

# Analysis

Cambridge University Mathematical Tripos: Part IA

17th May 2024

## Contents

<b>1</b>	<b>Limits and convergence</b>	<b>4</b>
1.1	Definition of limit . . . . .	4
1.2	Fundamental axiom of the real numbers . . . . .	4
1.3	Properties of limits . . . . .	4
1.4	Harmonic series . . . . .	5
1.5	Limits in the complex plane . . . . .	5
1.6	The Bolzano–Weierstrass theorem . . . . .	5
1.7	Cauchy sequences . . . . .	6
<b>2</b>	<b>Series</b>	<b>7</b>
2.1	Definition . . . . .	7
2.2	Geometric series . . . . .	8
<b>3</b>	<b>Convergence tests</b>	<b>9</b>
3.1	Comparison test . . . . .	9
3.2	Cauchy’s root test . . . . .	9
3.3	D’Alembert’s ratio test . . . . .	10
3.4	Cauchy’s condensation test . . . . .	10
3.5	Alternating series . . . . .	12
<b>4</b>	<b>Absolute convergence</b>	<b>12</b>
4.1	Absolute convergence . . . . .	12
4.2	Conditional convergence and rearrangement . . . . .	13
<b>5</b>	<b>Continuity</b>	<b>14</b>
5.1	Definitions . . . . .	14
5.2	Making continuous functions . . . . .	14
5.3	Composition of continuous functions . . . . .	15
<b>6</b>	<b>Limit of a function</b>	<b>15</b>
6.1	Definition . . . . .	15
6.2	Properties . . . . .	16
6.3	Intermediate value theorem . . . . .	16
6.4	Bounds of a continuous function . . . . .	17
6.5	Inverse functions . . . . .	18

<b>7</b>	<b>Differentiability</b>	<b>18</b>
7.1	Definitions . . . . .	18
7.2	Differentiation of sums and products . . . . .	19
7.3	Differentiating polynomial terms . . . . .	20
7.4	Chain rule . . . . .	21
7.5	Rolle's theorem . . . . .	21
7.6	Mean value theorem . . . . .	22
7.7	Properties of a function from its derivative . . . . .	22
7.8	Inverse function theorem . . . . .	23
7.9	Derivative of rational powers . . . . .	23
7.10	Mean value theorem applied to limits . . . . .	23
7.11	Cauchy's mean value theorem . . . . .	24
<b>8</b>	<b>Taylor's theorem</b>	<b>24</b>
8.1	Lagrange's and Cauchy's remainders . . . . .	24
8.2	Bounding error terms . . . . .	26
8.3	Binomial series . . . . .	26
<b>9</b>	<b>Power series</b>	<b>28</b>
9.1	Complex differentiation . . . . .	28
9.2	Definition of power series . . . . .	28
9.3	Radius of convergence . . . . .	28
9.4	Infinite differentiability . . . . .	31
9.5	Defining standard functions . . . . .	34
9.6	Exponential and logarithmic functions . . . . .	34
9.7	Real numbered exponents . . . . .	35
9.8	Trigonometric functions . . . . .	36
9.9	Circle constants . . . . .	36
<b>10</b>	<b>Integration</b>	<b>38</b>
10.1	Geometry of trigonometric functions . . . . .	38
10.2	Hyperbolic functions . . . . .	38
10.3	Defining the Riemann integral . . . . .	38
10.4	Determining integrability . . . . .	40
10.5	Monotonic and continuous functions . . . . .	40
10.6	Complicated integrable functions . . . . .	42
10.7	Properties of Riemann integral . . . . .	43
<b>11</b>	<b>Fundamental theorem of calculus</b>	<b>44</b>
11.1	Breaking an interval . . . . .	44
11.2	Fundamental theorem of calculus . . . . .	45
<b>12</b>	<b>Integration techniques</b>	<b>47</b>
12.1	Integration by parts . . . . .	47
12.2	Integration by substitution . . . . .	47
<b>13</b>	<b>Integrals in Taylor's theorem</b>	<b>48</b>
13.1	Integral remainder form of Taylor's theorem . . . . .	48
13.2	Mean value theorem for integrals . . . . .	49
13.3	Deriving Lagrange's and Cauchy's remainders for Taylor's theorem . . . . .	49

<b>14</b>	<b>Uses of integration</b>	<b>50</b>
14.1	Improper integration . . . . .	50
14.2	Integral test for series convergence . . . . .	52
14.3	Piecewise continuous functions . . . . .	53

# 1 Limits and convergence

## 1.1 Definition of limit

**Definition.** We say that the sequence  $a_n \rightarrow a$  as  $n \rightarrow \infty$  if given  $\varepsilon > 0$ ,  $\exists N$  such that  $|a_n - a| < \varepsilon$  for all  $n \geq N$ . Note that this  $N$  is actually a function of  $\varepsilon$ ; we may need to choose a very large  $N$  if the  $\varepsilon$  provided is very small, for instance.

**Definition.** An increasing sequence is a sequence for which  $a_n \leq a_{n+1}$ , and a decreasing sequence is a sequence for which  $a_n \geq a_{n+1}$ . Such increasing and decreasing sequences are called monotone. A strictly increasing sequence or a strictly decreasing sequence simply strengthens the inequalities to not include the equality case.

## 1.2 Fundamental axiom of the real numbers

If we have some increasing sequence  $a_n \in \mathbb{R}$ , where  $\exists A \in \mathbb{R}$  such that  $\forall n \geq 1, a_n \leq A$ , then  $\exists a \in \mathbb{R}$  such that  $a_n \rightarrow a$  as  $n \rightarrow \infty$ . This is also known as the 'least upper bound' axiom or property. This axiom applies equivalently to decreasing sequences of real numbers bounded below. We can also rephrase the axiom to state that every non-empty set of real numbers that is bounded above has a supremum.

**Definition.** We say that the supremum  $\sup S$  of a non-empty, bounded above set  $S$  is  $K$  if

- (i)  $x \leq K$  for all  $x \in S$
- (ii) given  $\varepsilon > 0$ ,  $\exists x \in S$  such that  $x > K - \varepsilon$

Note that the supremum (and hence the infimum) is unique.

## 1.3 Properties of limits

**Lemma.** The following properties about real sequences hold.

- (i) The limit is unique. That is, if  $a_n \rightarrow a$  and  $a_n \rightarrow b$ , then  $a = b$ .
- (ii) If  $a_n \rightarrow a$  as  $n \rightarrow \infty$  and  $n_1 < n_2 < \dots$ , then  $a_{n_j} \rightarrow a$  as  $j \rightarrow \infty$ . In other words, subsequences converge to the same limit.
- (iii) If  $a_n = c$  for all  $n$ , then  $a_n \rightarrow c$  as  $n \rightarrow \infty$ .
- (iv) If  $a_n \rightarrow a$  and  $b_n \rightarrow b$ , then  $a_n + b_n \rightarrow a + b$ .
- (v) If  $a_n \rightarrow a$  and  $b_n \rightarrow b$ , then  $a_n b_n \rightarrow ab$ .
- (vi) If  $a_n \rightarrow a$ ,  $a_n \neq 0$  for all  $n$ , and  $a \neq 0$ , then  $\frac{1}{a_n} \rightarrow \frac{1}{a}$ .
- (vii) If  $a_n \rightarrow a$ , and  $a_n \leq A$  for all  $n$ , then  $a \leq A$ .

*Proof.* We prove the some of these statements here.

- (i) Given  $\varepsilon > 0$ ,  $\exists n_1$  such that  $|a_n - a| < \varepsilon$  for all  $n \geq n_1$ , and  $\exists n_2$  such that  $|a_n - b| < \varepsilon$  for all  $n \geq n_2$ . So let  $N = \max(n_1, n_2)$ , so both inequalities hold. Then for all  $n \geq N$ , using the triangle inequality,  $|a - b| \leq |a_n - a| + |a_n - b| < 2\varepsilon$ . So  $a = b$ .

(ii) Given  $\varepsilon > 0$ ,  $\exists N$  such that  $|a_n - a| < \varepsilon$  for all  $n \geq N$ . Since  $n_j \geq j$  (by induction),  $|a_{n_j} - a| < \varepsilon$  for all  $j \geq N$ .

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

If  $a_n \rightarrow a$ , then given  $\varepsilon > 0$ ,  $\exists N_1$  such that  $|a_n - a| < \varepsilon$  for all  $n \geq N_1$ . (\*)

If  $b_n \rightarrow b$ , then given  $\varepsilon > 0$ ,  $\exists N_2$  such that  $|b_n - b| < \varepsilon$  for all  $n \geq N_2$ .

Using (\*), if  $n \geq N_1(1)$  (i.e.  $\varepsilon = 1$ ),  $|a_n - a| < 1$ , so  $|a_n| \leq |a| + 1$ .

Therefore  $|a_n b_n - ab| \leq \varepsilon(|a| + 1 + |b|)$  for all  $n \geq N_3(\varepsilon) = \max\{N_1(1), N_1(\varepsilon), N_2(\varepsilon)\}$ .

□

## 1.4 Harmonic series

**Lemma.** The sequence  $\frac{1}{n}$  tends to zero as  $n \rightarrow \infty$ .

*Proof.* We know that  $\frac{1}{n}$  is a decreasing sequence, and it is bounded below by zero. Hence it converges to a limit  $a$ . We will prove now that  $a = 0$ .  $\frac{1}{2n} = \frac{1}{2} \cdot \frac{1}{n}$ , and by property (v) above,  $\frac{1}{2n}$  tends to  $\frac{1}{2} \cdot a$ . But  $\frac{1}{2n}$  is a subsequence of  $\frac{1}{n}$ , and so by property (ii) it converges to  $a$ . So by property (i),  $\frac{1}{2} \cdot a = a$  hence  $a = 0$ . □

## 1.5 Limits in the complex plane

*Remark.* The definition of the limit of a sequence makes perfect sense for  $a_n \in \mathbb{C}$ .

**Definition.**  $a_n \rightarrow a$  if given  $\varepsilon > 0$ ,  $\exists N$  such that  $\forall n \geq N$ ,  $|a_n - a| < \varepsilon$ .

From this definition, it is easy to check that properties (i)–(vi) hold for complex numbers.

However, property (vii) makes no sense in the world of the complex numbers since they do not have an ordering.

## 1.6 The Bolzano–Weierstrass theorem

**Theorem.** If  $x_n$  is a sequence of real numbers, and there exists some  $k$  such that  $|x_n| \leq k$  for all  $n$ , then we can find  $n_1 < n_2 < n_3 < n_4 < \dots$  and  $x \in \mathbb{R}$  such that  $x_{n_j} \rightarrow x$  as  $j \rightarrow \infty$ . In other words, any bounded sequence has a convergent subsequence.

*Remark.* This theorem does not state anything about the uniqueness of such a subsequence; indeed, there could exist many subsequences that have possibly different limits. For example,  $x_n = (-1)^n$  gives  $x_{2n+1} \rightarrow -1$  and  $x_{2n} \rightarrow 1$ .

*Proof.* Let  $[a_1, b_1]$  be the range of the sequence, i.e.  $[-k, k]$ . Then let the midpoint  $c_1 = \frac{a_1 + b_1}{2}$ . Consider the following alternatives:

- (i)  $x_n \in [a_1, c]$  for infinitely many values of  $n$ .
- (ii)  $x_n \in [c, b_1]$  for infinitely many values of  $n$ .

Note that cases 1 and 2 could hold at the same time. If case 1 holds, we set  $a_2 = a_1$  and  $b_2 = c$ . If case 1 fails, then case 2 must hold, so we can set  $a_2 = c$  and  $b_2 = b_1$ . We have now constructed a subsequence whose range is half as large as the original sequence, and it contains infinitely many values of  $x_n$ .

We can proceed inductively to construct sequences  $a_n, b_n$  such that  $x_m \in [a_n, b_n]$  for infinitely many values of  $m$ . This is known as a ‘bisection method’. By construction,  $a_{n-1} \leq a_n \leq b_n \leq b_{n-1}$ . Since we are dividing by two each time,

$$b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1}) \quad (*)$$

Note that  $a_n$  is a bounded, increasing sequence; and  $b_n$  is a bounded, decreasing sequence. By the Fundamental Axiom of the Real Numbers,  $a_n$  and  $b_n$  converge to limits  $a \in [a_1, b_1]$  and  $b \in [a_1, b_1]$ . Using (\*),  $b - a = \frac{b-a}{2} \implies b = a$ .

Since  $x_m \in [a_n, b_n]$  for infinitely many values of  $m$ , having chosen  $n_j$  such that  $x_{n_j} \in [a_j, b_j]$ , there is  $n_{j+1} > n_j$  such that  $x_{n_{j+1}} \in [a_{j+1}, b_{j+1}]$ . Informally, this works because we have an unlimited supply of such  $x$  values. Hence

$$a_j \leq x_{n_j} \leq b_j$$

So this  $x_{n_j} \rightarrow a$ , so we have constructed a convergent subsequence. □

## 1.7 Cauchy sequences

**Definition.** A sequence  $a_n$  is called a Cauchy sequence if given  $\varepsilon > 0$  there exists  $N > 0$  such that  $|a_n - a_m| < \varepsilon$  for all  $n, m \geq N$ . Informally, the terms of the sequence grow ever closer together such that there are infinitely many consecutive terms within a small region.

**Lemma.** If a sequence converges, it is a Cauchy sequence.

*Proof.* If  $a_n \rightarrow a$ , given  $\varepsilon > 0$  then  $\exists N$  such that  $\forall n \geq N, |a_n - a| < \varepsilon$ . Then take  $m, n \geq N$ , and we have

$$|a_n - a_m| \leq |a_n - a| + |a_m - a| < 2\varepsilon$$

□

**Theorem.** Every Cauchy sequence converges.

*Proof.* First, we note that if  $a_n$  is a Cauchy sequence then it is bounded. Let us take  $\varepsilon = 1$ , so  $N = N(1)$  in the Cauchy property. Then

$$|a_n - a_m| < 1$$

for all  $m, n \geq N(1)$ . So by the triangle inequality,

$$|a_m| \leq |a_m - a_N| + |a_N| < 1 + |a_N|$$

So the sequence after this point is bounded by  $1 + |a_N|$ . The remaining terms in the sequence are only finitely many, so we can compute the maximum of all of those terms along with  $1 + |a_N|$  to produce a bound  $k$  for all  $n$ .

By the Bolzano–Weierstrass Theorem, this sequence  $a_n$  has a convergent subsequence  $a_{n_j} \rightarrow a$ . We want to prove that  $a_n \rightarrow a$ . Given  $\varepsilon > 0$ , there exists  $j_0$  such that  $|a_{n_j} - a| < \varepsilon$  for all  $j \geq j_0$ . Also,  $\exists N(\varepsilon)$  such that  $|a_m - a_n| < \varepsilon$  for all  $m, n \geq N(\varepsilon)$ . Combining these, we can take a  $j$  such that  $n_j \geq \max\{N(\varepsilon), n_{j_0}\}$ . Then, if  $n \geq N(\varepsilon)$ , using the triangle inequality,

$$|a_n - a| \leq |a_n - a_{n_j}| + |a_{n_j} - a| < 2\varepsilon$$

□

Therefore, on  $\mathbb{R}$ , a sequence is convergent if and only if it is a Cauchy sequence. This is sometimes referred to as the general principle of convergence, however this is a relatively old-fashioned name. This property is very useful, since we don't need to know what the limit actually is.

## 2 Series

### 2.1 Definition

Let  $a_n$  be a real or complex sequence. We say that  $\sum_{j=1}^{\infty} a_j$  converges to  $s$  if the sequence of partial sums  $s_N$  converges to  $s$  as  $N \rightarrow \infty$ , i.e.

$$s_N = \sum_{j=1}^N a_j \rightarrow s$$

If the sequence of partial sums does not converge, then we say that the series diverges. Note that any problem on series can be turned into a problem on sequences, by considering their partial sums.

- Lemma.** (i) If  $\sum_{j=1}^{\infty} a_j$  and  $\sum_{j=1}^{\infty} b_j$  converge, then so does  $\sum_{j=1}^{\infty} (\lambda a_j + \mu b_j)$ , where  $\lambda, \mu \in \mathbb{C}$ .
- (ii) Suppose  $\exists N$  such that  $a_j = b_j$  for all  $j \geq N$ . Then either  $\sum_{j=1}^{\infty} a_j$  and  $\sum_{j=1}^{\infty} b_j$  both converge, or they both diverge. In other words, the initial terms do not matter for considering convergence (but the sum will change).

