1. A **sequence** of real numbers is a function

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

We denote $f(n)$ as x_n and write the sequence as

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

The elements of a sequence can be given by a formula for the n-th term, a recurrence relation, or a pattern given by the first few terms.

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

Some sequences converge to a fixed value, or **limit**, meaning that as n gets larger and larger, x_n gets closer and closer to the limit. An important question to ask is how do we define if a sequence converges or not, or how do we find the limit if it is only reached after an infinite amount of steps? For example, the sequence $\left(\frac{1}{n}\right)_n$ from above intuitively converges to 0 as n tends to infinity, but how can we prove this mathematically?

Definition 4.2. Let $(x_n)_n$ be a sequence of real numbers and let $x \in \mathbb{R}$.

1. $(x_n)_n$ is a **convergent sequence** with **limit** x if

   $$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \ :\ n \geq N \implies |x_n - x| \leq \varepsilon.$$

   In english this says that 'for all $\varepsilon > 0$, there exists a natural number N such that for $n \geq N$ the distance between x_n and the limit is less than ε'. We say $(x_n)_n$ **tends to x as n tends to infinity**.

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

$$\forall K > 0,\ \exists N \in \mathbb{N}\ :\ n \geq N \implies x_n \geq K.$$

3. Likewise, $(x_n)_n$ **tends to minus infinity** as n tends to infinity if

$$\forall K > 0,\ \exists N \in \mathbb{N}\ :\ n \geq N \implies x_n \leq -K.$$

4. A sequence is **divergent** if the limit does not exist or it tends to plus/minus infinity.

Let's take a closer look at definition 4.2.1. Being unfamiliar with techniques in analysis, it might look very confusing, but it is actually quite a simple and clever idea. The key thing to realise is that converging sequences *never actually reach their limit* (unless they are constant), so in order to capture this idea of reaching a limit after infinitely many terms in concrete mathematics, we introduce this idea of bounding the elements of the sequence within a certain distance of the limit. Therefore, what this definition is really saying is that *no matter how small* epsilon is, there exists a step in the sequence such that from that point onwards all values in the sequence are inside the interval $(x - \varepsilon, x + \varepsilon)$. Once this idea is understood, everything in analysis starts to make sense! Let us now look at a few examples.

Example 4.1. Claim: The sequence $\left(\frac{1}{n}\right)_n$ tends to 0 as n tends to infinity.
Proof. Let $\varepsilon > 0$ and choose $N > \frac{1}{\varepsilon}$. Then for $n \geq N$ we have

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

Example 4.2. Claim: $\left(\frac{1}{n^2}\right)_n \to 0$ as $n \to \infty$.
Proof. Let $\varepsilon > 0$ and choose $N > \frac{1}{\sqrt{\varepsilon}}$. Then for $n \geq N$ we have

$$|x_n - x| = \left|\frac{1}{n^2} - 0\right| = \frac{1}{n^2} \leq \frac{1}{N^2} < \varepsilon.$$

It is important to note that we do not have to always use the most optimal value for N. In this proof, choosing $N > \frac{1}{\varepsilon}$ would have been sufficient because $\frac{1}{N^2} \leq \frac{1}{N}$.

Example 4.3. Claim: $x_n = \frac{2n+1}{3n+5}$, $(x_n)_n \to \frac{3}{2}$ as $n \to \infty$.

Proof. Let $\varepsilon > 0$ and choose $N > \frac{7}{9}\varepsilon$. Then for $n \geq N$ we have

$$\begin{aligned}|x_n - x| &= \left|\frac{2n+1}{3n+5} - \frac{2}{3}\right| = \left|\frac{3(2n+1) - 2(3n+5)}{3(3n+5)}\right| \\ &= \left|\frac{6n - 6n + 3 - 10}{9n+15}\right| = \left|\frac{-7}{9n+15}\right| = \frac{7}{9n+15} \\ &\leq \frac{7}{9N+15} \leq \frac{7}{9N} < \varepsilon.\end{aligned}$$

Example 4.4. Claim: $(n^3)_n \to \infty$ as $n \to \infty$.
Proof. Let $K > 0$ and choose $N > \sqrt[3]{K}$. Then for $n \geq N$ we have

$$x_n = n^3 \geq N^3 > K.$$

Example 4.5. Claim: $x_n = (-1)^n$, $(x_n)_n$ is divergent.
Proof. Suppose by way of contradiction that there is a limit $x \in \mathbb{R}$.
Let $\varepsilon = \frac{1}{2}$, then since the limit exists there exists N such that $n \geq N$ implies $|x_n - x| < \frac{1}{2}$.
Since there are arbitrarily large even numbers, there is always some $n \geq N$ (for example $n = 2N$) such that $x_n = 1$, so $|1 - x| < \frac{1}{2} \implies x > \frac{1}{2}$.
Also, since there are arbitrarily large odd numbers, there is always some $n \geq N$ ($n = 2N + 1$) such that $x_n = -1$, so $|-1 - x| < \frac{1}{2} \implies x < -\frac{1}{2}$.
But then $\frac{1}{2} < x < -\frac{1}{2}$, which is impossible, so the limit cannot exist.

Theorem 4.1 (Standard Sequences) *Let $a \in \mathbb{R}$.*

(i) *Let $x_n = a$, then $(x_n)_n \to a$ as $n \to \infty$ (constant sequence).*

(ii) *If $|a| < 1$, then $(a^n)_n \to 0$ as $n \to \infty$.*

(iii) *If $a > 1$, then $(a^n)_n \to \infty$ as $n \to \infty$.*

Proof.

(i) Let $\varepsilon > 0$. Then given $N \in \mathbb{N}$, $n \geq N$ implies

$$|x_n - x| = |x - x| = 0 < \varepsilon.$$

(ii) Let $\varepsilon > 0$ and choose $N > \frac{\ln \varepsilon}{\ln |a|}$. Then for $n \geq N$ we have

$$|x_n - x| = |a^n - 0| = |a^n| \leq |a^N| < \varepsilon.$$

(iii) Let $K > 0$ and choose $N > \frac{\ln K}{\ln a}$. Then for $n \geq N$ we have

$$a^n \geq a^N > K.$$

■

4.2 Properties of Limits

Now that we have established how to prove the existence of limits, we will consider the uniqueness of them, as well as some algebraic properties of convergent sequences.

Theorem 4.2 *If a sequence converges, its limit is unique.*

Proof. Let $(x_n)_n$ be a convergent sequence and suppose that $(x_n)_n \to x$ and $(x_n)_n \to y$. Let $\varepsilon > 0$. Then there exists $N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |x_n - x| < \varepsilon$, but there also exists $N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |x_n - y| < \varepsilon$. Now,

$$\begin{aligned} |x - y| &= |x - x_n + x_n - y| && (\text{for } n \geq \max\{N_1, N_2\}) \\ &\leq |x - x_n| + |x_n - y| \\ &< \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

Therefore, $|x - y| = 0$, so $x = y$ (since $\forall \varepsilon > 0,\ 0 \leq a < \varepsilon \implies a = 0\ \forall a \in \mathbb{R}$). Thus, the two limits are in fact the same. ■

The next theorem will serve as a stepping stone that will make some more useful theorems much easier to prove. First, recall that a set is bounded if and only if it is bounded both above and below. We extend this definition to the elements of sequences, saying that a sequence is **bounded** if there exists $M \geq 0$ such that $|x_n| \leq M\ \forall n \in \mathbb{N}$. This is a very quick and easy way to establish the range of a sequence that will come in handy later.

Theorem 4.3 *Every convergent sequence is bounded*[a].

[a]Note that the converse of this theorem is obviously not true, $((-1)^n)_n$ is bounded, but does not converge.

Proof. Let $(x_n)_n$ be a convergent sequence with $x_n \to x$. Then there exists $N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < 1$. Consider the terms in the sequence $x_1, x_2, \ldots, x_{N-1}$ and let $M = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|\}$. Now, for all $n \geq N$, $|x_n| \leq \max\{M, |x| + 1\}$ (since $|x_n - x| \geq |x_n| - |x| < 1 \implies |x_n| < 1 + |x|$), which is a constant, hence $(x_n)_n$ is bounded. ■

Now that we have this tool, we can proceed to some more interesting results about how limits of sequences behave when we add and multiply sequences together.

Theorem 4.4 *Let $a \in \mathbb{R}$, $(x_n)_n$ and $(y_n)_n$ converging sequences with $x_n \to x$ and $y_n \to y$.*

(i) $x_n + y_n \to x + y$

(ii) $x_n y_n \to xy$

(iii) *If* $\forall n, y_n \neq 0$ *and* $y \neq 0$, *then* $\frac{1}{y_n} \to \frac{1}{y}$

(iv) *If* $\forall n, y_n \neq 0$ *and* $y \neq 0$, *then* $\frac{x_n}{y_n} \to \frac{x}{y}$

(v) $ax_n \to ax$

(vi) *If* $\forall n, x_n \leq y_n$, *then* $x \leq y$.

Proof. For all parts of this proof, begin by letting $\varepsilon > 0$.

(i) Since $x_n \to x$, $\exists N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |x_n - x| < \frac{\varepsilon}{2}$ and since $y_n \to y$, $\exists N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |y_n - y| < \frac{\varepsilon}{2}$. Choose $N = \max\{N_1, N_2\}$, now for $n \geq N$ we have

$$\begin{aligned} |(x_n + y_n) - (x + y)| &= |(x_n - x) + (y_n - y)| \\ &\leq |x_n - x| + |y_n - y| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(ii) Since $(x_n)_n$ is bounded, $\exists M \geq 0$ such that $|x_n| \leq M$ $\forall n \in \mathbb{N}$. Since $x_n \to x$, $\exists N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |x_n - x| < \frac{\varepsilon}{2M}$ and since $y_n \to y$, $\exists N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |y_n - y| < \frac{\varepsilon}{2|y|}$. Now, choosing $N = \max\{N_1, N_2\}$, for $n \geq N$ we have

$$\begin{aligned} |x_n y_n - xy| &= |x_n y_n - x_n y + x_n y - xy| \\ &= |x_n(y_n - y) + y(x_n - x)| \\ &\leq |x_n||y_n - y| + |y||x_n - x| \\ &< |x_n|\frac{\varepsilon}{2M} + |y|\frac{\varepsilon}{2|y|} \\ &\leq M\frac{\varepsilon}{2M} + |y|\frac{\varepsilon}{2|y|} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(iii) Since $y_n \to y$, $\exists N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |y_n - y| < \frac{|y|}{2}$. Then for $n \geq N_1$

$$|y| = |y_n + (y - y_n)| \leq |y_n| + |y_n - y| < |y_n| + \frac{|y|}{2},$$

so $|y_n| > \frac{|y|}{2}$. Now, let N_2 be such that $n \geq N_2 \implies |y_n - y| < \frac{\varepsilon |y|^2}{2}$, then for $n \geq \max\{N_1, N_2\}$ we have

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

(iv) Write $\frac{x_n}{y_n}$ as $x_n \cdot \frac{1}{y_n}$, then the result follows from parts (ii) and (iii).

