Part IA Analysis I

András Zsák

Lent 2018

Contents

1 Real and Complex Numbers

We assume properties of the sets $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$ of natural numbers, integers and rational numbers, respectively.

$$\begin{aligned} \mathbb{N} &= \{1, 2, 3, \dots\} \\ \mathbb{Z} &= \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\} \\ \mathbb{Q} &= \left\{\tfrac{a}{b} : a, b \in \mathbb{Z},\ b \neq 0\right\} = \left\{\tfrac{a}{b} : a \in \mathbb{Z},\ b \in \mathbb{N}\right\} \end{aligned}$$

In $\mathbb{N}$ every non-empty subset has a least element. This is equivalent to induction. In $\mathbb{Q}$ we have operations of addition and multiplication that satisfy familiar properties. The set $\mathbb{R}$ of real numbers will be defined as the unique ordered field in which the least upper bound axiom holds.

Definition. A *field* is a set F together with two binary operations: addition (denoted by $+$) and multiplication (denoted by $\cdot$ or simply by juxtaposition) that satisfy the following.

(i) $(F, +)$ is an abelian group with identity denoted by 0,

(ii) $(F \setminus \{0\}, \cdot)$ is an abelian group,

(iii) $x(y + z) = xy + xz$ for all $x, y, z \in F$ (distributive law).

Some notation. 1. The identity element of $(F \setminus \{0\}, \cdot)$ is denoted by 1. We sometimes write $0_F, 1_F$ instead of $0, 1$ to emphasize the field.

2. $-x$ is the inverse of x in $(F, +)$.

3. x^{-1} is the inverse of $x \neq 0$ in $(F \setminus \{0\}, \cdot)$.

4. $x - y = x + (-y)$ $(x, y \in F)$ and $\frac{x}{y} = x \cdot y^{-1}$ $(x, y \in F,\ y \neq 0)$.

Examples. 1. $\mathbb{Q}$, $\mathbb{Z}/p\mathbb{Z}$ for a prime p, the set $\mathbb{Q}(X)$ of rational functions with coefficients in $\mathbb{Q}$:

$$\mathbb{Q}(X) = \left\{ \tfrac{p}{q} : p, q \text{ are polynomials with coefficients in } \mathbb{Q},\ q \neq 0 \right\} .$$

2. $\mathbb{Z}$ or the set $\mathbb{Q}[X]$ of polynomials over $\mathbb{Q}$ are not fields.

Some properties. In a field F we have

$$0 \cdot x = 0\ , \quad (-1)x = -x\ , \quad (-x)y = x(-y) = -(xy)$$

for all $x, y \in F$.

Definition. An *ordered field* is a pair $(F, <)$, where F is a field and $<$ is a total (or linear) order on F such that for all $x, y, z \in F$

(i) if $x < y$ then $x + z < y + z$,

(ii) if $x < y$ and $0 < z$ then $xz < yz$.

Examples. 1. $\mathbb{Q}$ is an ordered field with order

$$\frac{p}{q} < \frac{r}{s} \iff ps < qr \qquad (p, r \in \mathbb{Z},\ q, s \in \mathbb{N})\ .$$

2. $\mathbb{Q}(X)$ is an ordered field with $f < g$ if and only if for some $t_0 \in \mathbb{Q}$ we have $f(t) < g(t)$ for all $t > t_0$.

3. For a prime p the field $\mathbb{Z}/p\mathbb{Z}$ cannot be an ordered field under any ordering.

Properties (for an ordered field F). 1. $0_F < 1_F$ (this requires proof!)

2. F contains $\mathbb{Q}$. More precisely, there is a map $\theta \colon \mathbb{Q} \to F$ such that

(i) $\theta(1) = 1_F$,

(ii) $\theta(x + y) = \theta(x) + \theta(y)$ for all $x, y \in \mathbb{Q}$,

(iii) $\theta(xy) = \theta(x)\theta(y)$ for all $x, y \in \mathbb{Q}$.

Moreover, this map is unique and satisfies

$$x < y \iff \theta(x) < \theta(y) \qquad \text{for all } x, y \in \mathbb{Q}\ .$$

We identify $\mathbb{Q}$ with its image $\theta(\mathbb{Q})$ in F, so we can now write $\mathbb{Q} \subset F$.

Notation. In an ordered field we write $x \leqslant y$ to mean $x < y$ or $x = y$. Then for all $x, y, z \in F$

(i) if $x \leqslant y$ then $x + z \leqslant y + z$,

(ii) if $x \leqslant y$ and $0 \leqslant z$ then $xz \leqslant yz$.

Definitions. Let F be an ordered field and $A \subset F$.

We say A is *bounded above* if there exists $x \in F$ such that $a \leqslant x$ for all $a \in A$. Such an element x is called an *upper bound for* A, and we say A is *bounded above by* x.