*Proof.* (i) We have

$$\begin{aligned} s_N &= \sum_{j=1}^N (\lambda a_j + \mu b_j) \\ &= \sum_{j=1}^N \lambda a_j + \sum_{j=1}^N \mu b_j \\ &= \lambda c_N + \mu d_N \\ \therefore s_N &\rightarrow \lambda c + \mu d \end{aligned}$$

(ii) For any  $n \geq N$ , we have

$$s_N = \sum_{j=1}^n a_j = \sum_{j=1}^{N-1} a_j + \sum_{j=n}^N a_j$$

$$d_N = \sum_{j=1}^n b_j = \sum_{j=1}^{N-1} b_j + \sum_{j=n}^N b_j$$

Taking the difference, we get

$$s_N - d_N = \sum_{j=1}^{N-1} a_j - \sum_{j=1}^{N-1} b_j$$

which is finite. So  $s_N$  converges if and only if  $d_N$  also converges. □

## 2.2 Geometric series

Let  $a_n = x^{n-1}$ , where  $n \geq 1$ . Then

$$s_n = \sum_{j=1}^n a_j = 1 + x + x^2 + \cdots + x^{n-1}$$

Then

$$s_n = \begin{cases} \frac{1-x^n}{1-x} & \text{if } x \neq 1 \\ n & \text{if } x = 1 \end{cases}$$

This can be shown by observing that

$$xs_n = x + x^2 + \cdots + x^n = s_n - 1 + x^n \implies s_n(1-x) = 1 - x^n$$

If  $|x| < 1$ , then  $x^n \rightarrow 0$  as  $x \rightarrow \infty$ . So  $s_n \rightarrow \frac{1}{1-x}$ . If  $x > 1$ , then  $x^n \rightarrow \infty$  and so  $s_n \rightarrow \infty$ . If  $x < -1$ ,  $s_n$  oscillates. For completeness, if  $x = -1$ ,  $s_n$  oscillates between 0 and 1.

Note that the statement  $s_n \rightarrow \infty$  means that given  $a \in \mathbb{R}$ ,  $\exists N$  such that  $s_n > a$  for all  $n \geq N$ , and a similar statement holds for negative infinity (swapping the inequality). If  $s_n$  does not converge or tend to  $\pm\infty$ , we say that  $s_n$  oscillates.

Thus the geometric series converges if and only if  $|x| < 1$ . Note that to prove that  $x^n \rightarrow 0$  if  $|x| < 1$ , we can consider the case  $0 < x < 1$  and write  $1/x = 1 + \delta$  for some positive  $\delta$ . Then  $x^n = \frac{1}{(1+\delta)^n} \leq \frac{1}{1+\delta n}$  from the binomial expansion, and this tends to zero as required.

**Lemma.** If  $\sum_{j=1}^{\infty} a_j$  converges, then  $\lim_{j \rightarrow \infty} a_j = 0$ .

*Proof.* Given  $s_n = \sum_{j=1}^n a_j$ , we have  $a_n = s_n - s_{n-1}$ . If  $s_n \rightarrow a$ , then  $a_n \rightarrow 0$  since  $s_{n-1}$  also tends to  $a$ . □



*Remark.* The converse is not true. For example, the harmonic series diverges, but the terms approach zero. Consider

$$\begin{aligned} s_{2n} &= s_n + \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \\ &> s_n + \frac{1}{2n} + \frac{1}{2n} + \cdots + \frac{1}{2n} \\ &= s_n + \frac{1}{2} \end{aligned}$$

So as  $n \rightarrow \infty$ , if the sequence is convergent then the sequences  $s_n$  and  $s_{2n}$  tend to the same limit, but they clearly do not.

### 3 Convergence tests

#### 3.1 Comparison test

In this section, we will let  $a_n \in \mathbb{R}, a_n \geq 0$ . In other words, all series contain only non-negative real terms.

**Theorem.** Suppose  $0 \leq b_n \leq a_n$  for all  $n$ . If  $\sum_{j=1}^{\infty} a_j$  converges, then  $\sum_{j=1}^{\infty} b_j$  converges.

*Proof.* Let  $s_N$  be the  $N$ th partial sum over the  $a_n$ , and let  $d_N$  be the  $N$ th partial sum over the  $b_n$ . Since  $b_n \leq a_n, d_N \leq s_N$ . But  $s_N \rightarrow s$ , so  $d_N \leq s_N \leq s$ . So  $d_N$  is an increasing sequence that is bounded above by  $s$ , so it converges.  $\square$

For example, let us analyse the behaviour of the sum of the sequence  $\frac{1}{n^2}$ . Note that

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

for  $n \geq 2$ . By the comparison test, it is sufficient to show that the series on the right hand side converges, in order to show that the original series converges.

$$\sum_{j=2}^N a_j = 1 - \frac{1}{N} \rightarrow 1$$

as required. So the original series tends to some value less than or equal to 2.

#### 3.2 Cauchy's root test

**Theorem.** Suppose we have a sequence of non-negative terms  $a_n$ . Suppose that  $a_n^{1/n} \rightarrow a$  as  $n \rightarrow \infty$ . Then if  $a < 1$ , the series  $\sum a_n$  converges. If  $a > 1$ , the series  $\sum a_n$  diverges.

*Remark.* Nothing can be said if  $a = 1$ . There is an example later of this fact.

*Proof.* If  $a < 1$ , let us choose an  $r$  such that  $a < r < 1$ . By the definition of the limit,  $\exists N$  such that  $\forall n \geq N$ ,  $a_n^{1/n} < r$ . This implies that  $a_n < r^n$ . The geometric series  $\sum r^n$  converges. By comparison, the series  $a_n$  converges.

If  $a > 1$ , for all  $n \geq N$ ,  $a_n^{1/n} > 1$  which implies  $a_n > 1$ , thus  $\sum a_n$  diverges, since  $a_n$  does not tend to zero.  $\square$

### 3.3 D'Alembert's ratio test

**Theorem.** Suppose  $a_n > 0$ , and  $\frac{a_{n+1}}{a_n} \rightarrow \ell$ . If  $\ell < 1$ , then the series  $\sum a_n$  converges. If  $\ell > 1$ , then the series  $\sum a_n$  diverges.

*Remark.* Like before, no conclusion can be drawn if  $\ell = 1$ .

*Proof.* Suppose  $\ell < 1$ . We can choose  $\ell < r < 1$ ,  $\exists N$  such that  $\forall n \geq N$ ,  $\frac{a_{n+1}}{a_n} < r$ . Therefore  $a_n < r^{n-N} a_N$ . Hence,  $a_n < kr^n$  where  $k$  is independent of  $n$ . Applying the comparison test, the series  $\sum a_n$  must converge.

If  $\ell > 1$ , we can choose  $\ell > r > 1$ . Then  $\exists N$  such that  $\forall n \geq N$ ,  $\frac{a_{n+1}}{a_n} > r$ . As before,  $a_n > r^{n-N} a_N$ . But the  $r^{n-N}$  diverges, so the original series diverges.  $\square$

**Example.** Consider  $\sum_1^\infty \frac{n}{2^n}$ . We have

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)/2^{n+1}}{n/2^n} \rightarrow \frac{1}{2}$$

So we have convergence, by the ratio test. Now, consider  $\sum_1^\infty \frac{1}{n}$  and  $\sum_1^\infty \frac{1}{n^2}$ . In both cases, the ratio test gives limit 1. So the ratio test is inconclusive if the limit is 1. Since  $n^{1/n} \rightarrow 1$ , the root test is also inconclusive when the limit is 1. To check this limit, we can write

$$\begin{aligned} n^{1/n} &= 1 + \delta_n; \quad \delta_n > 0 \\ n &= (1 + \delta_n)^n > \frac{n(n-1)}{2} \delta_n^2 \end{aligned}$$

using the binomial expansion.

$$\implies \delta_n^2 < \frac{2}{n-1} \implies \delta_n \rightarrow 0$$

The root test is a good candidate for series that contain powers of  $n$ , for example

$$\sum_1^\infty \left[ \frac{n+1}{3n+5} \right]^n$$

In this instance, for example, we have convergence.

### 3.4 Cauchy's condensation test

**Theorem.** Let  $a_n$  be a decreasing sequence of positive terms. Then  $\sum_1^\infty a_n$  converges if and only if  $\sum_1^\infty 2^n a_{2^n}$  converges.

*Proof.* First, note that if  $a_n$  is decreasing, then

$$a_{2^k} \underset{(*)}{\leq} a_{2^{k-1}+i} \underset{(\dagger)}{\leq} a_{2^{k-1}}; \quad 1 \leq i \leq 2^{k-1}; \quad k \geq 1$$

Now let us assume that  $\sum a_n$  converges to  $A \in \mathbb{R}$ . Then, by  $(*)$ ,

$$\begin{aligned} 2^{n-1} a_{2^n} &= a_{2^n} + a_{2^n} + \cdots + a_{2^n} \\ &\leq a_{2^{n-1}+1} + a_{2^{n-1}+2} + \cdots + a_{2^n} \\ &= \sum_{m=2^{n-1}+1}^{2^n} a_m \end{aligned}$$

Thus,

$$\sum_{n=1}^N 2^{n-1} a_{2^n} \leq \sum_{n=1}^N \sum_{m=2^{n-1}+1}^{2^n} a_m = \sum_{n=2}^{2^N} a_m$$

Therefore,

$$\sum_{n=1}^N 2^n a_{2^n} \leq 2 \sum_{n=2}^{2^N} a_m \leq 2(A - a_1)$$

Thus  $\sum_{n=1}^N 2^n a_{2^n}$  converges, since it is increasing and bounded above. For the converse, we will assume that  $\sum 2^n a_{2^n}$  converges to  $B$ . Using  $(\dagger)$ ,

$$\begin{aligned} \sum_{m=2^{n-1}}^{2^n} a_m &= a_{2^{n-1}} + a_{2^{n-1}+1} + \cdots + a_{2^n} \\ &\leq a_{2^{n-1}} + a_{2^{n-1}} + \cdots + a_{2^{n-1}} \\ &= 2^{n-1} a_{2^{n-1}} \end{aligned}$$

So we have

$$\sum_{m=2}^{2^N} a_m = \sum_{n=1}^N \sum_{m=2^{n-1}+1}^{2^n} a_m \leq \sum_{n=1}^N 2^{n-1} a_{2^{n-1}} \leq \frac{1}{2} B$$

Therefore,  $\sum_{m=1}^N a_m$  is a bounded, increasing sequence and hence converges.  $\square$

Let us consider an example of this test. Consider the series definition of the Riemann zeta function

$$\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k}$$

For what  $k \in \mathbb{R}, k > 0$  does this series converge? This is equivalent to asking if the following series converges.

$$\sum_{n=1}^{\infty} 2^n \left[ \frac{1}{2^n} \right]^k = \sum_{n=1}^{\infty} (2^{1-k})^n$$

Hence it converges if and only if  $2^{1-k} < 1 \iff k > 1$ .

### 3.5 Alternating series

An alternating series is a series where the sign on each term switches between positive and negative.

**Theorem** (Alternating Series Test). If  $a_n$  decreases and tends to zero as  $n \rightarrow \infty$ , then the alternating series

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

converges.

*Proof.* Let us consider the partial sum

$$s_n = a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{n+1} a_n$$

In particular,

$$s_{2n} = (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n})$$

Since the sequence is decreasing, each parenthesised block is positive. Then  $s_{2n} \geq s_{2n-2}$ . We can also write the partial sum as

$$s_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n}$$

Each parenthesised block here is negative. So  $s_{2n} \leq a_1$ . So  $s_{2n}$  is increasing and bounded above, so it must converge. Now, note that

$$s_{2n+1} = s_{2n} + a_{2n+1} \rightarrow s_{2n}$$

since  $a_{2n+1} \rightarrow 0$ . So  $s_{2n+1}$  also converges, in fact to the same limit. Hence  $s_n$  converges to this same limit.  $\square$

## 4 Absolute convergence

### 4.1 Absolute convergence

**Definition.** Let  $a_n \in \mathbb{C}$ . Then if  $\sum_{n=1}^{\infty} |a_n|$  converges, then the series is called absolutely convergent.

*Remark.* Since  $|a_n| \geq 0$ , we can use the previous tests to check for absolute convergence.

**Theorem.** Let  $a_n \in \mathbb{C}$ . If this series is absolutely convergent, it is convergent.

*Proof.* Suppose first that  $a_n$  is a sequence of real numbers. Then let

$$v_n = \begin{cases} a_n & \text{if } a_n \geq 0 \\ 0 & \text{if } a_n < 0 \end{cases}; \quad w_n = \begin{cases} 0 & \text{if } a_n \geq 0 \\ -a_n & \text{if } a_n < 0 \end{cases}$$

Hence,

$$v_n = \frac{|a_n| + a_n}{2}; \quad w_n = \frac{|a_n| - a_n}{2}$$

Clearly,  $v_n, w_n \geq 0$ , and  $a_n = v_n - w_n$ , and  $|a_n| = v_n + w_n$ . If  $\sum |a_n|$  converges, then by comparison  $\sum v_n$  and  $\sum w_n$  also converge, and hence  $\sum a_n$  converges. Now, let us consider the case where  $a_n$  is complex. Then we can write  $a_n = x_n + iy_n$  where  $x_n, y_n$  are real sequences. Note that  $|x_n|, |y_n| \leq |a_n|$ . So by comparison  $x_n$  and  $y_n$  converge, so  $a_n$  converges.  $\square$

Here are some examples.

- (i) The alternating harmonic series  $\sum \frac{(-1)^n}{n}$  is convergent, but not absolutely convergent.
- (ii)  $\sum \frac{z^n}{2^n}$  is absolutely convergent when  $|z| < 2$ , because it reduces to a real geometric series. If  $|z| \geq 2$ , then  $|a_n| \geq 1$ , so we do not have absolute convergence.

## 4.2 Conditional convergence and rearrangement

If the series is convergent but not absolutely convergent, it is called *conditionally* convergent. The sum to which a series converges depends on the order in which the terms are added.

**Definition.** Let  $\sigma$  be a bijection of the positive integers to itself, then

$$a'_n = a_{\sigma(n)}$$

is a rearrangement of  $a_n$ .

**Theorem.** If  $\sum_1^\infty a_n$  is absolutely convergent, then every rearrangement of this series converges to the same value.

*Proof.* First, let us consider the real case. Let  $\sum a'_n$  be a rearrangement of  $\sum a_n$ . Let  $s_n = \sum_1^n a_n$ , and  $t_n = \sum_1^n a'_n$ . Let  $s_n$  converge to  $s$ . Suppose first that  $a_n \geq 0$ . Then given any  $n \in \mathbb{N}$ , we can find some  $q \in \mathbb{N}$  such that  $s_q$  contains every term of  $t_n$ . Since the  $a_n \geq 0$ ,

$$t_n \leq s_q \leq s$$

As  $n \rightarrow \infty$ , the  $t_n$  is an increasing sequence bounded above, so it must tend to a limit  $t$ , where  $t \leq s$ . Note, however, that this argument is symmetric; we can equally derive that  $s \leq t$ . Therefore  $s = t$ .

Now, let us drop the condition that  $a_n \geq 0$ . We can now consider  $v_n, w_n$  from above:

$$v_n = \frac{|a_n| + a_n}{2}; \quad w_n = \frac{|a_n| - a_n}{2}$$

Since  $\sum |a_n|$  converges, both  $\sum v_n, \sum w_n$  converge. Since all  $v_n, w_n \geq 0$ , we can deduce that  $\sum v_n = \sum v'_n$  and  $\sum w'_n = \sum w_n$ . The claim follows since  $a_n = v_n - w_n$ .

For the case  $a_n \in \mathbb{C}$ , we can write  $a_n = x_n + iy_n$ , noting that  $|x_n|, |y_n| \leq |a_n|$ . By comparison, the series  $\sum x_n, \sum y_n$  are absolutely convergent, and by the previous case,  $\sum x_n = \sum x'_n$  and  $\sum y'_n = \sum y_n$ . Since  $a'_n = x'_n + iy'_n$ ,  $\sum a_n = \sum a'_n$  as required.  $\square$

## 5 Continuity

### 5.1 Definitions

Let  $E \subseteq \mathbb{C}$  be a non-empty set, and  $f : E \rightarrow \mathbb{C}$  be any function, and let  $a \in E$ . Certainly, this includes the case in which  $f$  is a real-valued function and  $E \subseteq \mathbb{R}$ .

**Definition.**  $f$  is continuous at  $a$  if for every sequence  $z_n \in E$  that converges to  $a$ , we have  $f(z_n) \rightarrow f(a)$ .

We can use an alternative definition:

**Definition** ( $\varepsilon$ - $\delta$  definition).  $f$  is continuous at  $a$  if given  $\varepsilon > 0$ ,  $\exists \delta > 0$  such that for every  $z \in E$ , if  $|z - a| < \delta$ , then  $|f(z) - f(a)| < \varepsilon$ .