(v) Let $y_n = a$ (so $(y_n)_n$ is the constant sequence $(a, a, \dots)$ with limit a), then the result follows from part (ii).

(vi) We will make use of a minor result to assist in proving this.

Lemma *Suppose $(a_n)_n$ is a convergent sequence with $a_n \to a$. If $a_n > 0 \ \forall n \in \mathbb{N}$ then $a \geq 0$.*

Proof. Suppose by way of contradiction that $a < 0$ and let $\varepsilon = -a > 0$. Then since $(a_n)_n$ converges, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |a_n - a| < -a$. This means that

$$\begin{aligned}& a < a_n - a < -a \\ \implies & a_n - a + a < -a + a \\ \implies & a_n < 0\end{aligned}$$

which is a contradiction, hence a must be greater than 0. ■

Now, consider $y_n - x_n \geq 0 \ \forall n \in \mathbb{N}$. The sequence $(y_n - x_n)_n$ converges to $y - x$ by part (i) (and part (v) with $a = -1$) and so by our claim $y - x \geq 0$, which implies $x \leq y$ as required.

■

A little observation from the last part of this theorem is the **sandwich principle**, which is that if we have three sequences — $(a_n)_n$, $(x_n)_n$ and $(b_n)_n$ with limits a, x and b respectively — with $a_n \leq x_n \leq b_n \ \forall n \in \mathbb{N}$, then $a \leq x \leq b$. In the special case that $a = b$, we can use this principle to determine the limit of an unknown sequence $(x_n)_n$ as in this case $a = x = b$. We will finish off this section with a couple of examples about how to use the arithmetic properties of limits.

Example 4.6. Let $(x_n)_n = \left(1 + \frac{1}{n^p}\right)_n$ with $p \in \mathbb{N}$. Show that $x_n \to 1$ as $n \to \infty$. First, let $a_n = 1$ and $b_n = \frac{1}{n^p}$, then note that we already know that the sequence $\left(\frac{1}{n}\right)_n$ converges to 0, so by theorem 4.4 (ii) (applied repeatedly),

$$\underbrace{\frac{1}{n} \cdot \frac{1}{n} \cdots \frac{1}{n}}_{p \text{ times}} = \frac{1}{n^p} \to 0 \text{ as } n \to \infty.$$

Then $(x_n)_n = (a_n + b_n)_n$ which, by theorem 4.4 (i), converges to 1 since $a_n \to 1$ and $b_n \to 0$ as $n \to \infty$.

Example 4.7. Show that the sequence $(x_n)_n = \left(\frac{n+5n^2+2n^3}{n^3+3}\right)$ converges to 2.

$$\frac{2n^3 + 5n^2 + n}{n^3 + 3} = \frac{2 + 5\left(\frac{1}{n}\right) + \left(\frac{1}{n}\right)^2}{1 + 3\left(\frac{1}{n}\right)^3} \to \frac{2 + 5(0) + 0}{1 + 3(0)} = 2 \text{ as } n \to \infty.$$

4.3 Monotone Sequences

So far, in all the examples of sequences we have seen so far we have had to know what the limit of the sequence is in order to prove the existence of it, but sometimes this is tricky to know beforehand if we are dealing with a novel sequence. We will now introduce some tools and theorems that will enable us to prove the existence of a limit for a sequence without knowing what the limit is.

Definition 4.3. Let $(x_n)_n$ be a sequence of real numbers.

- $(x_n)_n$ is **increasing** if $x_1 \leq x_2 \leq x_3 \leq \dots$
- $(x_n)_n$ is **decreasing** if $x_1 \geq x_2 \geq x_3 \geq \dots$
- $(x_n)_n$ is **monotone** if it is either increasing or decreasing.

Example 4.8. Here are some examples of sequences with their characterisation by this definition.

- $(a_n)_n = (n)_n$ is increasing.
- $(b_n)_n = \left(\frac{1}{n}\right)_n$ is decreasing.
- $(c_n)_n = (p, p, p, \dots),\ p \in \mathbb{R}$ is both increasing and decreasing.
- $(d_n)_n = (-2n)_n$ is decreasing.

- $(e_n)_n = ((-1)^n)_n$ is not monotone.
- $(f_n)_n = \left(\frac{(-1)^n}{n}\right)_n$ is not monotone.

We can make an important observation from these examples. The bounded sequences $((b_n)_n, (c_n)_n, (e_n)_n, (f_n)_n)$ that are also monotone $((b_n)_n, (c_n)_n)$ are all convergent, whereas none of the unbounded sequences are convergent (which we would expect by theorem 4.3). Although a sequence does not have to be monotone to converge (look at the last example above), it is a general result that if a monotone sequence is bounded then it does converge, which we will now prove.

Theorem 4.5 (Monotone Convergent Theorem (MCT)) *Let $(x_n)_n$ be a monotone sequence. Then $(x_n)_n$ is convergent if and only if $(x_n)_n$ is bounded.*

Proof. $(\implies)$: This follows from theorem 4.3.
$(\impliedby)$: Suppose $(x_n)_n$ is bounded and without meaningful loss of generality assume $(x_n)_n$ is increasing. Recall that since $\mathbb{R}$ is *complete*, every bounded set of real numbers has a supremum that is also a real number, i.e. if we let

$$X \coloneqq \{x_n : n \in \mathbb{N}\}$$

then X is a bounded subset of the real numbers and we can let $x = \sup X \in \mathbb{R}$. Now we want to show that $(x_n)_n$ converges to x. Let $\varepsilon > 0$ and note that $x - \varepsilon$ is not an upper bound since x is the supremum. Therefore, there exists $N \in \mathbb{N}$ such that $x_N \geq x - \varepsilon$. Now since $(x_n)_n$ is increasing, for $n \geq N$ we have

$$|x_n - x| = x - x_n \leq x - x_N \leq \varepsilon,$$

and hence $x_n \to x$ as $n \to \infty$. ■

This theorem makes intuitive sense, if we have a sequence which is always increasing or decreasing and is bounded then it must reach a limit before it surpasses the bound, and in fact this limit is the least upper/greatest lower bound of the set of elements of the sequence.

Corollary 4.5.1 *Suppose $(x_n)_n$ is a bounded sequence, then*

(i) *If $(x_n)_n$ is increasing,* $\lim_{n\to\infty} x_n = \sup_n x_n$.

(ii) *If $(x_n)_n$ is decreasing,* $\lim_{n\to\infty} x_n = \inf_n x_n$.

Proof. The proof follows directly from the proof of MCT. ■

Example 4.9. Let $(x_n)_n$ be a sequence given by

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

Show that this sequence is convergent.
First notice that $(x_n)_n$ is increasing,

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

and also the elements of $(x_n)_n$ are bounded between 0 and 2.

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

Hence by MCT, this sequence must converge. In fact, it can be shown that this sequence converges to $\frac{\pi^2}{6}$, but the proof of that is a bit more involved.

4.4 Subsequences

Now we want to introduce the idea of a subsequence, which is a sequence contained within another sequence. We can do this by taking a regular old sequence of real numbers and 'picking' certain entries out that we want to be part of the new subsequence, for example: every 2nd entry, every 5th entry, every entry whose index is a power of 10. We can do this by defining a new sequence of natural numbers, and then using this sequence to index our original one as follows.

Definition 4.4. Let $(x_n)_n$ be a sequence of real numbers and $(m_k)_k$ be a sequence of natural numbers which is strictly increasing ($m_1 < m_2 < m_3 < \dots$). Then $(x_{m_k})_k$ is a **subsequence** of $(x_n)_n$.

So, for example, we could start with the sequence $(x_n)_n = \left(\frac{1}{n^2}\right)_n$ and take $m_k = 2k$, then $x_{m_k} = x_{2k} = \frac{1}{(2k)^2}$. One important thing to note from this definition is that we *always* have $m_k \geq k \ \forall k \in \mathbb{N}$, this will be a useful fact going forward. Now, how can we know if a subsequence converges? We can use the methods we have discovered already, but also it turns out that if a sequence is convergent then any subsequence of it will also converge to the same limit.

Theorem 4.6 *Let $(x_n)_n$ be a convergent subsequence with $x_n \to x$ as $n \to \infty$. Then any subsequence $(x_{m_k})_k$ also has $x_{m_k} \to x$ as $k \to \infty$.*

Proof. Let $\varepsilon > 0$. Since $x_n \to x$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon$. Since $(x_{m_k})_k$ is a subsequence of $(x_n)_n$, we can let $K \in \mathbb{N}$ such that for $k \geq K$ we have $m_k \geq N$ (for example, choose $K = N$ since $m_k \geq k \geq N$). Then we have

$$|x_{m_k} - x| < \varepsilon \ \forall k \geq K,$$

which means that the subsequence converges to x. ■

Theorem 4.7 (Monotone Subsequence Theorem) *Any sequence of real numbers has a monotone subsequence.*

Proof. Let $(x_n)_n$ ba a sequence of real numbers. Before proving this theorem we make a quick definition. We say that a natural number p is a *peak* if $x_p \geq x_n \ \forall n \geq p$, i.e. it is the index of the greatest number in the sequence from that point onwards. Now, there are two cases for a general sequence: either it has infinitely many peaks, or it has finitely many.
Case 1: $(x_n)_n$ has infinitely many peaks. List out the peaks in order of which they occur — $p_1 < p_2 < p_3 \ldots$ — and notice that by definition we have

$$x_{p_1} \geq x_{p_2} \geq x_{p_3} \geq \ldots,$$

and thus $(x_{p_k})_k$ is a decreasing subsequence (which is monotone) as required.
Case 2: $(x_n)_n$ has finitely many peaks. Again we can list them in order: $p_1 < p_2 < p_3 < \cdots < p_N$. Now let $t_1 > p_N$ be a natural number. Since t_1 is not a peak by definition, 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 $x_{t_3} > x_{t_2}$. Continuing in this way we generate a sequence $t_1 < t_2 < t_3 < \ldots$ with $x_{t_1} < x_{t_2} < x_{t_3} < \ldots$, hence $(x_{t_k})_k$ is an increasing subsequence (which is monotone) as required. ■

Notice that if we wanted to create an increasing subsequence from a sequence with infinitely many peaks, or a decreasing subsequence from one with finitely many, we need only switch round the definition of a peak to that of a 'trough' by flipping the order relation and also flipping all the order relations in the body of the proof. Thus no generality is lost. We will now prove one of the major results of real analysis using all of the tools we have gathered so far.