A *supremum for* A is an upper bound x for A such that if y is any upper bound for A then $x \leqslant y$. If such a supremum for A exists then it is unique and we denote it by $\sup A$.

Note. The empty set is bounded above by all $x \in F$, and hence it has no supremum.

Definition. Let F be an ordered field. The *least upper bound axiom for* F is the statement: every non-empty bounded above subset of F has a supremum.

Theorem (not proven). There is a unique ordered field, denoted by $\mathbb{R}$, in which the least upper bound axiom holds. Elements of $\mathbb{R}$ are called *real numbers*.

Examples. We have already seen that $\mathbb{Q} \subset \mathbb{R}$. What else is in $\mathbb{R}$? In the course Numbers & Sets you proved that there exists $x \in \mathbb{R}$ such that $x^2 = 2$, and moreover $x \notin \mathbb{Q}$.

More generally, for any $a \geqslant 0$ there is a unique $x \geqslant 0$ such that $x^2 = a$ which we denote by $\sqrt{a}$. Thus the polynomial $p(x) = x^2 - a$ has a root in $\mathbb{R}$.

Note. It is easy to see that $\mathbb{N}$ is not bounded above in $\mathbb{Q}$. However, the following is *not* obvious.

Theorem 1. $\mathbb{N}$ is not bounded above in $\mathbb{R}$.

Note. There are ordered fields in which $\mathbb{N}$ *is* bounded above. *E.g.,* in $\mathbb{Q}(X)$.

Definitions. Let F be an ordered field and $A \subset F$.

We say A is *bounded below* if there exists $x \in F$ such that $x \leqslant a$ for all $a \in A$. Such an element x is called a *lower bound for* A, and we say A is *bounded below by* x.

An *infimum for* A is a lower bound x for A such that if y is any lower bound for A then $y \leqslant x$. If such an infimum for A exists then it is unique and we denote it by $\inf A$.

Note. The empty set is bounded below by all $x \in F$, and hence it has no infimum.

Definition. Let F be an ordered field. The *greatest lower bound axiom for* F is the statement: every non-empty bounded below subset of F has an infimum.

Proposition 2. For an ordered field F TFAE.

(i) The least upper bound axiom holds in F.

(ii) The greatest lower bound axiom holds in F.

Definition. Let F be an ordered field and $A \subset F$. We say A is *bounded* if it is bounded above and below, *i.e.,* there exist $u, v \in F$ such that $u \leqslant a \leqslant v$ for all $a \in A$.

Modulus. The *modulus* $|x|$ of an element x of an ordered field F is defined as

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

Properties. 1. $|x| \leqslant y$ if and only if $-y \leqslant x \leqslant y$.

2. $|xy| = |x| \cdot |y|$

3. $|x + y| \leqslant |x| + |y|$ (triangle inequality)

4. $\big||x| - |y|\big| \leqslant |x - y|$

Note. A subset A of an ordered field F is bounded if and only if there exists $w \in F$ such that $|a| \leqslant w$ for all $a \in A$.

Note. For any $x \in F$ we have $x^2 \geqslant 0$. Thus the equation $x^2 + 1 = 0$ has no solution in any ordered field, in particular in $\mathbb{R}$. This is remedied as follows.

Complex numbers. Let $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\}$ and define two binary operations on $\mathbb{C}$ as follows. For $x, y, u, v \in \mathbb{R}$ set

$$(x, y) + (u, v) = (x + u, y + v) \qquad \text{and} \qquad (x, y) \cdot (u, v) = (xu - yv, xv + yu) \ .$$

A routine verification shows that these operations turn $\mathbb{C}$ into a field.

Note that $\mathbb{R} \subset \mathbb{C}$. More precisely, $\{(x, 0) : x \in \mathbb{R}\}$ is a copy of $\mathbb{R}$ in $\mathbb{C}$.

Let $\mathrm{i} = (0, 1)$ and observe that $\mathrm{i}^2 = (-1, 0)$. Identifying $\mathbb{R}$ with its copy in $\mathbb{C}$, we have $\mathrm{i}^2 = -1$ and if $z = (x, y) \in \mathbb{C}$ then

$$z = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0) = x + \mathrm{i}y \ .$$

We let $\mathrm{Re}z = x$ (*real part of* z) and $\mathrm{Im}z = y$ (*imaginary part of* z).

We now have $\mathbb{C} = \{x + \mathrm{i}y : x, y \in \mathbb{R}\}$.

Note. $\mathbb{C}$ cannot be made into an ordered field.

Modulus. For $z = x + \mathrm{i}y \in \mathbb{C}$ $(x, y \in \mathbb{R})$ we define the modulus $|z|$ of z by $|z| = \sqrt{x^2 + y^2}$. This is a well-defined non-negative number. If $z \in \mathbb{R}$, then this coincides with the previous definition of $|z|$.