We will immediately prove that both definitions are equivalent. First, let us prove that the  $\varepsilon$ - $\delta$  definition implies the first definition.

*Proof.* We know that given  $\varepsilon > 0$ ,  $\exists \delta > 0$  such that for all  $z \in E$ ,  $|z - a| < \delta$  implies  $|f(z) - f(a)| < \varepsilon$ . Let  $z_n \rightarrow a$ , then by the definition of the limit of the sequence then there exists  $n_0$  such that for all  $n \geq n_0$  we have  $|z_n - a| < \delta$ . But this implies that  $|f(z_n) - f(a)| < \varepsilon$ , i.e.  $f(z_n) \rightarrow f(a)$ .  $\square$

We now prove the converse, that the first definition implies the second.

*Proof.* We know that for every sequence  $z_n \in E$  that converges to  $a$ ,  $f(z_n) \rightarrow f(a)$ . Suppose  $f$  is not continuous at  $a$ , according to the  $\varepsilon$ - $\delta$  definition. Then there exists some  $\varepsilon$  such that for all  $\delta > 0$ , there exists  $z \in E$  such that  $|z - a| < \delta$  but  $|f(z) - f(a)| \geq \varepsilon$ . So, let us construct a sequence of  $\delta$  values to substitute into this definition. Let  $\delta = 1/n$ . Then the  $z_n$  given by this  $\delta$  is such that  $|z_n - a| < 1/n$  and  $|f(z_n) - f(a)| \geq \varepsilon$ . Clearly,  $z_n \rightarrow a$ , but  $f(z_n)$  does not tend to  $f(a)$  because the difference between the two is always greater than  $\varepsilon$ . This is a contradiction, since we assumed that  $f$  is continuous by the first definition. So  $f$  is continuous by the  $\varepsilon$ - $\delta$  definition.  $\square$

### 5.2 Making continuous functions

We can create new continuous functions from old ones by manipulating them in a number of ways.

**Proposition.** Let  $g, f : E \rightarrow \mathbb{C}$  be continuous functions at a point  $a \in E$ . Then all of the functions

- $f(z) + g(z)$
- $f(z)g(z)$
- $\lambda f(z)$  for some constant  $\lambda$

are all continuous. In addition, if  $f(z) \neq 0$  everywhere in  $E$ , then  $\frac{1}{f}$  is a continuous function at  $a$ .

*Proof.* Using the first definition, this is obvious using the fact that limits of sequences behave analogously.  $\square$

Trivially, the function  $f(z) = z$  is continuous. From this, we can derive that every polynomial is continuous at every point in  $\mathbb{C}$ . Note that we say that  $f$  is continuous on the entire set  $E$  if it is continuous at every point  $a \in E$ .

### 5.3 Composition of continuous functions

**Theorem.** Let  $f : A \rightarrow \mathbb{C}$  and  $g : B \rightarrow \mathbb{C}$  where  $A, B \subseteq \mathbb{C}$  be two functions that can be composed, i.e.  $f(A) \subseteq B$ . If  $f$  is continuous at  $a \in A$  and  $g$  is continuous at  $f(a) \in B$ , then  $g \circ f : A \rightarrow \mathbb{C}$  is continuous at  $a$ .

*Proof.* Take any sequence  $z_n \rightarrow a$ . By assumption,  $f(z_n) \rightarrow f(a)$ . Now, let us define a new sequence  $w_n = f(z_n)$ . Then  $w_n \in B$  and  $w_n \rightarrow f(a)$ . Thus,  $g(f(z_n)) = g(w_n) \rightarrow g(f(a))$  by continuity, as required.  $\square$

Consider the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  defined by

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

This is assuming the knowledge of  $\sin(x)$  being a continuous function  $\mathbb{R} \rightarrow \mathbb{R}$ , which we will prove later. So  $f(x)$  is certainly continuous at every point on  $\mathbb{R}$  excluding 0, since it is the composition of two continuous functions. We can prove it is discontinuous at  $x = 0$  by providing a sequence, for example

$$\frac{1}{x_n} = \left(2n + \frac{1}{2}\right)\pi$$

Then  $x_n \rightarrow 0$ , and  $f(x_n) = 1$ . But  $f(0) \neq 1$ , so it is discontinuous. Let us modify the example as follows.

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We can prove that this sequence is continuous at 0. For an arbitrary sequence  $x_n \rightarrow 0$ , then  $|f(x_n)| \leq |x_n|$  because  $|\sin x| \leq 1$ . So  $f(x_n)$  is bounded by  $x_n$ , which tends to zero, so  $f(x_n)$  tends to zero as required. Now for a final example, let

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

This is discontinuous at every point. If  $x \in \mathbb{Q}$ , take a sequence  $x_n \rightarrow x$  with all  $x_n$  irrational, then  $f(x_n) = 0$  but  $f(x) = 1$ . Similarly, if  $x \notin \mathbb{Q}$ , take a sequence  $x_n \rightarrow x$  with all  $x_n$  rational, then  $f(x_n) = 1$  but  $f(x) = 0$ .

## 6 Limit of a function

### 6.1 Definition

Let  $f : E \subseteq \mathbb{C} \rightarrow \mathbb{C}$ . We would like to define what is meant by  $\lim_{z \rightarrow a} f(z)$ , even when  $a \notin E$ . Further, if we have a set with an isolated point, for example  $E = \{0\} \cup [1, 2]$ , it does not make sense to talk about limits tending to 0 since there are no points in  $E$  close to 0.

**Definition.** Let  $E \subseteq \mathbb{C}$ ,  $a \in \mathbb{C}$ .  $a$  is a limit point of  $E$  if for any  $\delta > 0$ , there exists  $z \in E$  such that  $0 < |z - a| < \delta$ .

First, note that  $a$  is a limit point if and only if there exists a sequence  $z_n \in E$  such that  $z_n \rightarrow a$ , but notably  $z_n \neq a$  for all  $n$ .

**Definition.** Let  $f : E \subseteq \mathbb{C} \rightarrow \mathbb{C}$ , and let  $a \in \mathbb{C}$  be a limit point of  $E$ . We say that  $f \rightarrow \ell$  as  $z \rightarrow a$ , if given  $\varepsilon > 0$  there exists  $\delta > 0$  such that whenever  $0 < |z - a| < \delta$  and  $z \in E$ ,  $|f(z) - \ell| < \varepsilon$ . Equivalently,  $f(z_n) \rightarrow \ell$  for every sequence  $z_n \in E$ , such that  $z_n \rightarrow a$  but  $z_n \neq a$ .

Therefore if  $a \in E$  is a limit point, then  $\lim_{z \rightarrow a} f(z) = f(a)$  if and only if  $f$  is continuous at  $a$ . If  $a \in E$  is isolated (not a limit point) then  $f$  at  $a$  is trivially continuous, since there are no points near  $a$  but  $a$  itself.

## 6.2 Properties

The limit of a function has very similar properties when compared to the limit of a sequence.

- (i) It is unique.  $f(z) \rightarrow A, f(z) \rightarrow B$  implies  $A = B$ .
- (ii)  $f(z) \rightarrow A, g(z) \rightarrow B$  implies
  - (a)  $f(z) + g(z) \rightarrow A + B$
  - (b)  $f(z) \cdot g(z) \rightarrow AB$
  - (c) If  $B \neq 0$ ,  $\frac{f(z)}{g(z)} \rightarrow \frac{A}{B}$

## 6.3 Intermediate value theorem

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous function where  $f(a) \neq f(b)$ . Then  $f$  takes all values in the interval  $[f(a), f(b)]$ .

*Proof.* Without loss of generality, let us assume  $f(a) < f(b)$ . Let us take an  $\eta$  such that  $f(a) < \eta < f(b)$ . We want to prove that there exists some value  $c \in [a, b]$  with  $f(c) = \eta$ . Let  $s$  be the set of points defined by

$$s = \{x \in [a, b] : f(x) < \eta\}$$

$a \in s$  therefore the set  $s$  is non-empty. The set is also clearly bounded above by  $b$ . So there is a supremum of this set, say  $\sup s = c$  where  $c \leq b$ . This point  $c$  can be visualised as the last point at which  $y = f(x)$  crosses the line  $y = \eta$ . We intend to show that the function at this rightmost point is  $\eta$ .

By the definition of the supremum, given  $n$  there exists  $x_n \in s$  such that  $c - \frac{1}{n} < x_n \leq c$ . So the sequence  $x_n$  tends to  $c$ . We know that  $f(x_n) < \eta$  for all  $x_n$  by definition of the set  $s$ . By the continuity of  $f$ ,  $f(x_n) \rightarrow f(c)$ . Thus,

$$f(c) \leq \eta \tag{*}$$



Now, let us consider the fact that  $c \neq b$ . If  $c = b$ , then  $f(b) \leq \eta$  which is a contradiction since  $\eta < f(b)$ . So for a large  $n$ , we can ensure that  $c + \frac{1}{n} \in [a, b]$ . So by continuity of the function,  $f(c + \frac{1}{n}) \rightarrow f(c)$ . But since  $c + \frac{1}{n} > c$ , then necessarily  $f(c + \frac{1}{n}) \geq \eta$  because  $c$  is the supremum of  $s$ . Thus

$$f(c) \geq \eta$$

Combining this with (\*) we get  $f(c) = \eta$ . □

This theorem is very useful for finding zeroes and fixed points. For example, we can prove the existence of the  $N$ th root of a positive real number  $y$ . Let

$$f(x) = x^N$$

Then  $f$  is certainly continuous on the interval  $[0, 1 + y]$ , since

$$0 = f(0) < y < (1 + y)^N = f(1 + y)$$

By the intermediate value theorem, there exists a point  $c \in (0, 1 + y)$  such that  $f(c) = c^N = y$ . So  $c$  is a positive  $N$ th root of  $y$ . We can also prove the uniqueness of such a point. Suppose  $d^N = y$  with  $d > 0$  and  $d \neq c$ . Without loss of generality, suppose  $d < c$ . Then  $d^N < c^N$  so  $d^N \neq y$ , which is a contradiction.

## 6.4 Bounds of a continuous function

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous. Then the function is bounded, i.e. there exists  $k \in \mathbb{R}$  such that  $|f(x)| \leq k$  for every point  $x \in [a, b]$ .

*Proof.* Suppose that such a function  $f$  is not bounded. Then in particular, given any integer  $n \geq 1$ , there exists  $x_n \in [a, b]$  such that  $|f(x_n)| > n$ . By the Bolzano–Weierstrass theorem, the sequence  $x_n$ , which is bounded by  $a \leq x_n \leq b$ , has a convergent subsequence  $x_{n_j} \rightarrow x$ , such that  $x \in [a, b]$ . Then by continuity of  $f$ ,  $f(x_{n_j}) \rightarrow f(x)$ . But  $|f(x_{n_j})| > n_j \rightarrow \infty$ . This is a contradiction. □

We can actually improve this statement.

**Theorem.** Suppose  $f : [a, b] \rightarrow \mathbb{R}$  is a continuous function. Then there exist  $x_1, x_2 \in [a, b]$  such that

$$f(x_1) \leq f(x) \leq f(x_2)$$

for all  $x \in [a, b]$ . In other words, a continuous function on a closed bounded interval is bounded and attains its bounds.

*Proof.* Let  $A = \{f(x) : x \in [a, b]\}$  be the image of  $[a, b]$  under  $f$ . By the above theorem,  $A$  is bounded. It is also non-empty, hence it has a supremum  $M = \sup A$  (and analogously an infimum  $\inf A$ , whose proof is almost identical). Then by the definition of the supremum, given an integer  $n \geq 1$  there exists  $x_n \in [a, b]$  such that  $M - \frac{1}{n} < f(x_n) \leq M$ . By the Bolzano–Weierstrass theorem, there exists a convergent subsequence  $x_{n_j} \rightarrow x \in [a, b]$ . Since  $f(x_{n_j}) \rightarrow M$ , then by continuity,  $f(x) = M$ . □

Here is an alternative proof of the same theorem.

*Proof.* As before, let  $A$  be the image of  $f$ , and  $M$  be the supremum of  $A$ . Suppose there is no  $x_2 \in [a, b]$  such that  $f(x_2) = M$ . Then let  $g(x) = \frac{1}{M-f(x)}$  for  $x \in [a, b]$ . Since there exists no  $x$  such that  $M = f(x)$ ,  $g(x)$  is continuous since we are never dividing by zero. So  $g$  is bounded. So by the previous theorem, there is some  $k > 0$  such that  $g(x) \leq k$  for all  $x \in [a, b]$ . This means that  $f(x) \leq M - \frac{1}{k}$  on  $[a, b]$  for this  $k$ , but this cannot happen since  $M$  is the supremum.  $\square$

Note that these theorems are certainly false if the interval is not closed: consider the counterexample  $(0, 1]$  and the function  $x \mapsto x^{-1}$ .

## 6.5 Inverse functions

**Definition.**  $f$  is increasing for  $x \in [a, b]$  if  $f(x_1) \leq f(x_2)$  for all  $x_1 \leq x_2 \in [a, b]$ . If  $f(x_1) < f(x_2)$  then the function is strictly increasing. A function may be called decreasing or strictly decreasing analogously.

**Definition.** A function  $f$  is called monotone if it is either increasing or decreasing.

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous and strictly increasing for  $x \in [a, b]$ . Let  $c = f(a)$ ,  $d = f(b)$ . Then  $f : [a, b] \rightarrow [c, d]$  is bijective, and the inverse  $g := f^{-1} : [c, d] \rightarrow [a, b]$  is continuous and strictly increasing.

A similar theorem holds for strictly decreasing functions.

*Proof.* Let  $c < k < d$ . From the intermediate value theorem, there exists  $h$  such that  $f(h) = k$ . This  $h$  must be unique since the function is strictly increasing. Then we can define  $g(k) = h$ , giving us an inverse  $g : [c, d] \rightarrow [a, b]$  for  $f$ .

First, note that  $g$  is strictly increasing. Indeed, for  $y_1 < y_2$  then  $y_1 = f(x_1)$ ,  $y_2 = f(x_2)$ . This means that if  $x_2 \geq x_1$ , then since  $f$  is increasing  $y_2 \leq y_1$  which is a contradiction.

Now, note that  $g$  is continuous. Indeed, given  $\varepsilon > 0$ , we can let  $k_1 = f(h - \varepsilon)$  and  $k_2 = f(h + \varepsilon)$ . If  $f$  is strictly increasing, then  $k_1 < k < k_2$ . Then  $h - \varepsilon < g(y) < h + \varepsilon$ . So let  $\delta = \min(k_2 - k, k - k_1)$  where  $k \in (c, d)$ , establishing continuity as claimed.  $\square$

## 7 Differentiability

### 7.1 Definitions

Let  $f : E \subseteq \mathbb{C} \rightarrow \mathbb{C}$ . Mostly we will take  $E$  to be an interval in the real numbers, or a disc in the complex plane.

**Definition.** Let  $x \in E$  be a point such that there exists a sequence  $x_n \in E$  with  $x_n \neq x$ , but

$x_n \rightarrow x$ , i.e.  $x$  is a limit point.  $f$  is said to be differentiable at  $x$  with derivative  $f'(x)$  if

$$\lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} = f'(x)$$

If  $f$  is differentiable at each point in  $E$ , we say that  $f$  is differentiable on  $E$ .

*Remark.* One interpretation of the definition is to write it in the form

$$\varepsilon(h) := f(x + h) - f(x) - hf'(x); \quad \lim_{h \rightarrow 0} \frac{\varepsilon(h)}{h} = 0$$

so  $\varepsilon$  is  $o(h)$ . Hence,

$$f(x + h) = f(x) + hf'(x) + \varepsilon(h)$$

We could have made an alternative definition for differentiability.  $f$  is differentiable at  $x$  if there exists  $A$  and  $\varepsilon$  such that

$$f(x + h) = f(x) + hA + \varepsilon(h) \text{ where } \lim_{h \rightarrow 0} \frac{\varepsilon(h)}{h} = 0$$

If such an  $A$  exists, then it is unique, since  $A$  is the limit

$$A = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

We could have alternatively written the definition as

$$f(x + h) = f(x) + hf'(x) + h\varepsilon_f(h) \text{ where } \lim_{h \rightarrow 0} \varepsilon_f(h) = 0$$

or perhaps

$$f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon_f(x) \text{ where } \lim_{x \rightarrow a} \varepsilon_f(x) = 0$$

Note further that if  $f$  is differentiable at  $x$ ,  $f$  is certainly continuous at  $x$ . This follows from the fact that  $\varepsilon(h) \rightarrow 0$ , and hence  $f(x + h) \rightarrow f(x)$  as  $h \rightarrow 0$ .