Theorem 4.8 (Bolzano-Weierstrass Theorem) *Any bounded sequence of real numbers has a convergent subsequence.*

Proof. Let $(x_n)_n$ be any bounded sequence of real numbers. Then by the Monotone Subsequence Theorem (4.7), $(x_n)_n$ has a monotone subsequence $(x_{m_k})_k$. But now this subsequence is both monotone and bounded, so by the Monotone Convergence Theorem (4.5), this subsequence is convergent. ■

A simple example of this theorem in action would be the sequence $((-1)^n)_n$, which is bounded but does not converge, however the subsequence $((-1)^{2k})_k = (1)_k$ does converge.

4.5 Cauchy Sequences

Our goal was to develop tools that allow us to determine if a sequence is convergent without knowing what the limit of sequence is. We know that if a sequence is monotone and bounded then it is convergent, and that if a sequence is a subsequence of a bounded sequence, then it is convergent. But what if a sequence does not fulfill these properties? Our current definition of convergence requires knowledge of the limit in order to prove that a sequence is convergent, so how about we create a new definition of a convergent sequence where we don't need to know what the limit is. The intuition for this defintion is that we want to express mathematically that for a sequence to converge the terms have to get closer and closer together as $n \to \infty$. We will name this new type of convergent sequence *Cauchy* after the 19th-century French mathematician Augustin-Louis Cauchy.

Definition 4.5. Let $(x_n)_n$ be a sequence of real numbers. Then $(x_n)_n$ is a **Cauchy sequence** if

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N}\ :\ n, m \geq N \implies |x_m - x_n| \leq \varepsilon.$$

What this defintion is encoding for us is that for any small ε, there exists a step in the sequence such that from that step onwards, **all terms** of the sequence are ε-close to each other. It is actually quite important that we specify 'all terms' as we want this definition to be as close as possible to our original definition of convergence and if we say that, for example, only consecutive terms must get ε-close to each other then we admit some sequences which we would not consider convergent under the original definition. For example, the sequence $(x_n)_n = (\sqrt{n})_n$. Successive terms get arbitrarily close to each other ($x_{n+1} - x_n = \sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1}+\sqrt{n}} < \frac{1}{2\sqrt{n}}$) but not *all others* (the terms still get arbitrarily large as $n \to \infty$ so for any index n and distance M we can always find a large enough index m such that $x_m - x_n > M$), and thus the sequence is not Cauchy.

Example 4.10. Show that $\left(\frac{1}{n}\right)_n$ is a Cauchy sequence.

Let $\varepsilon > 0$ and choose $N > \frac{2}{\varepsilon}$. Then for $m, n \geq N$ we have

$$|x_m - x_n| = \left|\frac{1}{m} - \frac{1}{n}\right| \leq \frac{1}{m} + \frac{1}{n} \leq \frac{1}{N} + \frac{1}{N} = \frac{2}{N} < \varepsilon.$$

Example 4.11. Show that $\left(\frac{n}{1+n}\right)_n$ is a Cauchy sequence.
Let $\varepsilon > 0$ and choose $N > \frac{1}{\varepsilon}$. Then for $m \geq n \geq N$ we have

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

Example 4.12. Show that the sequence given by $x_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ is not Cauchy.
To prove this, we will show that the sequence fulfills the negation of the Cauchy sequence definition ($\exists \varepsilon > 0$ such that $\forall N \in \mathbb{N}\ \exists m, n \geq N$ such that $|x_m - x_n| > \varepsilon$).
Choose $\varepsilon < \frac{1}{2}$, let $N \in \mathbb{N}$, and choose $n = N$ and $m = 2N$. Then

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

Notice that in these examples, the sequences that are Cauchy are also convergent, and the sequence that is not Cauchy is not convergent. This correspondence turns out to be general and in fact it is an important result that the criteria for being convergent and being Cauchy are equivalent.

Theorem 4.9 *Let $(x_n)_n$ be a sequence of real numbers. Then $(x_n)_n$ is a convergent sequence if and only if it is also a Cauchy sequence.*

Proof.
($\implies$): We will make use of a lemma.

Lemma *Any Cauchy sequence is bounded.*

Proof. Let $(x_n)_n$ be a Cauchy sequence. Then $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x_N| < 1$ (this is equivalent to the definition of Cauchy with $m = N$). Then

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

Therefore, for

$$M = \max\{1 + |x_N|, |x_1|, |x_2|, \ldots, |x_{N-1}|\},$$

and thus for *all* $n \in \mathbb{N}$ we have $|x_n| \leq M$, so $(x_n)_n$ is bounded. ∎

Now suppose $(x_n)_n$ is a Cauchy sequence, then by the lemma above, $(x_n)_n$ is bounded and thus by the Bolzano-Weierstrass Theorem (4.8) there exists a convergent subsequence $(x_{m_k})_k$ which converges to some limit x. We now want to show that the Cauchy sequence $(x_n)_n$ also converges to x. Let $\varepsilon > 0$, Since $(x_n)_n$ is Cauchy $\exists N \in \mathbb{N}$ such that $m, n \geq N \implies |x_m - x_n| < \frac{\varepsilon}{2}$. Also, since $x_{m_k} \to x$, $\exists K \in \mathbb{N}$ such that $k \geq K \implies |x_{m_k} - x| < \frac{\varepsilon}{2}$. Then, for $n \geq \max\{N, K\}$ we have

$$|x_n - x| = |x_n - x_{m_n} + x_{m_n} - x| \leq |x_n - x_{m_k}| + |x_{m_k} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and thus $(x_n)_n$ is convergent with limit x.
$(\impliedby)$: Suppose $(x_n)_n$ is a convergent sequence with limit x. Let $\varepsilon > 0$, then $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \frac{\varepsilon}{2}$. For $m, n \geq N$ we have

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

and thus $(x_n)_n$ is a Cauchy sequence. ∎

Chapter 5

Series

5.1 Series

Outside the world of mathematics the words 'sequence' and 'series' can often mean the same thing, however, within the context of real analysis this is not the case. We define a series to be an infinite summation, which we can think of as a sum of all of the terms in a sequence. For example, we can take any sequence $(x_n)_n$ and form the series

$$\sum_{n=1}^{\infty} x_n = x_1 + x_2 + x_3 + \dots.$$

But when is it valid to write that a series has some finite value? If we add up infinitely many things there will be many cases where the sum is infinite, for example a series constructed from any constant sequence would just be that constant added up infinitely many times, which obviously has no finite value. To solve this problem, we introduce a new tool: the sequence of partial sums.

Definition 5.1. Let S be a series $\sum_{n=1}^{\infty} x_n$ generated by some sequence $(x_n)_n$. Then we define the n^{th} **partial sum** of S as

$$S_n = \sum_{k=1}^{n} = x_1 + x_2 + \dots + x_n.$$

It is clear that the n^{th} partial sums always have a finite value (if the elements of the sequence are well-defined) since they are only finite sums. This means we can consider the partial sums as a sequence in their own right $(S_n)_n$, and we can say that if the sequence of partial sums converges to some value, then the infinite series takes on that exact value.

Definition 5.2. Let S be a series $\sum_{n=1}^{\infty} x_n$. Then we say S **converges** if $(S_n)_n$

converges, in which case

$$S = \sum_{n=1}^{\infty} x_n = \lim_{n\to\infty} S_n.$$

If $(S_n)_n$ diverges, then $\sum_{n=1}^{\infty}$ diverges and S is either not defined, or we can say that $S = \infty$ or $S = -\infty$, depending on the nature of the divergence of $(S_n)_n$.

Example 5.1. Consider the sequence given by $x_n = \frac{1}{n(n+1)}$. The sequence of partial sums of the series generated by this sequence is given by

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

We can see that $(S_n)_n$ converges to 1, and hence $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ is a convergent series with the value 1.

Example 5.2. Consider the seqeuence given by $x_n = (-1)^n$. Then

$$S_1 = -1,\ S_2 = -1 + (-1)^2 = 0,\ S_3 = -1 + (-1)^2 + (-1)^3 = -1 \ldots$$

Thus the nth partial sum is given by

$$S_n = \begin{cases} -1 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

This sequence $(S_n)_n$ is divergent, and so the series $\sum_{n=1}^{\infty}(-1)^n$ is not defined.

In the first example we can see that the sequence which generated the convergent series converged to 0 (not the partial sums!). This turns out to be a general fact which we will now show.

Theorem 5.1 *Let $S = \sum_{n=1}^{\infty} x_n$ be a convergent series. Then $x_n \to 0$ as $n \to \infty$.*

Proof. Since S is convergent, the sequence $(S_n)_n$ converges. Note that $x_n = S_n -$

S_{n-1} for all $n \geq 2$, and since both $S_n \to S$ and $S_{n-1} \to S$ as $n \to \infty$, we have

$$x_n \to S - S = 0 \text{ as } n \to \infty.$$

■

It makes sense that this should be the case. If the generating sequence did not tend to 0, then we would be adding together infinitely many nonzero terms, and it would be very strange if this somehow converges! It would be useful if the converse of this theorem was also true ($x_n \to 0 \implies \sum_{n=1}^{\infty}$ converges), but this turns out not to be the case and leads to one of the most (in)famous counterexamples in mathematics.

Example 5.3 (The Harmonic Series)**.** The series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges to $+\infty$.
Fist note that $(S_n)_n$ is an increasing sequence since $S_n - S_{n-1} = \left(\sum_{k=1}^{n} \frac{1}{k}\right) - \left(\sum_{k=1}^{n-1} \frac{1}{k}\right) = \frac{1}{n} \geq 0$. Now consider the subsequence $(S_{2^m})_m$. The general term of this sequence is given by

$$\begin{aligned} S_{2^m} =& 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \cdots + \frac{1}{16}\right) \\ &+ \cdots + \left(\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}\right), \end{aligned}$$

where the brackets indicate how the terms are added on in each partial sum. The ith bracketed expression in each term is given by

$$t_i = \frac{1}{2^{i-1}+1} + \frac{1}{2^{i-1}+2} + \cdots + \frac{1}{2^i}.$$

There are $2^i - 2^{i-1}$ values in this sum, and they are all greater than or equal to $\frac{1}{2^i}$, so $t_i \geq 2^{i-1}\frac{1}{2^i} = \frac{1}{2}$. Hence

$$S_{2^m} = 1 + t_1 + t_2 + \cdots + t_m \geq 1 + \frac{m}{2}.$$

Now, let $K \in \mathbb{N}$, then for $n \geq 2^{2K}$ we have

$$S_n \geq S_{2^{2K}} \geq 1 + \frac{2K}{2} = 1 + K > K,$$

and so since the sequence of partials sums diverges to ∞, the series must also diverge.

Definition 5.3. A **geometric series** is an infinite series of terms which have a constant ratio between them. They can be expressed as

$$\sum_{n=1}^{\infty} ar^{n-1}$$

for some $a \in \mathbb{R}$ and $r \in \mathbb{R} \setminus \{0\}$.

Theorem 5.2 *The geometric series $\sum_{n=1}^{\infty} ar^{n-1}$ converges if and only if $|r| < 1$, and in this case $\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}$.*

Proof. First note the general term of the sequence of partial sums:

$$S_n = \sum_{k=1}^{n} ar^{k-1} = a + ar + ar^2 + \cdots + ar^{k-1} = \frac{a(1-r^n)}{1-r}. \qquad (r \neq 1)$$

If $|r| < 1$, then by theorem 4.1 $r^n \to 0$ as $n \to \infty$. Hence $S_n \to \frac{a}{1-r}$. If $r = 1$, then $S_n = an$ which tends to $\pm\infty$ depending on the sign of a. If $r = -1$, then S_n is 0 if n is even and a if n is odd, hence S_n diverges. It can be shown that if $|r| > 1$, S_n diverges, hence all the cases are covered and the only way the series can converge is if $|r| < 1$. ■

5.2 The Comparison Test

What might have become clear is that determining convergence of series is more complicated than for sequences because the terms themselves are more complicated. Because of this,we want to create tools which will allow us to more easily deduce the properties of a series. One of these is the comparison test, which allows us to determine if a series is convergent by comparing it to a known series such as $\sum_{n=1}^{\infty} \frac{1}{n}$ or $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Theorem 5.3 (Comparison Test) *Let $(x_n)_n$ and $(a_n)_n$ be sequence with no negative terms. Then*

(i) *If $\sum_{n=1}^{\infty} a_n$ is convergent and $\exists N \in \mathbb{N}$ such that $x_n \leq a_n$ $\forall n \geq N$, then $\sum_{n=1}^{\infty} x_n$ also converges.*

(ii) *If $\sum_{n=1}^{\infty} a_n$ is divergent and $\exists N \in \mathbb{N}$ such that $x_n \geq a_n$ $\forall n \geq N$, then $\sum_{n=1}^{\infty} x_n$ also diverges.*

Proof.

(i) Let $A = \sum_{n=1}^{\infty} a_n$ and note that since all the terms x_n are positive, $S_n =$

$\sum_{k=1}^{n} x_k$ is an increasing sequence. Then, since there exists some integer N such that for all $k \geq N$ we have $x_n \leq a_n$, we get

$$\sum_{k=N}^{n} x_k \leq \sum_{k=N}^{n} a_k \leq \sum_{k=1}^{\infty} a_k = A.$$

Therefore, for all $n \geq 1$ we have

$$\begin{aligned} S_n &= a_1 + a_2 + \cdots + a_{N-1} + \sum_{k=N}^{n} x_k \\ &\leq a_1 + a_2 + \cdots + a_{N-1} + A, \end{aligned}$$

and hence S_n is bounded above. But now S_n is both monotone and bounded and so by the MCT (4.5), S_n must be convergent and thus the series $\sum_{n=1}^{\infty} x_n$ converges.

(ii) Suppose by way of contradiction that $\sum_{n=1}^{\infty} x_n$ is convergent, then following the argument in part (i) implies that $\sum_{n=1}^{\infty} a_n$ is convergent, which is a contradiction, meaning $\sum_{n=1}^{\infty} x_n$ must be divergent.

$\blacksquare$

Note that we can also multiply the terms of the generating sequence by a positive constant, as this does not affect the convergence of the series, i.e. we can check that $\exists N \in \mathbb{N}, c > 0$ such that $x_n \leq ca_n$ or $x_n \geq ca_n$ $\forall n \geq N$.

We using the comparison test, often the most convenient series to compare to are the so-called 'p-series', that is, those of the form $\sum_{n=1}^{\infty} \frac{1}{n^p}$.

Example 5.4. Let $p \in \mathbb{R}$. Then the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$.
Suppose $p \geq 2$. Then $\frac{1}{n^p} \leq \frac{1}{n^2}$ for all $n \in \mathbb{N}$, so $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent by the comparison test.
Now suppose $p \leq 1$. Then $\frac{1}{n^p} \geq \frac{1}{n}$ for all $n \in \mathbb{N}$, so $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is divergent by the comparison test.
In the case of $p \in (1, 2)$, the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ can be proved to be convergent by the 'integral test'.

Theorem 5.4 (Ratio Test) *Let $(x_n)_n$ be a sequence of positive terms.*

(i) *If $\exists N \in \mathbb{N}$ and $r < 1$ such that for all $n \geq N$, $\frac{x_{n+1}}{x_n} \leq r$, then $\sum_{n=1}^{\infty} x_n$ is convergent.*

(ii) *If $\exists N \in \mathbb{N}$ and $r > 1$ such that for all $n \geq N$, $\frac{x_{n+1}}{x_n} \geq r$, then $\sum_{n=1}^{\infty} x_n$ is* *divergent.*

Proof.

(i) First note that $x_{N+1} \le r x_N$, $x_{N+2} \le r x_{N+1} = r^2 x_N$, etc. and so $x_{N+k} \le r^k x_N$. Now consider the sequence given by

$$a_n = r^{n-N} x_N = (x_N r^{-N}) \cdot r^n.$$

This is a geometric series which is convergent by theorem 5.2. Then notice that for all $n \ge N$ we have

$$x_n = x_{N+(n-N)} \le r^{n-N} x_N = a_n,$$

and hence by the comparison test $\sum_{n=1}^{\infty} x_n$ is convergent.

■

In effect, what the ratio test says is that if $\lim_{n\to\infty} \frac{x_{x+1}}{x_n}$ exists and is less than 1, then the series converges and if it is greater than 1 then the series diverges. If the limit is equal to 1, then we cannot determine the behaviour of the series from the ratio test. The ratio test works best for geometric series because of the cancellation that happens when dividing the exponents. It also works well for some series containing factorials of n for the same reason.

Example 5.5. Show that the series $\sum_{n=1}^{\infty} \frac{2^n}{n!}$ converges.
The generating sequence is given by $x_n = \frac{2^n}{n!}$, so

$$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{2^{n+1}}{(n+1)!}\right)}{\left(\frac{2^n}{n!}\right)} = \frac{2^{n+1} n!}{2^n (n+1)!} = \frac{2}{n+1} \to 0 \text{ as } n \to \infty.$$

Hence by the ratio test, the series converges.

5.3 Series of Positive & Negative Terms

Up to this point we have only looked at methods for analysing the behaviour of series generated by sequences which only contain positive terms. This is very limiting and we will now expand the scope to general sequences.

Definition 5.4. A series $\sum_{n=1}^{\infty} x_n$ is called **absolutely convergent** if $\sum_{n=1}^{\infty} |x_n|$ converges.

Theorem 5.5 *Every absolutely convergent series is also convergent.*

Proof. Consider an absolutely convergent series $\sum_{n=1}^{\infty} x_n$. Note that we have

$$x_n + |x_n| \leq 2|x_n|, \ \forall n \in \mathbb{N}.$$

Thus, by the comparison test (comparison with $\sum_{n=1}^{\infty}|x_n|$), the series $\sum_{n=1}^{\infty}(x_n + |x_n|)$ is convergent. Now, we can write

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty}(x_n + |x_n|) - \sum_{n=1}^{\infty}|x_n|,$$

and since $\sum_{n=1}^{\infty} x_n$ is the different of two convergent series, it is also convergent (follows from theorem 4.4). ∎

Example 5.6. Show that the series $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ is convergent.

Note that since $\left|\frac{(-1)^n}{n^2}\right| = \frac{1}{n^2}$, by comparison with the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ the series $\sum_{n=1}^{\infty}\left|\frac{(-1)^2}{n^2}\right|$ is convergent. Hence the series $\sum_{n=1}^{\infty} \frac{(-1)^2}{n^2}$ is absolutely convergent and therefore convergent.

Here are two more convergence tests which will be useful.

Theorem 5.6 (The Root Test) *Let $\sum_{n=1}^{\infty} x_n$ be a series.*

(i) *If $\exists N \in \mathbb{N}$ and $r < 1$ such that for all $n \geq N, |x_n|^{\frac{1}{n}} \leq r$, then the series converges.*

(ii) *If $\exists N \in \mathbb{N}$ and $r > 1$ such that for all $n \geq N, |x_n|^{\frac{1}{n}} \geq r$, then the series diverges.*

Proof.

(i) For $n \geq N$ we have $|x_n|^{\frac{1}{n}} \leq r \implies |x_n| \leq r^n$. Let $a_n = r^n$, then $(a_n)_n$ generates a convergent geometric series so by the comparison test the series $\sum_{n=1}^{\infty}|x_n|$ converges. Thus the series $\sum_{n=1}^{\infty} x_n$ is absolutely converges and therefore converges.

(ii) For all $n \geq N$ we have $|x_n|^{\frac{1}{n}} \geq r \implies |x_n| \geq r^n$, so x_n must be a divergent sequence ($\pm\infty$). In particular, x_n does not converge to 0 as $n \to \infty$, so by the contrapositive of theorem 5.1, $\sum_{n=1}^{\infty} x_n$ diverges.

∎

As with the ratio test, the root test says is that if $\lim_{n\to\infty}|x_n|^{\frac{1}{n}}$ exists and is less than 1, then the series converges, if it is greater than 1 then the series diverges and if it is equal to 1 then the root test gives us no information.

Theorem 5.7 (Alternating Series Test) *Let $(x_n)_n$ be a positive decreasing sequence with limit 0 ($\lim_{n\to\infty} x_n = 0$). Then*

$$\sum_{n=1}^{\infty}(-1)^n x_n$$

is a convergent series.

Proof. ■

Example 5.7. The sequence given by $x_n = \frac{1}{n}$ is positive and decreasing with limit 0. Thus by the alternating series test, $\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ is convergent. Note that it is not absolutely convergent since $\left|\frac{(-1)^n}{n}\right| = \frac{1}{n}$ generates the harmonic series which diverges.