As an example, let us consider  $f(x) = |x|$  for  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Is the function at the point  $x = 0$  differentiable? If  $x > 0$ , we have  $f'(x) = 1$ , but if  $x < 0$ , we have  $f'(x) = -1$ . These results can be checked directly using the definitions above. But we have produced two sequences for  $h \rightarrow 0$  which give different values, so the derivative is not defined here.

## 7.2 Differentiation of sums and products

**Proposition.** (i) If  $f(x) = c$  for all  $x \in E$ , then  $f$  is differentiable with  $f'(x) = 0$ .

(ii) If  $f$  and  $g$  are differentiable at  $x$ , then so is  $f + g$ , where  $(f + g)'(x) = f'(x) + g'(x)$ .

(iii) If  $f$  and  $g$  are differentiable at  $x$ , then so is  $fg$ , where  $(fg)'(x) = f'(x)g(x) + g'(x)f(x)$ .

(iv) If  $f$  is differentiable at  $x$  and  $f(x) \neq 0$ , then so is  $\frac{1}{f}$ , where  $(\frac{1}{f})'(x) = \frac{-f'(x)}{(f(x))^2}$ .

*Proof.* (i)  $\lim_{h \rightarrow 0} \frac{c - c}{h} = 0$  as required.

(ii) Since all relevant limits are well-defined,

$$\lim_{h \rightarrow 0} \frac{f(x+h) + g(x+h) - f(x) - g(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} + \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} = f'(x) + g'(x)$$

(iii) Let  $\phi(x) = f(x)g(x)$ . Then, since  $f$  is continuous at  $x$ ,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h} &= \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h \rightarrow 0} f(x+h) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \rightarrow 0} f(x) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h} \\ &= f(x)g'(x) + g(x)f'(x) \end{aligned}$$

(iv) Let  $\phi(x) = \frac{1}{f(x)}$ . Then,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h} &= \lim_{h \rightarrow 0} \frac{\frac{1}{f(x+h)} - \frac{1}{f(x)}}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x) - f(x+h)}{hf(x)f(x+h)} \\ &= \frac{-f'(x)}{f(x)f(x)} \end{aligned}$$

□

*Remark.* From (iii) and (iv), we can immediately find the quotient rule,

$$\left( \frac{f(x)}{g(x)} \right)' = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}$$

### 7.3 Differentiating polynomial terms

As an example of the differentiability properties we saw last lecture, we can find the derivative of  $f(x) = x^n$  for  $n \in \mathbb{Z}$ ,  $n > 0$ . If  $n = 1$ , clearly  $f'(x) = 1$ . We can show inductively that  $f'(x) = nx^{n-1}$ . Indeed,

$$\begin{aligned} (x^n)' &= x \cdot (x^{n-1})' + (x)' \cdot x^{n-1} \\ &= (n-1)x^{n-1} + x^{n-1} \\ &= nx^{n-1} \end{aligned}$$

We can now take  $f(x) = x^{-n}$ . Using the reciprocal law,

$$\begin{aligned} f'(x) &= \frac{-(x^n)'}{(x^n)^2} \\ &= \frac{-nx^{n-1}}{x^{2n}} \\ &= -nx^{-n-1} \end{aligned}$$

## 7.4 Chain rule

**Theorem.** Let  $f : U \rightarrow \mathbb{C}$  be such that  $f(x) \in V$  for all  $x \in U$ . If  $f$  is differentiable at  $a \in U$ , and  $g : V \rightarrow \mathbb{C}$  is differentiable at  $f(a) \in V$ , then  $g \circ f$  is differentiable at  $a$  with

$$gf'(a) = f'(a)g'(f(a))$$

*Proof.* We know that we can write

$$f(x) = f(a) + (x - a)f'(a) + \varepsilon_f(x)(x - a)$$

where  $\lim_{x \rightarrow a} \varepsilon_f(x) = 0$ . Further,

$$g(y) = g(b) + (y - b)g'(b) + \varepsilon_g(y)(y - b)$$

where  $\lim_{y \rightarrow b} \varepsilon_g(y) = 0$ , and  $b = f(a)$ . We will set  $\varepsilon_f(a) = 0$  and  $\varepsilon_g(b) = 0$ , so they are continuous at  $x = a$  and  $y = b$ , so that everything is well-defined when we begin to compose the functions. Now,  $y = f(x)$ , so

$$\begin{aligned} g(f(x)) &= g(b) + (f(x) - b)g'(b) + \varepsilon_g(f(x))(f(x) - b) \\ &= g(f(a)) + [(x - a)f'(a) + \varepsilon_f(x)(x - a)][g'(b) + \varepsilon_g(f(x))] \\ &= g(f(a)) + (x - a)f'(a)g'(b) + (x - a) \underbrace{[\varepsilon_f(x)g'(b) + \varepsilon_g(f(x))(f'(a) + \varepsilon_f(x))]}_{\sigma(x)} \end{aligned}$$

Now, we just need to show that  $\lim_{x \rightarrow a} \sigma(x) = 0$  in order to prove the theorem. Clearly

$$\sigma(x) = \underbrace{\varepsilon_f(x)}_{\rightarrow 0} g'(b) + \underbrace{\varepsilon_g(f(x))}_{\rightarrow 0} (f'(a) + \varepsilon_f(x))$$

Hence  $\sigma(x) \rightarrow 0$  as required. □

## 7.5 Rolle's theorem

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous function on  $[a, b]$  and differentiable on  $(a, b)$ . If  $f(a) = f(b)$ , then there exists  $c \in (a, b)$  such that  $f'(c) = 0$ .

*Proof.* Let  $M$  be the maximum point and  $m$  be the minimum point of the function. Recall that in Lecture 8 we proved that any function achieves its bounds. Let  $k = f(a)$ . If  $M = m = k$ , then  $f$  must be a constant, and clearly  $f'(c) = 0$  for every value  $c \in (a, b)$ . Otherwise, either  $M > k$  or  $m < k$ . Suppose  $M > k$  (the proof is very similar if  $m < k$ ). Then there exists some value  $c \in (a, b)$  such that  $f(c) = M$ . We would like to show that  $f'(c) = 0$ , so let us suppose that  $f'(c) \neq 0$ . If  $f'(c) > 0$ , then there are values  $d > c$  where  $f(d) > f(c)$ . Indeed,

$$f(h + c) - f(c) = h[f'(c) + \varepsilon(h)]$$

For a small, positive  $h$ , this value is positive. This contradicts the fact that  $M$  is the maximum. Similarly, if  $f'(c) < 0$  there are values  $d < c$  with  $f(d) > f(c)$ . Hence  $f'(c) = 0$ . □

## 7.6 Mean value theorem

We can make a small change to Rolle's theorem and obtain the mean value theorem.

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous function on  $[a, b]$  and differentiable on  $(a, b)$ . Then there exists  $c \in (a, b)$  such that

$$f(b) - f(a) = f'(c)(b - a)$$

*Proof.* Let  $\phi$  be a function defined by  $\phi(x) = f(x) - kx$ , choosing a  $k$  such that  $\phi(a) = \phi(b)$ . We can find that

$$f(b) - bk = f(a) - ak \implies k = \frac{f(b) - f(a)}{b - a}$$

By Rolle's theorem, there exists  $c \in (a, b)$  such that  $\phi'(c) = 0$ . Now, note that  $f'(x) = \phi'(x) + k$ , hence there exists  $c$  such that  $f'(c) = k$ .  $\square$

*Remark.* We will often rewrite the mean value theorem as follows.

$$f(a + h) = f(a) + hf'(a + \theta h)$$

where  $\theta \in (0, 1)$ . Note, however, that  $\theta$  is a function of  $h$ , so if we begin to shrink  $h$  then  $\theta$  may change.

## 7.7 Properties of a function from its derivative

We can deduce certain facts about a function by observing the properties its derivative exhibits. These results are mostly trivial corollaries to the mean value theorem, proven in the last lecture.

**Corollary.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous, and differentiable on  $(a, b)$ . Then we have

- (i) If  $f'(x) > 0$  for all  $x \in (a, b)$ , then  $f$  is strictly increasing on  $[a, b]$ ;
- (ii) If  $f'(x) \geq 0$  for all  $x \in (a, b)$ , then  $f$  is increasing on  $[a, b]$ ;
- (iii) If  $f'(x) = 0$  for all  $x \in (a, b)$ , then  $f$  is constant on  $[a, b]$ .

Part (iii) of this corollary is essentially solving the most simple differential equation; we are showing that the only possible solutions to this equation are the constant functions. Note that similar statements about decreasing functions hold.

*Proof.* (i) We have  $f(y) - f(x) = f'(c)(y - x)$  for some  $c \in (x, y)$ . If  $f'(c) > 0$ , then  $f(y) - f(x) > 0$ .

(ii) Analogously to before,  $f(y) - f(x) = f'(c)(y - x)$  for some  $c \in (x, y)$ . If  $f'(c) \geq 0$ , then  $f(y) - f(x) \geq 0$ .

(iii) By the mean value theorem on  $[a, x]$ , if  $f'(c) = 0$ , then  $f(x) - f(a) = 0$ .

$\square$

## 7.8 Inverse function theorem

**Theorem.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous function on  $[a, b]$  and differentiable on  $(a, b)$ , with  $f'(x) > 0$  everywhere on  $(a, b)$ . Let  $f(a) = c, f(b) = d$ . Then the function  $f : [a, b] \rightarrow [c, d]$  is bijective, and  $f^{-1} : [c, d] \rightarrow [a, b]$  is differentiable on  $(c, d)$  with

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$$

Note, in lecture 8 it was proven that a continuous strictly increasing function has a continuous inverse. This strengthens that claim to include the differentiability property if the original function was differentiable.

*Proof.* We know from lecture 8 that there exists  $g : [c, d] \rightarrow [a, b]$  which is a strictly increasing continuous function, which is the inverse of  $f$ . We must now show that  $g$  is differentiable and that its derivative has the required form as stated in the claim. Now, let  $y = f(x)$ . Given  $k \neq 0$ , let  $h$  be given by

$$y + k = f(x + h)$$

Alternatively, written in terms of  $g$ ,

$$x + h = g(y + k)$$

So clearly  $h \neq 0$ . Since  $g$  is continuous, if  $k \rightarrow 0$  then  $h \rightarrow 0$ . Then

$$\begin{aligned} \frac{g(y+k) - g(y)}{k} &= \frac{x+h-x}{f(x+h)-y} \\ &= \frac{h}{f(x+h)-f(x)} \\ \therefore \lim_{k \rightarrow 0} \frac{g(y+k) - g(y)}{k} &= \lim_{h \rightarrow 0} \frac{h}{f(x+h)-f(x)} \\ &= \frac{1}{f'(x)} \end{aligned}$$

as required. □

## 7.9 Derivative of rational powers

First, let  $g(x) = x^{1/q}$  for some positive integer  $q$ . We can find that  $f(x) = x^q$  has the derivative  $f'(x) = qx^{q-1}$ . By the inverse function theorem,  $g'(x) = \frac{1}{q}x^{1/q-1}$ . Now, if  $g(x) = x^{p/q}$ , where  $p$  is an integer and  $q$  is a positive integer, then by the chain rule  $g'(x) = \frac{p}{q}x^{p/q-1}$  which matches the expected result.

## 7.10 Mean value theorem applied to limits

Suppose  $f, g : [a, b] \rightarrow \mathbb{R}$  are continuous, and differentiable on  $(a, b)$ . Suppose further that  $g(a) \neq g(b)$ . The mean value theorem can be applied to both functions, and will give two points  $s, t \in (a, b)$  such that

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{(b-a)f'(s)}{(b-a)g'(t)} = \frac{f'(s)}{g'(t)}$$

This gives us a way to simplify a limit of the form of the left hand side (as  $b \rightarrow a$ ) by instead considering the right hand side. We can apply Cauchy's mean value theorem, seen in the next lecture.

## 7.11 Cauchy's mean value theorem

**Theorem.** If  $f, g : [a, b] \rightarrow \mathbb{R}$  are continuous, and differentiable on  $(a, b)$ , there exists  $t \in (a, b)$  such that

$$(f(b) - f(a))g'(t) = f'(t)(g(b) - g(a))$$

We can recover the normal mean value theorem from Cauchy's generalisation by taking  $g(x) = x$ .

*Proof.* Let

$$\phi(x) = \begin{vmatrix} 1 & 1 & 1 \\ f(a) & f(x) & f(b) \\ g(a) & g(x) & g(b) \end{vmatrix}$$

Certainly  $\phi(x)$  is continuous on  $[a, b]$  and differentiable on  $(a, b)$ , by using previous results. Also,  $\phi(a) = \phi(b) = 0$  by observing the linear dependence of the columns. By Rolle's theorem, there exists  $t \in (a, b)$  such that  $\phi'(t) = 0$ . We can expand  $\phi'(t)$  and this will show the required result.

$$\phi'(x) = f'(x)g(b) - g'(x)f(b) + f(a)g'(x) - g(a)f'(x) = f'(x)[g(b) - g(a)] + g'(x)[f(a) - f(b)]$$

□

**Example** (l'Hôpital's rule). The derivation of l'Hôpital's rule is on an example sheet, so here we will consider only a special case of it, using Cauchy's mean value theorem.

$$\ell = \lim_{x \rightarrow 0} \frac{e^x - 1}{\sin x}$$

We can write

$$\ell = \lim_{x \rightarrow 0} \frac{e^x - e^0}{\sin x - \sin 0} = \frac{e^t}{\cos t}$$

for some  $t \in (0, x)$ . So as  $x \rightarrow 0$ ,  $t \rightarrow 0$  and hence

$$\frac{e^t}{\cos t} \rightarrow 1$$

## 8 Taylor's theorem

### 8.1 Lagrange's and Cauchy's remainders

**Theorem** (Taylor's Theorem with Lagrange's Remainder). Suppose  $f$  and its derivatives up to order  $n - 1$  are continuous in  $[a, a + h]$ , and  $f^{(n)}$  exists for  $x \in (a, a + h)$ . Then

$$f(a + h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \dots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a) + \frac{h^n}{n!}f^{(n)}(a + \theta h)$$

where  $\theta \in (0, 1)$ .



Note that for  $n = 1$ , this is exactly the mean value theorem, so this can be seen as an  $n$ th order extension of the mean value theorem. We commonly write  $R_n$  for the final error term  $\frac{h^n}{n!} f^{(n)}(a + \theta h)$ . This is known as Lagrange's form of the remainder.

*Proof.* For  $0 \leq t \leq h$ , we define

$$\phi(t) = f(a+t) - f(a) - tf'(a) - \dots - \frac{t^{n-1}}{(n-1)!} f^{(n-1)}(a) - \frac{t^n}{n!} B$$

where we choose  $B$  suitably such that  $\phi(h) = 0$ . (Recall that in the proof of the mean value theorem, we used  $f(x) - kx$  and picked  $k$  suitably such that this allowed the use of Rolle's theorem. This is entirely analogous, but generalised to the  $n$ th derivative). Note that

$$\phi(0) = \phi'(0) = \dots = \phi^{(n-1)}(0) = 0$$

We can use Rolle's theorem inductively  $n$  times. Since  $\phi(0) = \phi(h) = 0$ , there is a point  $0 < h_1 < h$  such that  $\phi'(h_1) = 0$ . Since  $\phi'(0) = \phi'(h_1) = 0$ , there is a point  $0 < h_2 < h_1$  such that  $\phi''(h_2) = 0$ . This continues until we find a point  $0 < h_n < h$  such that  $\phi^{(n)}(h_n) = 0$ . Hence  $h_n = \theta h$  for some  $0 < \theta < 1$ . Now,  $\phi^{(n)}(t) = f^{(n)}(a+t) - B$ . We can see now that  $B = f^{(n)}(a + \theta h)$ , which gives the required result.  $\square$

We can prove an alternative version of Taylor's theorem with a different error term.

**Theorem** (Taylor's Theorem with Cauchy's Remainder). Suppose (equivalently to before)  $f$  and its derivatives up to order  $n-1$  are continuous in  $[a, a+h]$ , and  $f^{(n)}$  exists for  $x \in (a, a+h)$ . Then

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \dots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(a) + R_n$$

where

$$R_n = \frac{(1-\theta)^{n-1} h^n f^{(n)}(a+\theta h)}{(n-1)!}$$

for  $\theta \in (0, 1)$ .