Theorem 5.8 *These are some handy facts to keep in mind when using the root test.*

(i) $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$.

(ii) *If* $r > 0$, *then* $\lim_{n\to\infty} r^{\frac{1}{n}} = 1$.

(iii) *Suppose $(a_n)_n$ is a positive sequence and $\frac{a_{n+1}}{a_n} \to r$, then $a_n^{\frac{1}{n}} \to r$.*

Proof. ■

5.4 Power Series

A power series is a series of the powers of some unknown variable, basically an infinite polynomial. We make a clearer definition here.

Definition 5.5. Let $(a_n)_n$ be a sequence of real numbers. Then a power series in some variable x is given by

$$\sum_{n=1}^{\infty} a_n x^n.$$

If we compare the form of a power series to a geometric series, we can see that the convergence of a power series depends on the value of the variable x, so the power series is a function of x. In particular, it can be shown that a power series either converges for all $x \in \mathbb{R}$, or it converges in some interval of the real line symmetric about $x = 0$. For example, if $(a_n)_n$ is the constant sequence $(1)_n$, then the power series $\sum_{n=1}^{\infty} x^n$ converges for $-1 < x < 1$.

Theorem 5.9 *Consider a power series $\sum_{n=1}^{\infty} a_n x^n$ and suppose $\lim_{n\to\infty}|a_n|^{\frac{1}{n}} = \beta$. Then, taking $R = \frac{1}{\beta}$ (if $\beta = 0$, then take $R = \infty$; if $\beta = \infty$, then take $R = 0$), we have*

(i) *The power series converges for $x \in (-R, R)$.*

(ii) *The power series diverges for $x \notin (-R, R)$.*

R is called the ***radius of convergence*** *of the power series.*

Proof. Fix $x \in \mathbb{R}$. Then

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

(i) (a) $\beta = 0$ (so $R = \infty$ and $|x| < R$) implies $\beta|x| < 1$, so by the root test $\sum_{n=1}^{\infty} a_n x^n$ converges.
   (b) $\beta > 0$ and $\beta|x| < 1$ (so $|x| < R$). By the root test $\sum_{n=1}^{\infty} a_n x^n$ converges.

(ii) (a) $\beta = \infty$ (so $R = 0$). Then the root test implies that $\sum_{n=1}^{\infty} a_n x^n$ diverges.
   (b) $\beta > 0$ and $\beta|x| > 1$ (so $|x| > R$). Then by the root test $\sum_{n=1}^{\infty} a_n x^n$ diverges.

■

Note that when we are finding R, it can often be much easier to find the limit $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ than $\lim_{n\to\infty}|a_n|^{\frac{1}{n}}$, and theorem 5.8 (iii) says that they have the same value.

Example 5.8. For what values of x does the power series $\sum_{n=1}^{\infty} \frac{n+1}{2^n} x^n$ converge? Using the root test, we see that

$$|a_n|^{\frac{1}{n}} = \left(\frac{n+1}{2^n}\right)^{\frac{1}{n}} = \frac{(n+1)^{\frac{1}{n}}}{2} \to \frac{1}{2} \text{ as } n \to \infty.$$

So the radius of convergence of this power series is $R = 2$. What happens for the bounding values? For $x = 2$, the power series becomes $\sum_{n=1}^{\infty} \frac{n+1}{2^n} 2^n = \sum_{n=1}^{\infty}(n+1)$ which is divergent. For $x = -2$, the power series is $\sum_{n=1}^{\infty} \frac{n+1}{2^n}(-2)^n = \sum_{n=1}^{\infty}(-1)^n(n+1)$ which is divergent. Hence the power series converges $\forall x \in (-2, 2)$.

Example 5.9. For what values of x does the power series $\sum_{n=1}^{\infty} \frac{1}{n!} x^n$ converge? Inspecting the ratio of consecutive terms, we see

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

Hence $\beta = 0$ and $R = \infty$, so the power series converges $\forall x \in \mathbb{R}$.

Chapter 6

Continuous Functions

6.1 The Basic Definition of a Function

The concept of a function is one that is fundamental to all area of mathematics. We will be using functions thoroughly from now on so it's a good time to review exactly what a function is and how we can manipulate them.

Definition 6.1. Let D and R be non-empty sets. Then $f \subseteq D \times R$ is a function if $\forall x \in D, \forall y, z \in R,\ (x, y) \in f$ and $(x, z) \in f \implies y = z$.

In English this says that a function is a binary relation on two sets, and what makes a function different from any old binary relation is the property that every $x \in D$ is paired with a *unique* $y \in R$.

Definition 6.2. Let $f \subseteq D \times R$ be a function. Then we say $f : D \to R$ and denote the unique $y \in R$ associated with $x \in D$ as $f(x)$.

- D is **domain** of the function, denoted $\text{dom}(f)$.
- R is the **range**[a] of the function, denoted $\text{range}(f)$.
- The **image** of f, $\text{im}(f)$, is defined as $\{f(x) : x \in D\}$, or more explicitly $\{y \in R : \exists x \in D, f(x) = y\}$. The image is a subset of the range, and may or may not be a proper subset.
- For a given x, $f(x)$ may be referred to as the **image** of x under f, or the **value** of f at x, or the **output** of f for the **input** x.

[a] Some people use the term **range** in an ambiguous way. A less ambiguous term is **codomain**, which means the same thing as we have just defined for range.

To express what the image of some function f actually is, we need some kind of formula for the image of an arbitrary point $x \in D$. This can be given by a straightforward formula such as

$$f(x) = 2x \quad \text{or} \quad x \mapsto x^2 + 1,$$

or a piecewise definition like

$$f(x) = \begin{cases} x & x \leq 1 \\ x^2 + 1 & x > 1, \end{cases} \tag{6.1}$$

or something even more abstract, such as a power series or recurrence relation (Recall that sequences are defined as functions).

A real function is a function that has real numbers as its inputs and outputs. We can visualise a real functions behaviour by treating the ordered pairs $\big(x, f(x)\big) \in f$ as Cartesian coordinates in the plane, which is what we all know as a graph.

Example 6.1. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \sqrt{x}$.
This is not a function since $\sqrt{x}$ is not defined in $\mathbb{R}$ for $x < 0$. Even if we restrict the domain of f to $[0, \infty)$, it is still technically not a function since there are two possible values of $\sqrt{x}$ for all $x \in [0, \infty)$. We can fix this by taking the absolute value ($f : [0, \infty) \to \mathbb{R}$ given by $f(x) = |\sqrt{x}|$ is a function).

Example 6.2. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

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

This is again a valid function, and it has numerical values, but drawing a convincing graph of this function is left as a challenge to the reader.

6.2 Operations on Functions

Now we will define how to combine functions together to create new functions. Let f, g be functions with domains and ranges in $\mathbb{R}$.

Definition 6.3 (Basic Arithmetic)**.** Firstly, it is useful to have a working definition of equality.

- $f = g \iff D := \operatorname{dom}(f) = \operatorname{dom}(g)$ and $f(x) = g(x)\ \forall x \in D$.

For simple operations like the sum and product, we define the new domain simply as the common domain of the operands.

- $\operatorname{dom}(f + g) = \operatorname{dom}(f \cdot g) = \operatorname{dom}(f) \cap \operatorname{dom}(g)$.

- $(f+g)(x) = f(x) + g(x)$.
- $(f \cdot g)(x) = f(x)g(x)$.

In the special case of multiplying by a constant function c given by $x \mapsto k, k \in \mathbb{R}$,

- $(c \cdot f)(x) = c(x)f(x) = kf(x)$.

Using this definition with the constant function $x \mapsto -1$, we can define the differences of functions.

The reciprocal $1/g$ is defined where g is not zero.

- $\text{dom}\left(\frac{1}{g}\right) = \{x \in \text{dom}(g) : g(x) \neq 0\}$.
- $\left(\frac{1}{g}\right)(x) = \frac{1}{g(x)}$.

Using this we can define quotients of functions as a product of a function with a reciprocal.

Definition 6.4 (Composition)**.** The composition $f \circ g : \text{dom}(g) \to \text{range}(f)$ is defined only where $\text{im}(g)$ overlaps with $\text{dom}(f)$, so we say:

- $\text{dom}(f \circ g) = \{x \in \text{dom}(g) : g(x) \in \text{dom}(f)\}$.
- $(f \circ g)(x) = f\big(g(x)\big)$.

6.3 Classes of Functions

There are certain properties of functions that we can use to group functions together into similarly-behaving types. Let f be a function with domain and range in $\mathbb{R}$.

Definition 6.5.

- f is an **even** function $\iff$ $f(x) = f(-x)\ \forall x \in \text{dom}(f)$. This can be visualised as symmetry about the y-axis of the graph.
- f is an **odd** function $\iff$ $-f(x) = f(-x)\ \forall x \in \text{dom}(f)$. This can be visualised as $180°$ rotational symmetry about the origin of the graph.

It is possible for a function to be neither even nor odd.

Definition 6.6.

- f is an **injection** (or **one-to-one**) $\iff$

$$\forall x, x' \in \mathrm{dom}(f),\ f(x) = f(x') \implies x = x'.$$

Every element in the range is the image of *at most one* element from the domain.

- f is a **surjection** (or **onto**) $\iff$

$$\forall y \in \mathrm{range}(f),\ \exists x \in \mathrm{dom}(f) \text{ s.t. } y = f(x).$$

Every element in the range is reached by *at least one* element from the domain.

- f is a **bijection** (or **one-to-one correspondence**) $\iff$ f is **injective** and f is **surjective**. Each element of the range is mapped to by *exactly one* element from the domain.

It is possible for a function to be neither injective nor surjective.

6.4 Limits

6.5 Continuity

Very informally, a continuous function is one whose graph can be drawn without lifting the pen of the page. Intuitively this means that a small change in the input of the function leads to a small change to the output. To write a rigorous definition, we will use what we have learned from sequences.

Definition 6.7. Let $f : D \to R$, then f is **continuous at** $x_0 \in D$ if is **continuous at** $x_0 \in D$ if

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ such that } \forall x \in D,\ |x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon.$$

We say f is **continuous** if f is continuous at all $x_0 \in D$.

What this means is that for a fixed point x_0 and *any* value of ε we choose, we can find a δ such that when the inputs of the function are δ-close, the outputs of the function are ε-close. If this is true for all points in the functions domain, then we call the whole function continuous.

When we are doing a proof of continuity, we must first fix x_0, then ε, then we may choose δ similarly to the way we chose N when proving convergence of a sequence. Let's do an example.

Example 6.3. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = 2x$. Claim: f is continuous.
Let $x_0 \in \mathbb{R}$, let $\varepsilon > 0$, choose $\delta \in (0, \frac{\varepsilon}{2})$.
Consider $x \in \mathbb{R}$ with $|x - x_0| < \delta$, then

$$\begin{aligned} |f(x) - f(x_0)| &= |2x - 2x_0| \\ &= 2|x - x_0| \\ &< 2\delta \\ &< \varepsilon. \end{aligned}$$

Since the value of x_0 was arbitrary, f is continuous.

Just like when proving the convergence of a sequence, when we are first constructing a proof of continuity the value of δ is unknown until the proof is complete. Once the value is known, we go back and place it at the start so that the proof works.

It turns out that there is a very strong link between the continuous functions and convergent sequences that we will now prove.

Theorem 6.1 (Sequential Definition of Continuity) *Let $D \subseteq \mathbb{R}$, let $f : D \to \mathbb{R}$, then the following are equivalent:*

(i) *f is continuous at $x_0 \in D$.*

(ii) *If $(x_n)_n$ is a sequence in D with limit x_0, then the sequence $(f(x_n))_n$ in $\mathbb{R}$ has limit $f(x_0)$.*

Proof.
($\implies$): Suppose f is continuous at x_0, $(x_n)_n$ is a sequence in D with limit x_0.
Let $\varepsilon > 0$, since f is continuous at x_0, $\exists \delta > 0$ such that $|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$.
Now, since $x_n \to x_0$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x_0| < \delta$, therefore $|f(x_n) - f(x_0)| < \varepsilon$.
($\impliedby$): Suppose by way of contradiction that (ii) is true, but (i) is false. So $\exists \varepsilon > 0$ such that $\forall \delta > 0, \exists x \in D$ where $|x - x_0| < \delta \implies |f(x) - f(x_0)| \geq \varepsilon$ (negation of continuity).
Choose ε such that the above is satisfied, then for $n \in \mathbb{N}$, let $\delta = \frac{1}{n}$ and $x_n \in D$ such that $|x_n - x_0| < \frac{1}{n}$ but $|f(x_n) - f(x_0)| \geq \varepsilon$.
Thus $x_n \to x_0$, but $f(x_n) \not\to f(x_0)$ since ε is fixed, which is a contradiction of our assumptions that (ii) is true, hence f must be continuous at x_0. ■

What this theorem is telling us is that we now have two equivalent definitions of continuity; the $\epsilon - \delta$ definition which we defined originally and the sequential definition, "$\forall (x_n)_n \in D$ such that $x_n \to x_0$ we have $f(x_n) \to f(x_0)$". Which one we will want to use depends on the situation. Let's look at an example where using the sequential definition makes proving continuity much easier.

Example 6.4. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = 2x^2 + 1$. Claim: f is continuous.
Using $\varepsilon - \delta$ definition of continuity:
Let $x_0 \in \mathbb{R}$, let $\varepsilon > 0$, choose $\delta < \min\left\{1, \frac{\varepsilon}{4|x_0|+2}\right\}$.
Consider $x \in \mathbb{R}$ such that $|x - x_0| < \delta$, then

$$\begin{aligned}
|f(x) - f(x_0)| &= |2x^2 + 1 - (2x_0^2 + 1)| \\
&= 2|x^2 - x_0^2| \\
&= 2|(x - x_0)(x + x_0)| \\
&< 2\delta|x + x_0| \\
&< 2\delta(|x| + |x_0|) \\
&< 2\delta(2|x_0| + 1) && \text{(provided } \delta < 1 \text{ so } |x| < |x_0| + 1) \\
&< \varepsilon. && \text{(provided } \delta < \tfrac{\varepsilon}{4|x_0|+2})
\end{aligned}$$

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

$$f(x_n) = 2x_n^2 + 1 \to 2x_0^2 + 1 = f(x_0). \qquad \text{(by theorem 4.4)}$$

Therefore f is continuous.

Here are a couple of examples proving discontinuity using the sequential definition.

Example 6.5. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0. \end{cases}$ Claim: f is discontinuous at $x_0 = 0$.
Let $(x_n)_n = \left(-\frac{1}{n}\right)_n$ with $x_n \to 0$. Since $x_n < 0$, $f(x_n) = -1$. But $f(0) = 1 \neq -1$, so f cannot be continuous at $x = 0$.

Example 6.6. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$ Claim: f is continuous only at $x = 0$.

In these styles of proof, the arguments can often be greatly simplified by breaking down the given function into smaller building blocks and using previous results. Recall that this was one of the main benefits of the comparison test for proving the convergence of a series. In order to use this technique, we need to know if the continuity of functions is preserved under the operations we can do with them. Thankfully, they are.

Theorem 6.2 *Let $D \subseteq \mathbb{R}$, $x_0 \in D$, $f, g : D \to \mathbb{R}$ two functions continuous at x_0, $\lambda \in \mathbb{R}$.*

(i) $(f+g) : D \to \mathbb{R}$ *is continuous at* x_0.

(ii) $(fg) : D \to \mathbb{R}$ *is continuous at* x_0.

(iii) $(\lambda f) : D \to \mathbb{R}$ *is continuous at* x_0.

(iv) $\min\{f, g\} : D \to \mathbb{R}$ *is continuous at* x_0.

(v) *Similarly for* $\max\{f, g\}$.

(vi) $|f| : D \to \mathbb{R}$ *is continuous at* x_0.

(vii) *If* $g(x) \neq 0 \ \forall x \in D$, $\left(\frac{f}{g}\right) : D \to \mathbb{R}$ *is continuous at* x_0.

Proof.

(i) Using the $\varepsilon - \delta$ definition:
Let $\varepsilon > 0$, since f is continuous at x_0, $\exists \delta_f > 0$ so that $|x - x_0| < \delta_f \implies |f(x) - f(x_0)| < \frac{\varepsilon}{2}$. Similarly for g, $\exists \delta_g$.
Now choose $\delta = \min\{\delta_f, \delta_g\}$, then

$$\begin{aligned} |x - x_0| < \delta \implies |(f+g)(x) - (f+g)(x_0)| &= |f(x) + g(x) - (f(x_0) + g(x_0))| \\ &= |f(x) - f(x_0) + g(x) - g(x_0)| \\ &\leq |f(x) - f(x_0)| + |g(x) - g(x_0)| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Using the sequential definition:
Consider a sequence $(x_n)_n$ in D with limit x_0. Since f and g are continuous at x_0, $f(x_n) \to f(x_0)$ and $g(x_n) \to g(x_0)$.
Then $(f+g)(x_n) = f(x_n) + g(x_n) \to f(x_0) + g(x_0) = (f+g)(x_0)$.

- (ii), (iii), (vi), and (vii) follow similarly.
- To prove (iv) and (v), note that

$$\min\{f, g\} = \frac{f + g - |f - g|}{2}, \quad \max\{f, g\} = \frac{f + g + |f - g|}{2},$$

then use the results of (i), (iii), and (vi).

■

A notable consequence of this theorem is that since $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ is a continuous function (the proof is trivial), any polynomial is a continuous function.

Theorem 6.3 *Let $D_f, D_g \subseteq \mathbb{R}$, $f : D_f \to \mathbb{R}$ with $f(D_f) \subseteq (D_g)$, continuous at $x_0 \in D_f$, $g : D_g \to \mathbb{R}$ continuous at $f(x_0) \in D_g$. Then the composition $(g \circ f) : D_f \to \mathbb{R}$ is continuous at x_0.*

Proof. Let $(x_n)_n \in D_f$ with limit x_0. Since f is continuous at x_0, $f(x_n) \to f(x_0)$. Since g is continuous at $f(x_0)$,

$$(g \circ f)(x_n) = g(f(x_n)) \to g(f(x_0)) = (g \circ f)(x_0).$$

■

We will now move onto the meat of this chapter, some intuitive but nonetheless very important results about continuous functions which will be very important when we get into more advanced topics.

Definition 6.8. A function $f : D \to \mathbb{R}$ is **bounded** if $\exists M \in \mathbb{R}$ such that

$$\forall x \in D, \ |f(x)| \leq M.$$

The first theorem we will see is so simple that it is not clear why we need to prove it. It states that a continuous function on a closed interval is bounded and has a maximum and minimum value. As we will see after proving the theorem, there are discontinuous functions which satisfy these properties but there are also many which don't.

Theorem 6.4 (Extreme Value Theorem) *Given two real numbers a, b with $a < b$, let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then*

(i) *f is bounded.*

(ii) *f attains a maximum, i.e. $\exists x_0 \in [a, b]$ such that $f(x) \leq f(x_0)$ $\forall x \in [a, b]$.*

(iii) *f attains a minimum, i.e. $\exists x_0 \in [a, b]$ such that $f(x) \geq f(x_0)$ $\forall x \in [a, b]$.*

Proof.

(i) Assume by way of contradiction that f is not bounded. Then let $(x_n)_n$ be a sequence in $[a, b]$ such that $\forall n \in \mathbb{N}$ we have $|f(x_n)| > n$.
$(x_n)_n$ is a bounded sequence (all the terms are contained within an interval), so by the Bolzano-Weierstrass theorem (4.8), there exists a convergent subsequence $(x_{n_k})_k \in [a, b]$ with limit $x_0 \in [a, b]$.

Then since f is continuous, $f(x_{n_k}) \to f(x_0)$, hence $(f(x_{n_k}))_k$ is a bounded sequence by theorem 4.3, but this is a contradiction since $|f(x_{n_k})| > n_k \; \forall k$, so f must be bounded.

(ii) Let $M = \sup\{f(x) : \; x \in [a,b]\}$. M exists since all bounded sets in $\mathbb{R}$ have suprema and we know the set is bounded by (i).
By definition of sup, $\forall n \in \mathbb{N} \; \exists x_n \in [a,b]$ such that $M - \frac{1}{n} \leq f(x_n) \leq M$, which implies that $f(x_n) \to M$ as $n \to \infty$ (however, $(x_n)_n$ is by no means necessarily convergent, so we are not done yet).
By the Bolzano-Weierstrass theorem (4.8), there exists a convergent subequence $(x_{n_k})_k \in [a,b]$ with limit $x_0 \in [a,b]$, and since f is continuous, $f(x_{n_k}) \to f(x_0)$.
Therefore since $(x_{n_k})_k$ is a subsequence of $(x_n)_n$, $f(x_0) = M$.