*Proof.* For simplicity, in this proof we let  $a = 0$ , although the same argument applies when  $a \neq 0$ . Let us define

$$F(t) = f(h) - f(t) - (h-t)f'(t) - \dots - \frac{(h-t)^{n-1} f^{(n-1)}(t)}{(n-1)!}$$

for  $t \in [0, h]$ . Then

$$\begin{aligned} F'(t) &= -f'(t) + f'(t) - (h-t)f''(t) + (h-t)f''(t) - \frac{1}{2}(h-t)^2 f'''(t) + \frac{1}{2}(h-t)^2 f'''(t) \\ &\quad - \dots - \frac{(h-t)^{n-1}}{(n-1)!} f^{(n)}(t) \\ &= -\frac{(h-t)^{n-1}}{(n-1)!} f^{(n)}(t) \end{aligned}$$

Let

$$\phi(t) = F(t) - \left[ \frac{h-t}{h} \right]^p F(0)$$

where  $p \in \mathbb{N}$  and  $1 \leq p \leq n$ . Then

$$\phi(0) = \phi(h) = 0$$

By Rolle's theorem, there exists  $\theta \in (0, 1)$  such that

$$\phi'(\theta h) = 0$$

We can compute  $\phi'$  to find

$$\phi'(\theta h) = F'(\theta h) + \frac{p(1-\theta)^{p-1}}{h} F(0) = 0$$

Substituting everything back into  $F$  gives

$$0 = \frac{-h^{n-1}(1-\theta)^{n-1}}{(n-1)!} f^{(n)}(\theta h) + \frac{p(1-\theta)^{p-1}}{h} \left[ f(h) - f(0) - h'(0) - \dots - \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(0) \right]$$

Hence

$$f(h) = f(0) + hf'(0) + \frac{h^2}{2!} f''(0) + \dots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(0) + \underbrace{\frac{h^n(1-\theta)^{n-1} f^{(n)}(\theta h)}{(n-1)! \cdot p(1-\theta)^{p-1}}}_{R_n}$$

By letting  $p = n$ , we get Lagrange's remainder. If  $p = 1$ , we get Cauchy's remainder. □

## 8.2 Bounding error terms

Recall that Lagrange's remainder is

$$R_n = \frac{h^n}{n!} f^{(n)}(a + \theta h)$$

and Cauchy's remainder is

$$R_n = \frac{(1-\theta)^{n-1} h^n f^{(n)}(a + \theta h)}{(n-1)!}$$

and that we can write

$$f(h) = P_{n-1}(h) + R_n$$

where  $P_{n-1}$  is the Taylor polynomial to  $(n-1)$ th order. To get a Taylor series for a function  $f$ , we need to prove that the  $R_n$  tend to zero as  $n \rightarrow \infty$ . In general, this requires estimates for the  $R_n$  and it could take a lot of effort to prove whether this limit is zero or not. Note also that the theorems deducing the remainder terms work equally well in an interval  $[a + h, a]$  where  $h < 0$ .

## 8.3 Binomial series

**Proposition.** Let

$$f(x) = (1+x)^r$$

for some  $r \in \mathbb{Q}$ . If  $|x| < 1$ , then

$$f(x) = 1 + \binom{r}{1}x + \dots + \binom{r}{n}x^n + \dots$$

where

$$\binom{r}{n} = \frac{r(r-1)\cdots(r-n+1)}{n!}$$

*Proof.* Clearly,

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

These coefficients correspond exactly with that of the Taylor polynomial. If  $r \in \mathbb{N}$ , then  $f^{(r+1)}(x) \equiv 0$ , so clearly the  $R_n$  are zero as  $n \rightarrow \infty$ . In general, using Lagrange's form of the remainder,

$$R_n = \frac{x^n}{n!} f^{(n)}(\theta x) = \binom{r}{n} \frac{x^n}{(1+\theta x)^{n-r}}$$

Note that in principle,  $\theta$  depends both on  $x$  and  $n$ . For  $0 < x < 1$ ,  $(1+\theta x)^{n-r} > 1$  for  $n > r$ . Now observe that the series given by

$$\sum \binom{r}{n} x^n$$

is absolutely convergent for  $|x| < 1$ . Indeed, we can apply the ratio test and find that

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{(r-n)x}{n+1} \right|$$

which tends to  $|x|$  as  $n \rightarrow \infty$ . In particular therefore, the terms  $\binom{r}{n} x^n$  tend to zero for  $|x| < 1$ . Hence for  $n > r$  and  $0 < x < 1$ , we have

$$|R_n| \leq \left| \binom{r}{n} x^n \right| \rightarrow 0$$

So the claim is proven in the range  $0 \leq x < 1$ . If  $x < 0$ , then the step when we compare  $(1+\theta x)^{n-r}$  with 1 breaks down. Let us instead use the Cauchy form of the remainder to bypass this step.

$$R_n = \frac{(1-\theta)^{n-1} x^n f^{(n)}(\theta x)}{(n-1)!} = \frac{(1-\theta)^{n-1} r(r-1)\cdots(r-n+1)(1+\theta x)^{r-n} x^n}{(n-1)!}$$

By regrouping terms, we get

$$R_n = \frac{r(r-1)\cdots(r-n+1)}{(n-1)!} \cdot \frac{(1-\theta)^{n-1}}{(1+\theta x)^{n-r}} x^n = r \binom{r-1}{n-1} x^n (1+\theta x)^{r-1} \left( \frac{1-\theta}{1+\theta x} \right)^{n-1}$$

Hence

$$|R_n| \leq \left| r \binom{r-1}{n-1} x^n \right| (1+\theta x)^{r-1}$$

This will then tend to zero, after a bit more effort; we can bound the  $(1+\theta x)^{r-1}$  term by the maximum of 1 and  $(1+x)^{r-1}$ , which is independent of  $n$ , and then the result will follow.  $\square$

## 9 Power series

### 9.1 Complex differentiation

The complex derivative and the real derivative have the same core properties, for instance linearity, the product rule and the chain rule. However, the complex derivative is significantly more restrictive than the real derivative, since we can approach a point in any number of directions. If we can find a function that is complex differentiable with this restriction, we actually get a whole array of features for free. As an example of this restriction, consider the function  $f(z) = \bar{z}$ . This function is actually nowhere differentiable. If it were differentiable, then any sequence tending to  $z$  would yield the same limit when substituted into the definition of the derivative. Consider first the sequence

$$z_n = z + \frac{1}{n} \rightarrow z$$

Then

$$\frac{f(z_n) - f(z)}{z_n - z} = \frac{\bar{z} + \frac{1}{n} - \bar{z}}{z + \frac{1}{n} - z} = 1$$

Now consider the sequence

$$z_n = z + \frac{i}{n} \rightarrow z$$

Then

$$\frac{f(z_n) - f(z)}{z_n - z} = \frac{\bar{z} - \frac{i}{n} - \bar{z}}{z + \frac{i}{n} - z} = -1$$

Hence  $f(z)$  is nowhere differentiable. On the other hand, the real function  $f(x, y) = (x, -y)$  is clearly real differentiable, since it is linear; but in the complex world the function  $z \mapsto \bar{z}$  is not linear.

### 9.2 Definition of power series

A power series is a series of the form

$$\sum_{n=0}^{\infty} a_n z^n$$

where  $z \in \mathbb{C}$ , and the  $a_n$  is a given sequence of complex numbers. We can also take a power series of the form

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

but for simplicity we will take  $z_0 = 0$  in all of the analysis we will conduct on power series.

### 9.3 Radius of convergence

**Lemma.** If the series

$$\sum_{n=0}^{\infty} a_n z_1^n$$

converges for some point  $z_1$ , and  $|z| < |z_1|$ , then the series

$$\sum_{n=0}^{\infty} a_n z^n$$

also converges absolutely.

*Proof.* Since  $\sum_{n=0}^{\infty} a_n z_1^n$  converges,  $a_n z_1^n \rightarrow 0$ . Thus the sequence  $a_n z_1^n$  is bounded by some  $k > 0$ , i.e. for all  $n$ ,  $|a_n z_1^n| < k$ . Then

$$|a_n z^n| \leq k \left| \frac{z}{z_1} \right|^n$$

Since the geometric series  $\sum_0^{\infty} \left| \frac{z}{z_1} \right|^n$  converges, the lemma follows by comparison.  $\square$

Using this lemma, we can find that there exists a radius inside which any given power series converges absolutely. This radius might be zero, and it might be infinite.

**Theorem.** Any power series either

- (i) converges absolutely for all  $z$ , or
- (ii) converges absolutely for all  $z$  where  $|z| < R$  and diverges for all  $z$  where  $|z| > R$ , or
- (iii) converges for  $z = 0$  only.

The circle  $|z| = R$  is called the circle of convergence, and  $R$  is called the radius of convergence. Note that this theorem does not make any claim about the behaviour *on* the circle of convergence, just the behaviour inside it.

*Proof.* Let

$$S = \left\{ x \in \mathbb{R} : x \geq 0, \sum_0^{\infty} a_n x^n \text{ converges} \right\}$$

Clearly,  $0 \in S$ . By the above lemma, if  $x_1 \in S$ , then  $[0, x_1] \subseteq S$ . If  $S = [0, \infty)$ , then we have case (i) above due to the lemma.

If  $S \neq [0, \infty)$ , there exists a supremum  $0 \leq R = \sup S < \infty$ .

We must now just deal with case (ii), which is  $R > 0$ . For all  $z_1$  with  $|z_1| < R$  there exists  $R_0$  such that  $|z_1| < R_0 < R$ , and absolute convergence follows using the lemma. If  $|z_1| > R$ , there exists  $R_0$  such that  $|z_1| > R_0 > R$ . If the series with  $z_1$  converges, then by the lemma the same would be true for  $R_0$ . But  $R_0$  does not converge, so this is a contradiction.  $\square$

**Lemma.** If

$$\left| \frac{a_{n+1}}{a_n} \right| \rightarrow \ell$$

as  $n \rightarrow \infty$ , then  $R = \frac{1}{\ell}$ .

*Proof.* By the ratio test, we have absolute convergence if

$$\left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| < 1$$

So we have absolute convergence if  $|z| < \frac{1}{\ell}$  and divergence if  $|z| > \frac{1}{\ell}$  as required.  $\square$

**Lemma.** If

$$|a_n^{1/n}| \rightarrow \ell$$

as  $n \rightarrow \infty$ , then  $R = \frac{1}{\ell}$ .

This can be shown similarly using the root test.

**Example.** (i) Consider the series  $\sum_0^\infty \frac{z^n}{n!}$ . Using the ratio test, the series converges absolutely everywhere.

(ii) The geometric series  $\sum_0^\infty z^n$  gives  $R = 1$  by the ratio test. In this case,  $|z| = 1$  gives divergence.

(iii) The series  $\sum_0^\infty n!z^n$  has  $R = 0$ , which again can be seen using the ratio test.

(iv) Consider  $\sum_1^\infty \frac{z^n}{n}$ . This also has  $R = 1$  by the ratio test. Note that the series diverges for  $z = 1$  since we get the harmonic series. However, it converges when  $z = -1$  by the alternating series test. To work out the behaviour at other points on the circle of convergence, we could consider the series  $\sum_1^\infty \frac{z^n}{n}(1-z)$ , which converges exactly when the original series does. The partial sums are

$$\begin{aligned} S_N &= \sum_1^N \left[ \frac{z^n - z^{n+1}}{n} \right] \\ &= \sum_1^N \frac{z^n}{n} - \sum_1^N \frac{z^{n+1}}{n} \\ &= \sum_1^N \frac{z^n}{n} - \sum_2^{N+1} \frac{z^n}{n-1} \\ &= z - \frac{z^{N+1}}{N+1} + \sum_2^{N+1} \frac{-z^n}{n(n-1)} \end{aligned}$$

If  $|z| = 1$ , then the term  $\frac{z^{N+1}}{N+1}$  will vanish as  $N \rightarrow \infty$ . If  $z \neq 1$ , the term  $\sum_2^{N+1} \frac{-z^n}{n(n-1)}$  converges as  $N \rightarrow \infty$ . So  $S_N$  does indeed converge for  $|z| = 1, z \neq 1$ .

(v) Now, consider  $\sum_1^\infty \frac{z^n}{n^2}$ . This has  $R = 1$  by the ratio test, but it converges for all  $z$  with  $|z| = 1$ .

(vi) If we have  $\sum_0^\infty nz^n$ , we have  $R = 1$ , but diverges for all  $|z| = 1$ .

In conclusion, we cannot determine the behaviour at the boundary in the general case. Inside the radius of convergence, power series will behave as if they were simply polynomials.

## 9.4 Infinite differentiability

**Theorem.** Let  $f(z) = \sum_0^\infty a_n z^n$  have a radius of convergence  $R$ . Then  $f$  is complex differentiable at all points with  $|z| < R$ , with

$$f'(z) = \sum_1^\infty n a_n z^{n-1}$$

with the same radius of convergence as the original series.

This proof comprises the entire subsection. This whole subsection is non-examinable, but included for completeness. First, we will state two lemmas.

**Lemma.** If  $\sum_0^\infty a_n z^n$  has radius of convergence  $R$ , then both series

$$\sum_1^\infty n a_n z^{n-1}$$

and

$$\sum_2^\infty n(n-1) a_n z^{n-2}$$

also have radius of convergence  $R$ .

*Proof.* Let  $R_0$  be such that  $0 < |z| < R_0 < R$ . Since  $a_n R_0^n \rightarrow 0$ , the sequence  $a_n R_0^n$  is bounded. In other words there exists a  $k$  such that  $|a_n R_0^n| \leq k$  for all  $n \geq 0$ . Thus,

$$|a_n n z^{n-1}| = \frac{n}{|z|} |a_n R_0^n| \left| \frac{z}{R_0} \right|^n \leq \frac{kn}{|z|} \left| \frac{z}{R_0} \right|^n$$

But

$$\sum n \left| \frac{z}{R_0} \right|^n$$

converges by the ratio test, since the ratio is

$$\frac{n+1}{n} \left| \frac{z}{R_0} \right|^{n+1} \left| \frac{R_0}{z} \right|^n = \frac{n+1}{n} \left| \frac{z}{R_0} \right| \rightarrow \left| \frac{z}{R_0} \right| < 1$$

Hence, the original series  $\sum_1^\infty n a_n z^{n-1}$  is absolutely bounded above by a convergent series, and therefore is absolutely convergent. So it is known that the radius of convergence of this derivative series is *at least*  $R$ . Now, if  $|z| > R$ , the series diverges since  $|a_n z^n|$  is unbounded, and hence  $|n a_n z^{n-1}|$  is also unbounded. The same proof applies to the series for the second derivative.  $\square$

We will need this ‘second derivative’ condition in order to talk about the remainder term after the first derivative, which is related to the second derivative.

**Lemma.** First, for all  $2 \leq r \leq n$ .

$$\binom{n}{r} \leq n(n-1) \binom{n-2}{r-2}$$

Further, for all  $z \in \mathbb{C}, h \in \mathbb{C}$ ,

$$|(z+h)^n - z^n - nhz^{n-1}| \leq n(n-1)(|z| + |h|)^{n-2}|h|^2$$

*Proof.* For the first part, we can expand the definitions to get

$$\frac{\binom{n}{r}}{\binom{n-2}{r-2}} = \frac{n(n-1)}{r(r-1)} \leq n(n-1)$$

as required. For the second part, we can apply the binomial expansion to cancel the other two terms, and we get

$$\begin{aligned} (z+h)^n - z^n - nhz^{n-1} &= \left( \sum_{r=0}^n \binom{n}{r} z^{n-r} h^r \right) - z^n - nhz^{n-1} \\ &= \sum_{r=2}^n \binom{n}{r} z^{n-r} h^r \\ \therefore |(z+h)^n - z^n - nhz^{n-1}| &= \left| \sum_{r=2}^n \binom{n}{r} z^{n-r} h^r \right| \\ &\leq \sum_{r=2}^n \left| \binom{n}{r} z^{n-r} h^r \right| \\ &= \sum_{r=2}^n \binom{n}{r} |z|^{n-r} |h|^r \\ &\leq n(n-1) \underbrace{\left[ \sum_{r=2}^n \binom{n-2}{r-2} |z|^{n-r} |h|^{r-2} \right]}_{(|z|+|h|)^{n-2}} |h|^2 \\ &= n(n-1)(|z| + |h|)^{n-2}|h|^2 \end{aligned}$$

as required. □

Now, we can prove the original theorem.

*Proof.* By the first lemma, we may define  $f'(z)$  to be

$$f'(z) = \sum_1^{\infty} n a_n z^{n-1}$$



We now just need to prove that

$$\lim_{h \rightarrow 0} I = 0; \quad I = \frac{f(z+h) - f(z) - hf'(z)}{h}$$

We can substitute the expressions we have found for each power series:

$$\begin{aligned} I &= \frac{\sum_0^\infty a_n(z+h)^n - \sum_0^\infty a_n z^n - h \sum_1^\infty n a_n z^{n-1}}{h} \\ &= \frac{1}{h} \sum_0^\infty [a_n(z+h)^n - a_n z^n - h n a_n z^{n-1}] \\ &= \frac{1}{h} \sum_0^\infty a_n [(z+h)^n - z^n - h n z^{n-1}] \\ |I| &= \frac{1}{|h|} \left| \lim_{N \rightarrow \infty} \sum_0^N a_n [(z+h)^n - z^n - h n z^{n-1}] \right| \end{aligned}$$

Since the modulus function is continuous,

$$\begin{aligned} |I| &= \frac{1}{|h|} \lim_{N \rightarrow \infty} \left| \sum_0^N a_n [(z+h)^n - z^n - h n z^{n-1}] \right| \\ &\leq \frac{1}{|h|} \lim_{N \rightarrow \infty} \sum_0^N |a_n [(z+h)^n - z^n - h n z^{n-1}]| \\ &= \frac{1}{|h|} \sum_0^\infty |a_n| \cdot |(z+h)^n - z^n - h n z^{n-1}| \end{aligned}$$

By the second part of the second lemma above,

$$\begin{aligned} |I| &\leq \frac{1}{|h|} \sum_0^\infty |a_n| \cdot n(n-1)(|z| + |h|)^{n-2} |h|^2 \\ &= |h| \sum_0^\infty |a_n| \cdot n(n-1)(|z| + |h|)^{n-2} \end{aligned}$$

For  $|h|$  small enough,  $(|z| + |h|) < R$ . Therefore, by the first lemma above,

$$\sum_0^\infty |a_n| \cdot n(n-1)(|z| + |h|)^{n-2}$$

converges to some  $A(h)$ . But  $A(h)$  is monotonically decreasing, so

$$|I| \leq |h|A(h) \leq |h|A(r)$$

for some  $r$  such that  $|z| + r < R$ . We can now let  $h \rightarrow 0$ , giving

$$|I| \rightarrow 0$$

as required. □

## 9.5 Defining standard functions

We can now use this differentiability property to cleanly define the standard exponential, logarithmic and trigonometric functions. Let  $e : \mathbb{C} \rightarrow \mathbb{C}$  be defined by

$$e(z) = \sum_0^{\infty} \frac{z^n}{n!}$$

We have already seen that it has infinite radius of convergence. Straight from the above theorem,  $e$  is infinitely differentiable everywhere, and it is its own derivative. Note that if a function  $F : \mathbb{C} \rightarrow \mathbb{C}$  has  $F'(z) = 0$  for all  $z \in \mathbb{C}$ , then  $F$  is constant. Indeed, consider  $g(t) = F(tz) = u(t) + iv(t)$  where  $t, u, v \in \mathbb{R}$ . Then by the chain rule,  $g'(t) = F'(tz)z = 0$  and hence  $u'(t) + iv'(t) = 0$ , giving  $u'(t) = 0$  and  $v'(t) = 0$  everywhere. We can now apply the real-valued case, showing that  $u$  and  $v$  (and hence  $F$ ) are constant everywhere. Now, let  $a, b \in \mathbb{C}$ , and consider

$$F(z) = e(a + b - z)e(z)$$

Then

$$F'(z) = -e(a + b - z)e(z) + e(a + b - z)e'(z) = 0$$

Hence  $e(a + b - z)e(z)$  is constant for all  $z$ , hence

$$e(a + b - z)e(z) = e(a + b - 0)e(0) = e(a + b)$$

Since  $z$  is arbitrary, we can set  $z = b$  to recover the familiar relation

$$e(a + b - b)e(b) = e(a + b) \implies e(a)e(b) = e(a + b)$$

## 9.6 Exponential and logarithmic functions

Last lecture, we covered the power series form of the exponential function  $e : \mathbb{C} \rightarrow \mathbb{C}$ . Note that if we input a real number, the output is also real. Hence,  $e : \mathbb{R} \rightarrow \mathbb{R}$ . This restricted definition of the function has the following properties.

**Theorem.** (i)  $e : \mathbb{R} \rightarrow \mathbb{R}$  is everywhere differentiable, and  $e'(x) = e(x)$ .  
(ii)  $e(x + y) = e(x)e(y)$ .  
(iii)  $e(x) > 0$ .  
(iv)  $e$  is strictly increasing.  
(v)  $e(x) \rightarrow \infty$  as  $x \rightarrow \infty$ , and  $e(x) \rightarrow 0$  as  $x \rightarrow -\infty$ .  
(vi)  $e : \mathbb{R} \rightarrow (0, \infty)$  is a bijection.

*Proof.* The first two properties follow from the last lecture.

(iii) Clearly,  $e(x) > 0$  for all  $x \geq 0$  by considering the power series, which contains only positive terms for  $x > 0$ , and also  $e(0) = 1$ . Also,  $e(0) = e(x - x) = e(x)e(-x)$ , hence for all negative  $x$ ,  $e(x) > 0$ .

(iv) Since  $e'(x) = e(x)$ ,  $e'(x) = e(x) > 0$  everywhere.

(v) By considering partial sums, if  $x > 0$  we have  $e(x) > 1 + x$ , so if  $x \rightarrow \infty$ ,  $e(x) \rightarrow \infty$ . When  $x \rightarrow -\infty$ ,  $e(x) = \frac{1}{e(-x)} \rightarrow 0$ .

(vi) Injectivity follows from being strictly increasing. For surjectivity, we need to show that given any  $y \in (0, \infty)$  there exists some  $x$  such that  $e(x) = y$ . Due to property (v) above, we can certainly find real numbers  $a$  and  $b$  such that  $e(a) < y < e(b)$ . By the intermediate value theorem, there exists  $x \in \mathbb{R}$  such that  $e(x) = y$ .

□

*Remark.* We have essentially proven that  $e : (\mathbb{R}, +) \rightarrow ((0, \infty), \times)$  is a group isomorphism. This is exactly the same as showing that it is a bijection. Since  $e$  is a function, there exists an inverse function  $\ell : ((0, \infty), \times) \rightarrow (\mathbb{R}, +)$ .

**Theorem.** (i)  $\ell : (0, \infty) \rightarrow \mathbb{R}$  is a bijection, and  $\ell(e(x)) = x$  for all  $x \in \mathbb{R}$ , and  $e(\ell(x)) = x$  for all  $x \in (0, \infty)$ .  
(ii)  $\ell$  is differentiable and its derivative is  $\ell'(t) = \frac{1}{t}$ .  
(iii)  $\ell(xy) = \ell(x) + \ell(y)$ .

*Proof.* (i) This first property is obvious from the definition.

(ii) By the inverse function theorem,  $\ell$  is differentiable everywhere and  $\ell'(t) = \frac{1}{t}$  as required.

(iii) From IA Groups, if  $e$  is an isomorphism, so is its inverse.

□

## 9.7 Real numbered exponents

We will now define for  $\alpha \in \mathbb{R}$  and  $x > 0$  the function

$$r_\alpha(x) = e(\alpha\ell(x))$$

This can be taken as the definition of  $x$  raised to the power  $\alpha$ .

**Theorem.** Suppose  $x, y > 0$  and  $\alpha, \beta \in \mathbb{R}$ . Then

- (i)  $r_\alpha(xy) = r_\alpha(x)r_\alpha(y)$
- (ii)  $r_{\alpha+\beta}(x) = r_\alpha(x)r_\beta(x)$
- (iii)  $r_\alpha(r_\beta(x)) = r_{\alpha\beta}(x)$
- (iv)  $r_1(x) = x$ , and  $r_0(x) = 1$

*Proof.* (i)  $r_\alpha(xy) = e(\alpha\ell(xy)) = e(\alpha\ell(x) + \alpha\ell(y)) = e(\alpha\ell(x))e(\alpha\ell(y)) = r_\alpha(x)r_\alpha(y)$

(ii)  $r_{\alpha+\beta}(x) = e((\alpha + \beta)\ell(x)) = e(\alpha\ell(x) + \beta\ell(x)) = e(\alpha\ell(x))e(\beta\ell(x)) = r_\alpha(x)r_\beta(x)$

(iii)  $r_\alpha(r_\beta(x)) = e(\alpha\ell[e(\beta\ell(x))]) = e(\alpha\beta\ell(x)) = r_{\alpha\beta}(x)$

(iv)  $r_1(x) = e(\ell(x)) = x$ , and  $r_0(x) = e(0\ell(x)) = e(0) = 1$

□

Suppose we want to compute  $r_n(x)$ , where  $n \in \mathbb{Z}$ . Then  $r_n(x) = r_{1+\dots+1}(x) = x \cdots x$ , so we have agreement between  $r_n(x)$  and our previous definition of  $x^n$ . Similarly, since  $r_1(x)r_{-1}(x) = 1$ , we have

$r_{-1}(x) = \frac{1}{x}$ . Further,  $r_{\frac{1}{q}}(x) = x^{\frac{1}{q}}$ . Therefore,  $r_{\frac{p}{q}}(x) = x^{\frac{p}{q}}$ . So this definition is simply a more general definition for exponentiation by a real number.

From now, we will let  $\exp(x) \equiv e(x)$ ,  $\log(x) \equiv \ell(x)$ , and  $x^\alpha \equiv r_\alpha(x)$ . In fact,  $\exp(x) = e^x$  for a suitable number  $e$ , since  $e(x) = e(x \log(e)) = r_x(e) = e^x$  where  $e := e(1) = \sum_0^\infty \frac{1}{n!}$ .

Finally, we can compute the derivative of  $x^\alpha$  using the chain rule.

$$(x^\alpha)' = (e^{\alpha \log x})' = e^{\alpha \log x} \alpha \frac{1}{x} = \alpha x^\alpha x^{-1} = \alpha x^{\alpha-1}$$

as expected. Further, if  $f(x) = a^x$ , we can find

$$f'(x) = (e^{x \log a})' = e^{x \log a} \log a = a^x \log a$$

## 9.8 Trigonometric functions

We define

$$\begin{aligned} \cos z &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots = \sum_0^\infty \frac{(-1)^k z^{2k}}{(2k)!} \\ \sin z &= z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots = \sum_0^\infty \frac{(-1)^k z^{2k+1}}{(2k+1)!} \end{aligned}$$

Both power series have infinite radius of convergence, by the ratio test (the same proof from the exponential function can be used here). Hence  $\cos z$  and  $\sin z$  are differentiable, and  $\frac{d}{dz} \cos z = -\sin z$  and  $\frac{d}{dz} \sin z = \cos z$  as expected, by termwise differentiation. Further, we can deduce that

$$e^{iz} = \sum_0^\infty \frac{(iz)^n}{n!} = \sum_0^\infty \frac{(iz)^{2k}}{(2k)!} + \sum_0^\infty \frac{(iz)^{2k+1}}{(2k+1)!}$$

Note that

$$(iz)^{2k} = (-1)^k z^{2k}; \quad (iz)^{2k+1} = i(-1)^k z^{2k+1}$$

Hence,

$$e^{iz} = \cos z + i \sin z$$

Similarly,

$$e^{-iz} = \cos z - i \sin z$$

We can then write

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}); \quad \sin z = \frac{1}{2i}(e^{iz} - e^{-iz})$$

Many common trigonometric identities follow from this, such as the identity  $\cos^2 z + \sin^2 z \equiv 1$ . However, we have not deduced the period of the functions. Now, restricted to the real case,  $\sin x, \cos x \in \mathbb{R}$ , and the identity  $\cos^2 z + \sin^2 z \equiv 1$  gives that  $|\sin x| \leq 1$  and  $|\cos x| \leq 1$  for all real  $x$ .

## 9.9 Circle constants

**Proposition.** There is a smallest positive number  $\pi$  such that

$$\cos \frac{\pi}{2} = 0$$

and we have  $\sqrt{2} < \frac{\pi}{2} < \sqrt{3}$ .

*Proof.* If  $0 < x < 2$ ,

$$\sin x = \left(x - \frac{x^3}{3!}\right) + \left(\frac{x^5}{5!} - \frac{x^7}{7!}\right) + \dots$$

For this range of values, each parenthesised block is positive, so  $\sin x > 0$ . So in this range,

$$\frac{d}{dx} \cos x < 0$$

Hence,  $\cos x$  is a strictly decreasing function on this interval. Now,

$$\cos \sqrt{2} = 1 - \frac{\sqrt{2}^2}{2!} + \left(\frac{\sqrt{2}^4}{4!} - \frac{\sqrt{2}^6}{6!}\right) + \dots > 0$$

since each bracketed block is positive.

$$\cos \sqrt{3} = 1 - \frac{\sqrt{3}^2}{2!} + \frac{\sqrt{3}^4}{4!} - \left(\frac{\sqrt{3}^6}{6!} - \frac{\sqrt{3}^8}{8!}\right) + \dots < 0$$

since all the bracketed terms are positive, and being subtracted from a negative number. By the intermediate value theorem, the existence of such a  $\pi$  follows.  $\square$

**Corollary.** We have that  $\sin \frac{\pi}{2} = 1$ .

*Proof.* We know that  $\cos^2 \frac{\pi}{2} + \sin^2 \frac{\pi}{2} = 1$ , and  $\sin \frac{\pi}{2} > 0$ , so the result follows.  $\square$

**Theorem.** The following standard properties about the periodicity of trigonometric functions hold.

- (i)  $\sin\left(z + \frac{\pi}{2}\right) = \cos z$ , and  $\cos\left(z + \frac{\pi}{2}\right) = -\sin z$
- (ii)  $\sin(z + \pi) = -\sin z$ , and  $\cos(z + \pi) = -\cos z$
- (iii)  $\sin(z + 2\pi) = \sin z$ , and  $\cos(z + 2\pi) = \cos z$

The proofs are immediate from the angle addition formulae. This then implies that

$$e^{iz+2\pi i} = e^{iz}$$

Hence  $e^z$  is periodic with period  $2\pi i$ .

## 10 Integration

### 10.1 Geometry of trigonometric functions

Recall that given any two vectors  $\mathbf{x}$  and  $\mathbf{y}$  in  $\mathbb{R}^2$ , we can define the dot (scalar) product by

$$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2$$

By the Cauchy–Schwarz inequality, we have

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

where we define the Euclidean norm in the normal way. Thus, for  $\mathbf{x} \neq 0$ ,  $\mathbf{y} \neq 0$ , we have

$$-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1$$

We now define the angle between two vectors  $\mathbf{x}$  and  $\mathbf{y}$  as exactly the unique number  $\theta \in [0, \pi]$  such that

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

### 10.2 Hyperbolic functions

We define the functions  $\cosh$  and  $\sinh$  as follows.

$$\cosh z = \frac{1}{2}(e^z + e^{-z})$$

$$\sinh z = \frac{1}{2}(e^z - e^{-z})$$

Hence

$$\cosh z = \cos(iz); \quad \sinh z = -i \sin(iz)$$

We can then show that

$$\frac{d}{dz} \cosh z = \sinh z; \quad \frac{d}{dz} \sinh z = \cosh z$$

and further,

$$\cosh^2 z - \sinh^2 z \equiv 1$$

### 10.3 Defining the Riemann integral

**Definition.** A *dissection* or *partition*  $\mathcal{D}$  of  $[a, b]$  is a finite subset of  $[a, b]$  containing the end points  $a$  and  $b$ . We write

$$\mathcal{D} = \{x_0, x_1, \dots, x_n\}$$

with  $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ .

**Definition.** We define the *upper sum* of a bounded function  $f$  associated with a partition  $\mathcal{D}$

by

$$S(f, \mathcal{D}) = \sum_{j=1}^n (x_j - x_{j-1}) \sup_{x \in [x_{j-1}, x_j]} f(x)$$

The *lower sum* is defined similarly,

$$s(f, \mathcal{D}) = \sum_{j=1}^n (x_j - x_{j-1}) \inf_{x \in [x_{j-1}, x_j]} f(x)$$

Clearly then  $S \geq s$  for all  $\mathcal{D}$ .