The proof of (iii) follows similarly from (ii). ■

The next theorem is the most important result for continuous functions. It cements the intuitive notion that continuous functions have no gaps or jumps. The proof uses the completeness property of the real numbers, which ensures that there are no infinitesimal gaps in the function.

Theorem 6.5 (Intermediate Value Theorem) *Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ a continuous function. Let $a, b \in I$ with $a < b$ and let $y \in \mathbb{R}$ lie between $f(a)$ and $f(b)$ (i.e. $f(a) < y < f(b)$ or $f(b) < y < f(a)$). Then $\exists x \in (a,b)$ such that*

$$f(x) = y.$$

Proof. Assume without meaningful loss of generality that $f(a) < y < f(b)$. Let $S = \{x \in [a,b] : \; f(x) < y\}$ (since $a \in S$, $S \neq \emptyset$). Let $x_0 = \sup S$.
Claim: $f(x_0) = y$.

(i) $\forall n \in \mathbb{N} \; x_0 - \frac{1}{n} < x_0$, so $x_0 - \frac{1}{n}$ is not an upper bound for S. Hence $\exists s_n \in S$ such that $x_0 - \frac{1}{n} < s_n < x_0$. Then $s_n \to x_0$ as $n \to \infty$ and therefore since f is continuous $f(s_n) \to f(x_0)$.
Since $s_n \in S$, $f(s_n) < y$, therefore $f(x_0) \leq y$.

(ii) It can be shown similarly that $f(x_0) \geq y$, so $f(x_0) = y$.

Note that since $f(a) < f(x_0) < f(b)$, $x_0 \in (a,b)$. ■

It should be noted that although it may seem as if this theorem may be taken to be the definition of a continuous function, the converse of the theorem is not true.

Example 6.7. Let $f : [0,1] \to [0,1]$ be continuous. Then f has a fixed point ($\exists x \in [0,1]$ such that $f(x) = x$).

Proof. Let g be given by $g(x) = f(x) - x$. Note that by theorem 6.2, g is continuous on $[0,1]$. Also note that

$$\begin{aligned} g(0) &= f(0) - 0 = f(0) \geq 0 \\ g(1) &= f(1) - 1 \leq 1 - 1 = 0 \end{aligned}$$

So $0 \in [g(1), g(0)]$. If $g(0) = 0$ or $g(1) = 0$, then 0 or 1 are fixed points. Therefore, assume that $g(0) > 0 > g(1)$, and hence by the intermediate value theorem (6.5), $\exists x_0 \in (0,1)$ so that $g(x_0) = 0$, i.e. $f(x_0) - x_0 = 0$, so x_0 is a fixed point.

Corollary 6.5.1 *Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ a continuous function. Then the range of f, $f(I)$, is either an interval or a single point.*

Proof. Suppose there are two or more distinct points in $f(I)$, let $y \in (\inf f(I), \sup f(I))$. By the intermediate value theorem (6.5), $\exists x \in I$ such that $f(x) = y$, hence $f(I)$ is an interval. ■

Chapter 7

Differentiation

7.1 The Derivative

The derivative is an extremely important concept in applied mathematics, as it allows us to model real life situations with mathematical functions and their rates of change. For this reason, the derivative is the most interesting part of analysis for applied mathematicians and physicists. The purpose of this chapter is to establish what a derivative is, what its properties are, and what we can do with derivatives.

You probably have an intuition that the derivative of a function is the gradient of the tangent line at every point along the function's domain, and that in order for a function to have a derivative defined everywhere it must have no sharp corners or cusps. This is the basic idea, but of course in analysis we must be extra careful with our definitions. How do we know if we can even define a tangent line? To make a definition of the derivative that we can work with we will use equipment that we have developed in the last chapter.

Definition 7.1. Let $I \subseteq \mathbb{R}$ be an open interval and let $f : I \to \mathbb{R}$ be a function. Let $c \in I$. Then f is **differentiable** at c if

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \text{ exists.}$$

If this limit exists, then we denote it $f'(c)$, the **derivative** of f at c. We say f is **differentiable** if f is differentiable at all $c \in I$. Then $f' : I \to \mathbb{R}$ is also a function, called the **derivative** of f (also denoted $\frac{\mathrm{d}f}{\mathrm{d}x}$).

Example 7.1. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Claim: f is differentiable at $x = 3$ with $f'(3) = 6$.

Proof: Let $c = 3, x \neq 3$. Then

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

Hence,

$$f'(3) = \lim_{x \to 3} \frac{f(x) - f(3)}{x - 3} = \lim_{x \to 3} x + 3 = 6.$$

So f is differentiable at $x = 3$ with derivative 6. In fact, if $c \neq x$,

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

So f is differentiable everywhere with derivative $f' : \mathbb{R} \to \mathbb{R}$ given by $f'(x) = 2x$.

Example 7.2. Let $f : (0, \infty) \to \mathbb{R}$ be given by $f(x) = \sqrt{x}$. Claim: f is differentiable.
Proof: Let $c \neq x$, then

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

Hence $f' : (0, \infty) \to \mathbb{R}$ is given by

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

Example 7.3. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^n$ for some $n \in \mathbb{R}$. Claim: f is differentiable with derivative $f' : \mathbb{R} \to \mathbb{R}$ given by $f(x) = nx^{n-1}$.
Proof: Let $c \neq x$, then

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

Hence $f'(c)$ is given by

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

Now that we have defined strictly what a derivative is, we will explore some properties of them. Firstly, we will note that our notion of differentiability is **stronger** than the definition of continuity.

Theorem 7.1 *Let $I \subseteq \mathbb{R}$ be an open interval, let $f : I \to \mathbb{R}$ be a function and let $c \in I$. Then if f is differentiable at c, then f is continuous at c.*

Proof. Since f is differentiable, $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$. Hence

$$\begin{aligned}\lim_{x\to c}(f(x)-f(c)) &= \lim_{x\to c}\frac{f(x)-f(c)}{x-c}(x-c)\\ &= \left(\lim_{x\to c}\frac{f(x)-f(c)}{x-c}\right)\left(\lim_{x-c}(x-c)\right) &&\text{(By theorem ??)}\\ &= f'(c)\cdot 0 = 0.\end{aligned}$$

Therefore,

$$\lim_{x\to c} f(x) = f(c).$$

So by theorem **??**, f is continuous. ■

Note that the converse of this theorem is not true. Continuity does not imply differentiability (although).

Example 7.4. Let $f : (-1, 1) \to \mathbb{R}$ be given by $f(x) = |x|$. Claims:

(i) f is continuous.

(ii) f is not differentiable at $x = 0$.

Proof:

(i) This is true by theorem 6.2.

(ii) Consider $x \neq 0$, then

$$\frac{f(x)-f(0)}{x-0} = \frac{|x|}{x} = \begin{cases} 1 & x > 0 \\ -1 & x < 0. \end{cases}$$

Thus $f'(0)$ is not defined since the limit does not exist at $x = 0$, so f is not differentiable at $x = 0$.

Now we will prove some properties of derivatives under arithmetic operations similarly to how we have done with every concept we have covered so far. These will be familiar are they are the same rules for derivatives that you will have learned in school.

Theorem 7.2 *Let $I \subseteq \mathbb{R}$, let $f, g : I \to \mathbb{R}$ be two functions, differentiable at $c \in I$.*

(i) *$(f + g) : I \to \mathbb{R}$ is differentiable at c and $(f + g)'(c) = f'(c) + g'(c)$.*

(ii) *Given $\lambda \in \mathbb{R}$, $(\lambda f) : I \to \mathbb{R}$ is differentiable at c and $(\lambda f)'(c) = \lambda f'(c)$.*

(iii) $(fg) : I \to \mathbb{R}$ *is differentiable at* c *and* $(fg)'(c) = f(c)g'(c) + f'(c)g(c)$.

(iv) *If* $g(c) \neq 0$, *then* $\left(\frac{1}{g}\right) : I \to \mathbb{R}$ *is differentiable at* c *and* $\left(\frac{1}{g}\right)'(c) = \frac{-g'(c)}{g(c)^2}$.

Proof. In all cases below, let $x \in I$ with $x \neq c$.

(i)

$$\begin{aligned}
(f+g)'(c) &= \lim_{x\to c} \frac{(f+g)(x) - (f+g)(c)}{x-c} \\
&= \lim_{x\to c} \frac{f(x) + g(x) - f(c) - g(c)}{x-c} \\
&= \lim_{x\to c} \frac{f(x) - f(c)}{x-c} + \lim_{x\to c} \frac{g(x) - g(c)}{x-c} \qquad \text{(By theorem \textbf{??})} \\
&= f'(c) + g'(c).
\end{aligned}$$

(ii)

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

(iii) Note that since f and g are differentiable at c, they are continuous at c by theorem 7.1. So $\lim_{x\to c} f(x) = f(c)$.

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

(iv) Since $g(c) \neq 0$ and g is continuous at c, there exists $\delta > 0$ so that $x \in (c-\delta, c+\delta) \implies g(x) \neq 0$.

$$\begin{aligned}
\left(\frac{1}{g}\right)'(c) = \lim_{x\to c} \frac{\left(\frac{1}{g}\right)(x) - \left(\frac{1}{g}\right)(c)}{x-c} &= \lim_{x\to c} \left(\frac{1}{x-c}\right)\left(\frac{1}{g(x)} - \frac{1}{g(c)}\right) \\
&= \lim_{x\to c} \left(\frac{1}{x-c}\right)\left(\frac{g(c) - g(x)}{g(x)g(c)}\right) \\
&= \lim_{x\to c} \left(\frac{-1}{g(x)g(c)}\right)\frac{g(x) - g(c)}{x-c} \\
&= \frac{-g'(c)}{g(c)^2}.
\end{aligned}$$

■

Part (iii) of this theorem is the well-known product rule for derivatives. We can prove the quotient rule as a corollary.

Corollary 7.2.1 (Quotient Rule) *If $g(c) \neq 0$, $\left(\frac{f}{g}\right) : I \to \mathbb{R}$ is differentiable at c and $\left(\frac{f}{g}\right)'(c) = \frac{g(c)f'(c)-g'(c)f(c)}{g(c)^2}$.*

Proof. Since $\frac{f}{g} = f \cdot \frac{1}{g}$,

$$\begin{aligned}
\left(f \cdot \frac{1}{g}\right)'(c) &= \frac{f'(c)}{g(c)} + f(c)\left(\frac{1}{g}\right)'(c) && \text{(By theorem 7.2 (iii))} \\
&= \frac{f'(c)}{g(c)} - \frac{f(c)g'(c)}{g(c)^2} && \text{(By theorem 7.2 (iv))} \\
&= \frac{g(c)f'(c) - g'(c)f(c)}{g(c)^2}.
\end{aligned}$$

∎

Other corollaries of this theorem are that all polynomials are differentiable, and all rational functions are differentiable except where the denominator is zero.

Theorem 7.3 (Chain Rule) *Let $I, J \subseteq \mathbb{R}$ be intervals, $f : I \to \mathbb{R}$ a function differentiable at some $c \in I$ with $f(I) \subseteq J$, $g : J \to \mathbb{R}$ a function differentiable at $f(c) \in J$. Then $(g \circ f) : I \to \mathbb{R}$ is differentiable at c and $(g \circ f)'(c) = g'(f(c))f'(c)$.*

7.2 Important Theorems about Derivatives

The following theorems about the properties of differential functions are among the most important results in real analysis for applied mathematics and physics. The first theorem we will look at is Rolle's theorem, which essentially states that if a function is differentiable and has the same value at two points, then the derivative must be zero somewhere between those two points.

Theorem 7.4 (Rolle's Theorem) *Let $a, b \in \mathbb{R}$ with $a < b$, let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b) with $f(a) = f(b)$. Then there exists $\theta \in (a, b)$ such that $f'(\theta) = 0$.*

Proof. Firstly, we will prove that the derivative of a function is zero at the maximum or minimum points of a function. This is a fact that we are used to from school, and we will need it when proving this theorem.

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

(i) *If f attains its maximum at some $c \in I$, then $f'(c) = 0$.*

(ii) *If f attains its minimum at some $c \in I$, then $f'(c) = 0$.*

Proof.

(i) Since f has reached its maximum, $f(x) \leq f(c)\ \forall x \in I$. Hence

$$f(x) - f(c) \leq 0\ \forall x \in I.$$

Now, if $x < c$, then $x - c < 0$ and

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

On the other hand, if $x > c$, then $x - c > 0$ and

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

Since f is differentiable, $f'(c)$ is defined and so the left and right limits at c must be equal, which can only be true if

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

(ii) follows similarly.

■

Now there are three cases to consider. The first is the case where the function is a horizontal line between a and b, the second is the case where the function has a maximum between a and b, and the third is the case where there is a minimum.

(i) Suppose $\forall x \in [a, b]$, $f(x) = f(a)$. Then $\forall \theta \in (a, b)$, $f'(\theta) = 0$.

(ii) Suppose there exists some $y \in [a, b]$ with $f(y) > f(a) = f(b)$. Since f is continuous on $[a, b]$, the Extreme Value Theorem (6.4) says that f attains a maximum at some $\theta \in [a, b]$, therefore

$$f(x) \leq f(\theta)\ \forall x \in [a, b].$$

However since $f(y) > f(a) = f(b)$, θ must be at least y, i.e. $\theta \neq a$, $\theta \neq b$. So $\theta \in (a, b)$ and hence by the part (i)) of the lemma above, $f'(\theta) = 0$.

(iii) Suppose there exists some $y \in [a, b]$ with $f(y) < f(a) = f(b)$. Then this case follows similarly to (ii) in that f must attain a minimum in (a, b).

■

The next theorem is a generalisation of Rolle's theorem. It states that between two points of a function, there must be a least one point where the tangent to the function is parallel to a straight line between the two points.

Theorem 7.5 (Mean Value Theorem) *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b). Then $\exists \theta \in (a, b)$ such that*

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

Proof. This theorem is very easy to prove using Rolle's theorem (7.4). Let $g(x) = f(x) - \lambda x$, where λ is chosen so that $g(a) = g(b)$. So

$$\begin{aligned} f(a) - \lambda a &= f(b) - \lambda b \\ \lambda(b - a) &= f(b) - f(a) \\ \lambda &= \frac{f(b) - f(a)}{b - a}. \end{aligned}$$

Then by Rolle's theorem, there exists $\theta \in (a, b)$ such that $g'(\theta) = 0$. By theorem 7.2,

$$\begin{aligned} & g'(\theta) = f'(\theta) - \lambda = 0 \\ \implies & f'(\theta) = \lambda = \frac{f(b) - f(a)}{b - a}. \end{aligned}$$

■

Another way of writing the Mean Value theorem is that there exists $\theta \in (a, b)$ so that $f(b) = f(a) + (b - a)f'(\theta)$. This is another way of expressing what it means for a function to be differentiable. If we replace b with some point x in the interior of the domain of f then we get

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

This is a linear function, so we say that if a function is differentiable at a point then it has a good linear approximation at that point. This fact could be used to calculate the value of the function at lots of points if we know the value of the function at a few points and we know information about the derivative. We will touch on this in the last part of this chapter.

Example 7.5. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function with $f(0) = 0, f(1) = 1, f(2) = 2$. Show that $f'(x)$ takes the values 0 and $\frac{1}{2}$ for some $x \in (0, 2)$.
Proof: Note that $f(1) = f(2)$, so by Rolle's theorem there exists some $\theta \in (1, 2) \subseteq (0, 2)$ with $f'(\theta) = 0$. Also, by the Mean Value theorem there exists $\theta \in (0, 2)$ such that

$$f'(\theta) = \frac{f(2) - f(0)}{2 - 0} = \frac{1 - 0}{2 - 0} = \frac{1}{2}.$$

7.3 Higher Derivatives

If the derivative of a differentiable function can be considered a function in its own right, then we can investigate the differentiability of it as well. If the derivative f' of a function f is differentiable, then we can define another function f'' or $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$ as the **second derivative**, and we say that f is **twice differentiable**. We can carry on this way if the derivatives carry on being differentiable, and once the prime notation becomes unwieldy, we can write the n^{th} **derivative** of f as $f^{(n)}$ or $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$.

Theorem 7.6 *Let $I \subseteq \mathbb{R}$ and let $f, g : I \to \mathbb{R}$ be two functions which are twice differentiable at some point $c \in I$.*

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

(ii) *Given* $\lambda \in \mathbb{R}$, $(\lambda f)''(c) = \lambda f''(c)$.

(iii) $(fg)''(c) = f''(c)g(c) + 2f'(c)g'(c) + f(c)g''(c)$.

Proof.

(i) By theorem 7.2 (i), $(f + g)'(c) = f'(c) + g'(c)$. Therefore,

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

(ii) By theorem 7.2 (ii), $(\lambda f)'(c) = \lambda f'(c)$. So,

$$(\lambda f)''(c) = (\lambda f')'(c) = \lambda f''(c).$$

(iii) By theorem 7.2 (iii), $(fg)'(c) = f(c)g'(c) + f'(c)g(c)$. So,

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

∎

The first two parts of this theorem generalise very easily by induction to all higher derivatives. The final part is more involved.

Theorem 7.7 (General Leibniz Rule) *Let $I \subseteq \mathbb{R}$ and let $f, g : I \to \mathbb{R}$ be two n-times differentiable functions at some point $c \in I$. Then the n^{th} derivative of the product $(fg) : I \to \mathbb{R}$ at c takes the value*

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

The last theorem of this chapter is a generalisation of the Mean Value theorem. It allows us to approximate the value of a function at a point as a polynomial with the derivatives of the function as coefficients.

Theorem 7.8 (Taylor's Theorem) *Let I be some open interval and let $a, b \in I$. Let $f : I \to \mathbb{R}$ be a continuous function which is $(n+1)$-times differentiable at a. Then there exists θ between a and b such that*

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

Proof. First we define $F : [a, b] \to \mathbb{R}$ to be given by

$$F(t) = f(b) - f(t) - (b-t)f'(t) - \frac{(b-t)^2}{2!} f''(t) - \cdots - \frac{(b-t)^n}{n!} f^{(n)}(t).$$

Since the first n derivatives of f are continuous on $[a, b]$ and differentiable on (a, b), F inherits these properties by theorem 7.2. The derivative of F is then

$$\begin{aligned} F'(t) &= -f'(t) + [f'(t) - (b-t)f''(t)] + \left[(b-t)f''(t) - \frac{(b-t)^2}{2!} f'''(t)\right] \\ &\quad + \cdots + \left[\frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) - \frac{(b-t)^n}{n!} f^{(n+1)}(t)\right] \\ &= -\frac{(b-t)^n}{n!} f^{(n+1)}(t). \end{aligned}$$

Now, define $G : [a, b] \to \mathbb{R}$ to be given by

$$G(t) = F(t) - \left(\frac{b-t}{b-a}\right)^{n+1} F(a).$$

Notice that $G(a) = G(b) = 0$, so by Rolle's theorem (7.4) there exists $\theta \in (a, b)$

such that $G'(\theta) = 0$. Specifically,

$$\begin{aligned} G'(\theta) &= F'(\theta) + (n+1)\frac{(b-\theta)^n}{(b-a)^{n+1}}F(a) \\ &= -\frac{(b-\theta)^n}{n!}f^{(n+1)}(\theta) + (n+1)\frac{(b-\theta)^n}{(b-a)^{n+1}}F(a) \\ &= (n+1)\frac{(b-\theta)^n}{(b-a)^{n+1}}\left[F(a) - \frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(\theta)\right] = 0. \end{aligned}$$

The prefactor is never zero, hence the term in square brackets must be zero, so

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

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

■

Taylor's theorem allows us to make approximations that are much better than linear ones. If, like before, we replace b any point x in the interior of the domain of f, we get

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

where θ is now between x and a. This polynomial representation of f is known as the n^{th} **Taylor Polynomial** of f about a and the last term is known as the error or remainder, sometimes denoted $R_n(x)$. The idea is that for well-behaved functions, the remainder will get smaller as n gets larger.

Corollary 7.8.1 (Maclaurin's Theorem) *Suppose $0 \in I$. Let $x \in I$, then if $f : I \to \mathbb{R}$ is an $(n+1)$-times differentiable function at 0, then there exists θ between x and 0 such that*

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

where $R_n(x) = \frac{x^{n+1}}{(n+1)!}f^{(n+1)}(\theta)$.

Proof. This is easily shown by using Taylor's theorem with $a = 0$. ■

Taylor's theorem and Maclaurin's theorem have very important consequences that we will explore in more detail in later chapters.

Bibliography

[1] Jonathan Fraser. Lecture notes for MT2502 Analysis. University of St. Andrews, 2020.

[2] John M. Howie. *Real Analysis*. Springer London Ltd., 2001.

[3] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Book Company, 1953.

[4] Terence Tao. *Analysis I*. Hindustan Book Agency, 2016.

[5] Terence Tao. *Analysis II*. Hindustan Book Agency, 2016.