**Lemma.** If  $\mathcal{D}$  and  $\mathcal{D}'$  are dissections with  $\mathcal{D}' \supseteq \mathcal{D}$  ( $\mathcal{D}'$  is a refinement of  $\mathcal{D}$ ), then

$$S(f, \mathcal{D}) \underset{(i)}{\geq} S(f, \mathcal{D}') \underset{(ii)}{\geq} s(f, \mathcal{D}') \underset{(iii)}{\geq} s(f, \mathcal{D})$$

*Proof.* Inequality (ii) is obvious, we have already shown this to be true. Now, suppose  $\mathcal{D}'$  contains a single extra point  $y$  compared to  $\mathcal{D}$ , where  $y \in (x_{r-1}, x_r)$ . Clearly,

$$\sup_{x \in [x_{r-1}, y]} f(x), \sup_{x \in [y, x_r]} f(x) \leq \sup_{x \in [x_{r-1}, x_r]} f(x)$$

Then

$$(x_r - x_{r-1}) \sup_{x \in [x_{r-1}, x_r]} f(x) \geq (y - x_{r-1}) \sup_{x \in [x_{r-1}, y]} f(x) + (x_r - y) \sup_{x \in [y, x_r]} f(x)$$

Hence,

$$S(f, \mathcal{D}) \geq S(f, \mathcal{D}')$$

The same proof holds for inequality (iii), and inductively we can show that this works for any amount of extra points.  $\square$

**Lemma.** If  $\mathcal{D}_1, \mathcal{D}_2$  are arbitrary dissections, then

$$S(f, \mathcal{D}_1) \geq S(f, \mathcal{D}_1 \cup \mathcal{D}_2) \geq s(f, \mathcal{D}_1 \cup \mathcal{D}_2) \geq s(f, \mathcal{D}_2)$$

In particular,  $S(f, \mathcal{D}_1) \geq s(f, \mathcal{D}_2)$ .

*Proof.* Let  $\mathcal{D}' = \mathcal{D}_1 \cup \mathcal{D}_2$ , which is a refinement of both  $\mathcal{D}_1$  and  $\mathcal{D}_2$ , and apply the previous lemma.  $\square$

**Definition.** The *upper integral* of  $f$  is

$$I^*(f) = \inf_{\mathcal{D}} S(f, \mathcal{D})$$

Note that such an integral always exists, since the upper sums are always bounded below by

an arbitrary lower sum. Hence the infimum does indeed exist and is finite. Similarly,

$$I_*(f) = \sup_{\mathcal{D}} s(f, \mathcal{D})$$

Then by the lemmas above,  $I^*(f) \geq I_*(f)$ , since  $S(f, \mathcal{D}_2) \geq s(f, \mathcal{D}_1)$  for arbitrary dissections  $\mathcal{D}_1$  and  $\mathcal{D}_2$ .

**Definition.** A bounded function  $f : [a, b] \rightarrow \mathbb{R}$  is (Riemann) integrable if  $I^*(f) = I_*(f)$ . If this equality holds, we write

$$\int_a^b f(x) dx = I^*(f) = I_*(f) = \int_a^b f$$

## 10.4 Determining integrability

**Theorem.** A function  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable if and only if given  $\varepsilon > 0$ , there exists  $\mathcal{D}$  such that

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$$

*Proof.* For every dissection  $\mathcal{D}$ , we have that  $0 \leq I^*(f) - I_*(f) \leq S(f, \mathcal{D}) - s(f, \mathcal{D})$ . If the given condition holds,  $0 \leq I^*(f) - I_*(f) \leq S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$  for all  $\varepsilon > 0$ . This immediately implies that  $f$  is Riemann integrable since the upper integral and the lower integral match. Conversely, if  $f$  is integrable, by the definition of the supremum and infimum, there are partitions  $\mathcal{D}_1$  and  $\mathcal{D}_2$  such that

$$\int_a^b f - \frac{\varepsilon}{2} = I_*(f) - \frac{\varepsilon}{2} < s(f, \mathcal{D}_1)$$

Also,

$$\int_a^b f + \frac{\varepsilon}{2} = I^*(f) + \frac{\varepsilon}{2} > S(f, \mathcal{D}_2)$$

From last lecture, we can use the fact that  $\mathcal{D}_1 \cup \mathcal{D}_2$  is a refinement of both  $\mathcal{D}_1$  and  $\mathcal{D}_2$  to show that

$$S(f, \mathcal{D}_1 \cup \mathcal{D}_2) - s(f, \mathcal{D}_1 \cup \mathcal{D}_2) \leq S(f, \mathcal{D}_2) - s(f, \mathcal{D}_1)$$

Now,

$$S(f, \mathcal{D}_2) - s(f, \mathcal{D}_1) < \int_a^b f + \frac{\varepsilon}{2} - \int_a^b f - \frac{\varepsilon}{2} = \varepsilon$$

as required. □

## 10.5 Monotonic and continuous functions

We can use this theorem to show that monotonic and continuous functions are integrable. Note that monotonic and continuous functions (defined on a closed interval) are always bounded.



**Theorem.** Suppose a function  $f : [a, b] \rightarrow \mathbb{R}$  is monotonic. Then  $f$  is integrable.

*Proof.* Suppose  $f$  is increasing. Then

$$\sup_{x \in [x_{j-1}, x_j]} f(x) = f(x_j)$$

and similarly

$$\inf_{x \in [x_{j-1}, x_j]} f(x) = f(x_{j-1})$$

Thus,

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) = \sum_{j=1}^n (x_j - x_{j-1}) [f(x_j) - f(x_{j-1})]$$

Let us choose the dissection

$$\mathcal{D} = \left\{ a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n} + \dots + b \right\}$$

giving

$$x_j = a + j \frac{b-a}{n}$$

for  $0 \leq j \leq n$ . In this case,

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) = \frac{b-a}{n} \sum_{j=1}^n [f(x_j) - f(x_{j-1})] = \frac{b-a}{n} [f(b) - f(a)] \rightarrow 0$$

so then using the above theorem,  $f$  is integrable.  $\square$

To prove an analogous result for continuous function, we must first prove the following lemma.

**Lemma** (Uniform Continuity). Suppose a function  $f : [a, b] \rightarrow \mathbb{R}$  is continuous. Then given  $\varepsilon > 0$ ,  $\exists \delta > 0$  such that if  $|x - y| < \delta$ , we have  $|f(x) - f(y)| < \varepsilon$ .

Note that in this lemma, we are saying that there exists such a  $\delta$  that works for *every* pair of points within  $\delta$ . The definition of continuity only provides a  $\delta$  that depends on  $x$ , so this is stronger than the definition of continuity, and this property does not hold for all continuous functions.

*Proof.* Suppose there does not exist such a  $\delta$ . Then there exists some  $\varepsilon > 0$  such that for all  $\delta > 0$  there exist  $x, y \in [a, b]$  such that  $|x - y| < \delta$  but  $|f(x) - f(y)| \geq \varepsilon$ . Let  $\delta = \frac{1}{n}$ . For this choice, we can find sequences  $x_n$  and  $y_n$  with  $|x_n - y_n| < \frac{1}{n}$  but  $|f(x_n) - f(y_n)| \geq \varepsilon$ . By the Bolzano–Weierstrass theorem, since we are working in a closed bounded interval, the  $x_n$  and  $y_n$  have convergent subsequences that tend to  $c$  and  $d$ . Then by the triangle inequality,

$$|y_{n_k} - c| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - c| \rightarrow 0$$

So  $c = d$ . But  $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$ , and by continuity as  $k \rightarrow \infty$ ,  $|f(c) - f(c)| \geq \varepsilon$  which is a contradiction.  $\square$

**Theorem.** Suppose a function  $f : [a, b] \rightarrow \mathbb{R}$  is continuous. Then  $f$  is integrable.

*Proof.* We know that given  $\varepsilon > 0$ , there exists  $\delta > 0$  such that  $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$ . So now, let

$$\mathcal{D} = \left\{ a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n} + \dots + b \right\}$$

where  $n$  is chosen large enough such that  $\frac{b-a}{n} < \delta$ . Then, for any  $x, y \in [x_{j-1}, x_j]$ , we have that  $|f(x) - f(y)| < \varepsilon$ . We can now write

$$\max_{x \in [x_{j-1}, x_j]} f(x) - \min_{x \in [x_{j-1}, x_j]} f(x) = f(p) - f(q) < \varepsilon$$

Therefore, the upper sums and lower sums differ by at most  $(b-a)\varepsilon$ . Hence,  $f$  is integrable.  $\square$

## 10.6 Complicated integrable functions

In principle, many functions that are not continuous or monotonic can be integrated using the Riemann integral. For example, the function  $f : [0, 1] \rightarrow \mathbb{R}$  defined by

$$f(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \in (0, 1] \text{ in its lowest form} \\ 0 & \text{otherwise} \end{cases}$$

is Riemann integrable. We know that  $s(f, \mathcal{D}) = 0$  for all  $\mathcal{D}$ , since any interval will contain irrational numbers. We will show that given  $\varepsilon > 0$ , there exists  $\mathcal{D}$  such that  $S(f, \mathcal{D}) < \varepsilon$ . If this is true, then this function  $f$  really is Riemann integrable, with  $\int f = 0$ . We will choose  $N \in \mathbb{N}$  such that  $\frac{1}{N} < \frac{\varepsilon}{2}$ . Then

$$S = \left\{ x \in [0, 1] : f(x) \geq \frac{1}{N} \right\} = \left\{ \frac{p}{q} : 1 \leq q \leq N, 1 \leq p \leq q \right\}$$

This set  $S$  is a finite set, hence

$$S = \{0, t_1, \dots, t_R\}; \quad 0 < t_1 < \dots < t_R = 1$$

Consider a dissection  $\mathcal{D}$  such that

- (1) Each  $t_k$  is in some interval  $[x_{j-1}, x_j]$ , and
- (2) For all  $k$ , the unique interval containing  $t_k$  has length at most  $\frac{\varepsilon}{2R}$ .

Such a dissection can certainly be constructed. Then, in any interval that does not contain a  $t_k$ ,  $f$  in this interval is less than  $\frac{1}{N}$ . In any interval that does contain a  $t_k$ ,  $f \geq \frac{1}{N}$  but  $f < 1$  everywhere. Since there are  $R$  such intervals, each of which with length  $\frac{\varepsilon}{2R}$ , we have

$$S(f, \mathcal{D}) \leq \frac{1}{N} + \frac{\varepsilon}{2} < \varepsilon$$

## 10.7 Properties of Riemann integral

Consider functions  $f$  and  $g$  which are bounded and integrable on  $[a, b]$ .

- (1) If  $f \leq g$  on  $[a, b]$ , then  $\int f \leq \int g$ .
- (2)  $f + g$  is integrable on  $[a, b]$ , and  $\int (f + g) = \int f + \int g$ .
- (3) For any constant  $k$ ,  $kf$  is integrable, and  $\int kf = k \int f$ .
- (4)  $|f|$  is integrable, and  $\int |f| \leq \int |f|$ .
- (5)  $fg$  is integrable.

*Proof.* We will see proofs for some of these properties.

- (1) If  $f \leq g$ , then

$$\int f = I^*(f) \leq S(f, \mathcal{D}) \leq S(g, \mathcal{D})$$

Hence,

$$\int f = I^*(f) \leq I^*(g) = \int g$$

- (2) We have

$$\sup_{[x_{j-1}, x_j]} (f + g) \leq \sup_{[x_{j-1}, x_j]} f + \sup_{[x_{j-1}, x_j]} g$$

Therefore,

$$S(f + g, \mathcal{D}) \leq S(f, \mathcal{D}) + S(g, \mathcal{D})$$

Now, consider two dissections  $\mathcal{D}_1, \mathcal{D}_2$ . Now,

$$I^*(f + g) \leq S(f + g, \mathcal{D}_1 \cup \mathcal{D}_2) \leq S(f, \mathcal{D}_1 \cup \mathcal{D}_2) + S(g, \mathcal{D}_1 \cup \mathcal{D}_2) \leq S(f, \mathcal{D}_1) + S(g, \mathcal{D}_2)$$

We can then fix  $\mathcal{D}_1$  and take the infimum over  $\mathcal{D}_2$  to get

$$I^*(f + g) \leq S(f, \mathcal{D}_1) + I^*(g)$$

Taking the infimum over  $\mathcal{D}_1$  gives

$$I^*(f + g) \leq I^*(f) + I^*(g) = \int f + \int g$$

A completely similar argument will show that

$$I_*(f + g) \geq \int f + \int g$$

Combining this,  $f + g$  must be integrable, since  $I^*(f + g) \geq I_*(f + g)$ . This integral is then exactly  $\int f + \int g$ .

- (4) Consider first  $f_+(x) = \max(f(x), 0)$ . We want to show that  $f_+$  is integrable. We can check that

$$\sup_{[x_{j-1}, x_j]} f_+ - \inf_{[x_{j-1}, x_j]} f_+ \leq \sup_{[x_{j-1}, x_j]} f - \sup_{[x_{j-1}, x_j]} f$$

We know that given  $\varepsilon > 0$ , there exists  $\mathcal{D}$  such that

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$$

Hence,

$$S(f_+, \mathcal{D}) - s(f_+, \mathcal{D}) \leq S(f, \mathcal{D}) - s(f, \mathcal{D}) < \varepsilon$$

Therefore  $f_+$  is integrable. But  $|f| = 2f_+ - f$ , hence  $|f|$  is integrable by properties (2) and (3). Since  $-|f| \leq f \leq |f|$ , we can use monotonicity from (1) to find that

$$\left| \int f \right| \leq \int |f|$$

as claimed.

(5) Let  $f$  be integrable and positive. Then we can check that

$$\sup_{[x_{j-1}, x_j]} f^2 = \left( \underbrace{\sup_{[x_{j-1}, x_j]} f}_{M_j} \right)^2; \quad \inf_{[x_{j-1}, x_j]} f^2 = \left( \underbrace{\inf_{[x_{j-1}, x_j]} f}_{m_j} \right)^2$$

Then,

$$\begin{aligned} S(f^2, \mathcal{D}) - s(f^2, \mathcal{D}) &= \sum_{j=1}^n (x_j - x_{j-1})(M_j^2 - m_j^2) \\ &= \sum_{j=1}^n (x_j - x_{j-1})(M_j - m_j)(M_j + m_j) \end{aligned}$$

The function  $f$  is bounded by some constant  $k$ , therefore the bracket  $(M_j + m_j)$  is bounded by  $2k$ , which gives

$$S(f^2, \mathcal{D}) - s(f^2, \mathcal{D}) \leq 2k(S(f, \mathcal{D}) - s(f, \mathcal{D}))$$

So  $f^2$  is integrable. Now, considering any  $f$ ,  $|f| \geq 0$  is a non-negative integrable function. Since  $f^2 = |f|^2$ , we deduce that  $f^2$  is integrable for any integrable  $f$ . Finally, for  $fg$ , note that

$$4fg = (f + g)^2 - (f - g)^2$$

The right hand side is integrable, so the left hand side is integrable. □

## 11 Fundamental theorem of calculus

### 11.1 Breaking an interval

Let  $f$  be integrable on  $[a, b]$ . If  $a < c < b$ , then  $f$  is integrable over  $[a, c]$  and  $[c, b]$ , with

$$\int_a^b f = \int_a^c f + \int_c^b f$$

Conversely, if  $f$  is integrable on  $[a, c]$  and  $[c, b]$ , then  $f$  is integrable over  $[a, b]$  and the same equality holds for the combination of the integrals.

*Proof.* We first make two observations. First, if  $\mathcal{D}_1$  is a dissection of  $[a, c]$  and  $\mathcal{D}_2$  is a dissection of  $[c, b]$ , then  $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$  is a dissection of  $[a, b]$ , and

$$S(f, \mathcal{D}_1 \cup \mathcal{D}_2) = S\left(f\Big|_{[a,c]}, \mathcal{D}_1\right) + S\left(f\Big|_{[c,b]}, \mathcal{D}_2\right) \quad (*)$$

Also, if  $\mathcal{D}$  is a dissection of  $[a, b]$ , then

$$S(f, \mathcal{D}) \geq S(f, \mathcal{D} \cup \{c\}) = S\left(f\Big|_{[a,c]}, \mathcal{D}_1\right) + S\left(f\Big|_{[c,b]}, \mathcal{D}_2\right) \quad (\dagger)$$

Now,

$$(*) \implies I^*(f) \leq I^*\left(f\Big|_{[a,c]}\right) + I^*\left(f\Big|_{[c,b]}\right)$$

Further,

$$(\dagger) \implies I^*(f) \geq I^*\left(f\Big|_{[a,c]}\right) + I^*\left(f\Big|_{[c,b]}\right)$$

Hence,

$$I^*(f) = I^*\left(f\Big|_{[a,c]}\right) + I^*\left(f\Big|_{[c,b]}\right)$$

This argument also applies for the lower integral, therefore

$$0 \leq I^*(f) - I_*(f) = \underbrace{I^*\left(f\Big|_{[a,c]}\right) - I_*\left(f\Big|_{[a,c]}\right)}_A + \underbrace{I^*\left(f\Big|_{[c,b]}\right) - I_*\left(f\Big|_{[c,b]}\right)}_B$$

Note that  $A, B \geq 0$ . If  $f$  is integrable on  $[a, c]$  and  $[c, b]$ , then  $A = B = 0$ , hence  $I^*(f) = I_*(f)$  and it is integrable on  $[a, b]$ . If  $f$  is integrable on  $[a, b]$ , then we know  $I^*(f) = I_*(f)$ , so  $A = B = 0$  so  $f$  is integrable on  $[a, c]$  and  $[c, b]$ .  $\square$

Note that we take the following convention:

$$\int_a^b f = - \int_b^a f$$

and if  $a = b$ , then this value is zero. With this convention, if  $f$  is bounded with  $|f| \leq k$ , then

$$\left| \int_a^b f \right| \leq k|b - a|$$

## 11.2 Fundamental theorem of calculus

Suppose a function  $f : [a, b] \rightarrow \mathbb{R}$  is bounded and integrable. Then since it is integrable on any sub-interval, we can define

$$F(x) = \int_a^x f(t) dt$$

for  $x \in [a, b]$ .

**Theorem.**  $F$  is continuous.

*Proof.* We know that

$$F(x+h) - F(x) = \int_x^{x+h} f(t) dt$$

We want this quantity to vanish as  $h \rightarrow 0$ . We find, given that  $f$  is bounded by  $k$ ,

$$|F(x+h) - F(x)| = \left| \int_x^{x+h} f(t) dt \right| \leq k|h|$$

So the result follows as  $h \rightarrow 0$ . □

**Theorem.** If in addition  $f$  is continuous at  $x$ , then  $F$  is differentiable, with  $F'(x) = f(x)$ .

*Proof.* Consider

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right|$$

If this tends to zero, then the theorem holds.

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = \frac{1}{|h|} \left| \int_x^{x+h} f(t) dt - hf(x) \right| = \frac{1}{|h|} \left| \int_x^{x+h} [f(t) - f(x)] dt \right|$$

Since  $f$  is continuous at  $x$ , given  $\varepsilon > 0$ ,  $\exists \delta > 0$  such that  $|t - x| < \delta \implies |f(t) - f(x)| < \varepsilon$ . If  $|h| < \delta$ , then the integrand is bounded by  $\varepsilon$ . Hence,

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| \leq \frac{1}{|h|} |h\varepsilon| = \varepsilon$$

So we can make this value as small as we like. So the theorem holds. □

For example, consider the function

$$f(x) = \begin{cases} -1 & x \in [-1, 0] \\ 1 & x \in (0, 1] \end{cases}$$

This is a bounded, integrable function, with

$$F(x) = -1 + |x|$$

Note that this  $F$  is not differentiable at  $x = 0$ .

**Corollary.** If  $f = g'$  is a continuous function on  $[a, b]$ , then

$$\int_a^x f(t) dt = g(x) - g(a)$$

is a differentiable function on  $[a, b]$ .

*Proof.* From above,  $F - g$  has zero derivative in  $[a, b]$ , hence  $F - g$  is constant. Since  $F(a) = 0$ , we get  $F(x) = g(x) - g(a)$ .  $\square$

Note that every continuous function  $f$  has an ‘indefinite’ integral (or ‘antiderivative’) written  $\int f(x) dx$ , which is determined uniquely up to the addition of a constant. Note further that we have now essentially solved the differential equation

$$\begin{cases} y'(x) = f(x) \\ y(a) = y_0 \end{cases}$$

and shown that there is a unique solution to this ordinary differential equation.

## 12 Integration techniques

### 12.1 Integration by parts

We can use the fundamental theorem of calculus to deduce familiar integration techniques, such as integration by parts, and integration by substitution.

**Corollary.** Suppose  $f', g'$  exist and are continuous on  $[a, b]$ . Then

$$\int_a^b f'g = fg \Big|_a^b - \int_a^b fg'$$

*Proof.* By the product rule, we have

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

Then by the fundamental theorem of calculus,

$$\int_a^b (fg)' = fg \Big|_a^b = \int_a^b f'g + \int_a^b fg'$$

and the result follows.  $\square$

### 12.2 Integration by substitution

**Corollary.** Let  $g : [\alpha, \beta] \rightarrow [a, b]$  with  $g(\alpha) = a, g(\beta) = b$  and let  $g'$  exist and be continuous on  $[\alpha, \beta]$ . Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous. Then

$$\int_a^b f(x) dx = \int_\alpha^\beta f(g(t))g'(t) dt$$

*Proof.* Let  $F(x) = \int_a^x f(t) dt$ . Then let  $h(t) = F(g(t))$ . This is well defined since  $g$  takes values in

$[a, b]$ . Then,

$$\begin{aligned}
 \int_{\alpha}^{\beta} f(g(t))g'(t) dt &= \int_{\alpha}^{\beta} F'(g(t))g'(t) dt \\
 &= \int_{\alpha}^{\beta} h'(t) dt \\
 &= h(\beta) - h(\alpha) \\
 &= F(b) - F(a) \\
 &= F(b) \\
 &= \int_a^b f(x) dx
 \end{aligned}$$

□

## 13 Integrals in Taylor's theorem

### 13.1 Integral remainder form of Taylor's theorem

**Theorem.** Let  $f$  such that  $f^{(n)}(x)$  is continuous for  $x \in [0, h]$ . Then

$$f(h) = f(0) + \dots + \frac{h^{n-1}f^{(n-1)}(0)}{(n-1)!} + R_n$$

where

$$R_n = \frac{h^n}{(n-1)!} \int_0^1 (1-t)^{n-1} f^{(n)}(th) dt$$

Note that for this formulation of Taylor's theorem, we require continuity of  $f^{(n)}(x)$ , whereas with the previous remainders, the  $n$ th derivative need not be continuous.

*Proof.* First, by substituting  $u = th$ , we can see that it is sufficient to show

$$R_n = \frac{1}{(n-1)!} \int_0^h (h-u)^{n-1} f^{(n)}(u) du$$

Now, integrating by parts, we have

$$\begin{aligned}
 R_n &= \frac{-h^{n-1}f^{(n-1)}(0)}{(n-1)!} + \frac{1}{(n-2)!} \int_0^h (h-u)^{n-2} f^{(n-1)}(u) du \\
 &= \frac{-h^{n-1}f^{(n-1)}(0)}{(n-1)!} + R_{n-1}
 \end{aligned}$$

Hence,

$$R_n = -\frac{h^{n-1}f^{(n-1)}(0)}{(n-1)!} - \frac{h^{n-2}f^{(n-2)}(0)}{(n-2)!} - \dots - \underbrace{\int_0^h f'(u) du}_{f(h)-f(0)}$$

which is exactly all the other terms in the Taylor polynomial as required. □



## 13.2 Mean value theorem for integrals

**Theorem.** Let  $f, g : [a, b] \rightarrow \mathbb{R}$  be continuous with  $g(x) \neq 0$  for all  $x \in (a, b)$ . Then

$$\exists c \in (a, b) \text{ s.t. } \int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx$$

Note that if we let  $g(x) = 1$ , we get

$$\int_a^b f(x) dx = f(c)(b - a)$$

*Proof.* We will use Cauchy's mean value theorem to get this result. Let

$$F(x) = \int_a^x fg; \quad G(x) = \int_a^x g$$

Then there exists an intermediate point  $c$  such that

$$(F(b) - F(a))G'(c) = F'(c)(G(b) - G(a))$$

By the fundamental theorem of calculus,

$$\left( \int_a^b fg \right) g(c) = f(c)g(c) \left( \int_a^b g \right)$$

Now, since  $g \neq 0$  everywhere,

$$\int_a^b fg = f(c) \int_a^b g$$

□

## 13.3 Deriving Lagrange's and Cauchy's remainders for Taylor's theorem

We can use this new mean value theorem to recover the other forms of the remainders in Taylor's theorem. We have

$$R_n = \frac{h^n}{(n-1)!} \int_0^1 (1-t)^{n-1} f^{(n)}(th) dt$$

and we want to show that this is equal to

$$\frac{h^n}{n!} f^{(n)}(a + \theta h); \quad \frac{(1-\theta)^{n-1} h^n f^{(n)}(a + \theta h)}{(n-1)!}$$

First, let us apply the above mean value theorem with  $g \equiv 1$  and the entire integrand in  $R_n$  as  $f$ . Then

$$R_n = \frac{h^n}{(n-1)!} \int_0^1 (1-t)^{n-1} f^{(n)}(th) dt = \frac{h^n}{(n-1)!} (1-\theta)^{n-1} f^{(n)}(\theta h)$$

as required for Cauchy's remainder. To find Lagrange's remainder, we need to use the above mean value theorem with  $g = (1-t)^{n-1}$ , which is positive everywhere in  $(0, 1)$ , and  $f = f^{(n)}(th)$ . Then

$$R_n = \frac{h^n}{(n-1)!} f^{(n)}(\theta h) \int_0^1 (1-t)^{n-1} dt$$

This integral is simple to find by inspection:

$$R_n = \frac{h^n}{(n-1)!} f^{(n)}(\theta h) \frac{1}{n} = \frac{h^n}{n!} f^{(n)}(\theta h)$$

as required.

## 14 Uses of integration

### 14.1 Improper integration

**Definition.** Suppose  $f : [a, \infty) \rightarrow \mathbb{R}$  is integrable (and therefore bounded) on every interval of the form  $[a, R]$ , and further, as  $R \rightarrow \infty$ , we have  $\int_a^R f \rightarrow \ell$ . Then we say that the integral  $\int_a^\infty f$  exists (or converges), and its value is  $\ell$ . If this integral does not tend to a limit, we say that  $\int_a^\infty f$  diverges.

We can similarly define the integral  $\int_{-\infty}^a f$ . If  $\int_a^\infty f = \ell_1$  and  $\int_{-\infty}^a f = \ell_2$ , we can write

$$\int_{-\infty}^\infty f = \ell_1 + \ell_2$$

Note that this last condition is *not* the same as saying that  $\lim_{R \rightarrow \infty} \int_{-R}^R f$  exists. For this two-sided improper integral to exist, we need the stronger condition that the one-sided improper integrals exist on either side. For example, consider  $f(x) = x$ . Clearly  $\int_{-R}^R f = 0$ , but this function is not improper integrable. For example, consider

$$\int_1^\infty \frac{dx}{x^k}$$

This converges if and only if  $k > 1$ . Indeed, if  $k \neq 1$ ,

$$\int_1^R \frac{dx}{x^k} = \frac{x^{1-k}}{1-k} \Big|_1^R = \frac{R^{1-k} - 1}{1-k}$$

which is clearly finite in the limit if and only if  $k > 1$ . If  $k = 1$ , then we can find

$$\int_1^R \frac{dx}{x} = \log R \rightarrow \infty$$

as expected. Note the following observations.

- (1)  $\frac{1}{\sqrt{x}}$  is continuous (and bounded) on  $[\delta, 1]$  for all  $\delta > 0$ , and

$$\int_\delta^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x} \Big|_\delta^1 = 2 - 2\sqrt{\delta}$$

So as  $\delta \rightarrow 0$ , this integral tends to 2. This integral is defined, even though the value of the function at zero is unbounded. So we commonly write

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2$$

(2) Similarly, we write

$$\int_0^1 \frac{dx}{x} = \lim_{\delta \rightarrow 0} \int_{\delta}^1 \frac{dx}{x} = \lim_{\delta \rightarrow 0} \log x \Big|_{\delta}^1 = \lim_{\delta \rightarrow 0} (\log 1 - \log \delta)$$

Since this limit does not exist, the integral  $\int_0^1 \frac{dx}{x}$  does not exist.

(3) If  $f \geq 0$  and  $g \geq 0$  for  $x \geq a$ , and  $f(x) \leq kg(x)$ , where  $k$  is a constant for  $x \geq a$ , then

$$\int_a^{\infty} g \text{ converges} \implies \int_a^{\infty} f \text{ converges, and } \int_a^{\infty} f \leq \int_a^{\infty} g$$

This is similar to the comparison test for series. First, note that  $\int_a^R f \leq k \int_a^R g$ . Further,  $\int_a^R f$  is an increasing function of  $R$  since  $f \geq 0$ , and bounded above, since  $\int_a^{\infty} g$  converges. Let

$$\ell = \sup_{R \geq a} \int_a^R f < \infty$$

Then we want to show that  $\lim_{R \rightarrow \infty} \int_a^R f = \ell$ . Given  $\varepsilon > 0$ , by the definition of the supremum  $\exists R_0$  such that

$$\int_a^{R_0} f \geq \ell - \varepsilon$$

Thus for all  $R \geq R_0$  we have

$$\int_a^R f \geq \int_a^{R_0} f \geq \ell - \varepsilon$$

Hence,

$$0 \leq \ell - \int_a^R f \leq \varepsilon$$

As an example, consider the integral

$$\int_0^{\infty} \exp\left(-\frac{x^2}{2}\right) dx$$

Now, for  $x \geq 1$ , we can bound the integrand by  $\exp\left(-\frac{x}{2}\right)$ , and the integral of this bound is clearly bounded. Hence the original integral converges.

(4) If  $\sum a_n$  converges, then  $a_n \rightarrow 0$ . However, with improper integrals, this is not necessarily the case. Consider a convergent series  $a_n$ , where  $0 < a_n < 1$  for all  $n$ . Then define the function  $f$  defined by

$$f(n+r) = \begin{cases} 1 & \text{if } r < a_n \\ 0 & \text{otherwise} \end{cases}$$

where the input  $x$  is split into the integer part  $n$  and the remainder  $r$ . This function is essentially a sequence of rectangles of height 1 and width  $a_n$ , spaced so that each rectangle starts at an integer value of  $x$ . Clearly, we have

$$\int_0^n f = \sum_0^n a_n$$

where  $n$  is an integer. So the integral converges, but the integrand does not tend to zero.

## 14.2 Integral test for series convergence

**Theorem.** Let  $f(x)$  be a positive decreasing function for  $x \geq 1$ . Then,

- (1) The integral  $\int_1^\infty f(x) dx$  and the series  $\sum_1^\infty f(x)$  both converge or diverge. (Note that such a function is always Riemann integrable on a closed interval since it is bounded and decreasing.)
- (2) As  $n \rightarrow \infty$ ,  $\sum_{r=1}^n f(r) - \int_1^n f(x) dx$  tends to a limit  $\ell$  such that  $0 \leq \ell \leq f(1)$ .

*Proof.* If  $n-1 \leq x \leq n$ , then

$$f(n-1) \geq f(x) \geq f(n)$$

Hence,

$$f(n-1) \geq \int_{n-1}^n f(x) dx \geq f(n)$$

Adding up such integrals, we get

$$\sum_1^{n-1} f(r) \geq \int_1^n f(x) dx \geq \sum_2^n f(r)$$

Then the first claim is obvious. For the second claim, let

$$\phi(n) = \sum_1^n f(r) - \int_1^n f(x) dx$$

Then, using the inequalities established above,

$$\phi(n) - \phi(n-1) = f(n) - \int_{n-1}^n f(x) dx \leq 0$$

So  $\phi$  is a decreasing sequence. Further,

$$0 \leq \phi(n) \leq f(1)$$

$\phi$  is bounded, so it converges to some limit  $\ell$ . □

**Example.** First, consider the sum  $\sum_1^\infty \frac{1}{n^k}$ . By the integral test, this converges if and only if  $k > 1$ . As a more complicated example, consider  $\sum_2^\infty \frac{1}{n \log n}$ . Let  $f(x) = \frac{1}{x \log x}$ , and

$$\int_2^R \frac{dx}{x \log x} = \log(\log x) \Big|_2^R$$

which diverges, so by the integral test the series diverges.

**Corollary** (Euler–Mascheroni Constant). As  $n \rightarrow \infty$ ,

$$\sum_1^n \frac{1}{n} - \int_1^n \frac{1}{n} = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n \rightarrow \gamma$$

where  $\gamma \in [0, 1]$ . This is known as the Euler–Mascheroni constant. It is unknown whether  $\gamma$  is irrational.

### 14.3 Piecewise continuous functions

**Definition.** A function  $f : [a, b] \rightarrow \mathbb{R}$  is piecewise continuous if there is a dissection  $\mathcal{D}$  such that  $f$  is continuous on all intervals defined by this dissection, and that the one-sided limits

$$\lim_{x \rightarrow x_{j-1}^+} f(x); \quad \lim_{x \rightarrow x_{j-1}^-} f(x)$$

exist.

We can extend the class of Riemann integrable functions to include piecewise continuous functions as well. This is true since we use this dissection to construct the upper and lower sums. The one-sided limits are here to ensure that the function is bounded near these discontinuities. We might now ask how large the discontinuity set is allowed to be in order for  $f$  to still be Riemann integrable. As we have seen from examples before, it is possible to have a function which has countably many discontinuity points, but is still Riemann integrable.