Properties. For any $z, w \in \mathbb{C}$ we have

$$|z+w| \leqslant |z|+|w|, \ \big||z|-|w|\big| \leqslant |z-w| \ , |zw| = |z|\cdot|w| \ , |\mathrm{Re}z| \leqslant |z| \text{ and } |\mathrm{Im}z| \leqslant |z| \ .$$

Definition. We say $A \subset \mathbb{C}$ is *bounded* if there exists $C \in \mathbb{R}$ with $|z| \leqslant C$ for all $z \in A$, *i.e.,* if the subset $\{|z| : z \in A\}$ of $\mathbb{R}$ is bounded.

Remark. The Fundamental Theorem of Algebra states that every non-constant polynomial over $\mathbb{C}$ has a root in $\mathbb{C}$.

2 Sequences and Series

Definition. Informally, a sequence is a list $x_1, x_2, x_3, \dots$ of elements of some set. Formally, a *sequence* is a function x with domain $\mathbb{N}$. We often write x_n instead of $x(n)$ for $n \in \mathbb{N}$.

Notation. $(x_n)_{n\in\mathbb{N}}$, or $(x_n)_{n=1}^{\infty}$, or simply (x_n).

If a sequence x takes values in a set S, then x is a *sequence in S*, *i.e.,* a function $x\colon \mathbb{N} \to S$.

A *real sequence* is a sequence in $\mathbb{R}$, *i.e.,* a function $x\colon \mathbb{N} \to \mathbb{R}$.

A *complex sequence* is a sequence in $\mathbb{C}$, *i.e.,* a function $x\colon \mathbb{N} \to \mathbb{C}$.

Examples. 1. $x_n = 2n^2 - n$ defines a sequence in $\mathbb{N}$.

2. $x_n = \frac{1}{n}$ and $y_n = \left(1 + \frac{1}{n}\right)^n$ define sequences in $\mathbb{Q}$.

3. $x_n = \sqrt{n}$ defines a sequence in $\mathbb{R}$.

4. $x_n = \mathrm{e}^{\mathrm{i}n\theta}$ defines a sequence in $\mathbb{C}$ where θ is a fixed real number.

Definition. Let (x_n) be a real (or complex) sequence and let $\lambda \in \mathbb{R}$ (or, respectively, $\lambda \in \mathbb{C}$). We say x_n *converges to* λ (written $x_n \to \lambda$ as $n \to \infty$) if given any positive real number ε, there exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, if $n \geqslant N$ then $|x_n - \lambda| < \varepsilon$. Formally with symbols:

$$(\forall\, \varepsilon > 0)\ (\exists\, N \in \mathbb{N})\ (\forall\, n \in \mathbb{N})\ (n \geqslant N \implies |x_n - \lambda| < \varepsilon) \tag{1}$$

Picture. In $\mathbb{R}$ we have $|y - \lambda| < \varepsilon$ if and only if $\lambda - \varepsilon < y < \lambda + \varepsilon$. It follows that (x_n) converges to λ if and only if (x_n) is eventually in every *open* interval $(a, b) = \{t \in \mathbb{R} :\ a < t < b\}$ containing λ. In $\mathbb{C}$ we let

$$D(\lambda, \varepsilon) = \{z \in \mathbb{C} :\ |z - \lambda| < \varepsilon\}$$

be the *open disc of centre λ and radius ε*. So (x_n) converges to λ if and only if (x_n) is eventually in every open disc with centre λ.

Another interpretation. For large n we have $x_n = \lambda + (x_n - \lambda) \approx \lambda$. Thus, we think of x_n as an approximation to λ, the error of approximation $x_n - \lambda$ being small for large n.

Definition. A real or complex sequence (x_n) is said to be *convergent* if there exists $\lambda \in \mathbb{R}$ (respectively, $\lambda \in \mathbb{C}$) such that $x_n \to \lambda$ as $n \to \infty$. In this case we say λ is *a limit of* (x_n).

Note. From now on if we don't specify whether we work in $\mathbb{R}$ or $\mathbb{C}$, then the result or definition under consideration is valid in both fields. Since $\mathbb{R} \subset \mathbb{C}$, a real sequence (x_n) can also be thought of as a sequence in $\mathbb{C}$. However, if $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then (x_n) cannot converge to λ.

Lemma 1 (Uniqueness of limit). If $x_n \to \lambda$ and $x_n \to \mu$ as $n \to \infty$, then $\lambda = \mu$.

Notation. If (x_n) is a convergent sequence, then we write $\lim_{n\to\infty} x_n$ or simply $\lim x_n$ for its unique limit, and we say that $\lim_{n\to\infty} x_n$ *exists*.

Examples.
1. If $x_n = \lambda$ for all $n \in \mathbb{N}$, then $x_n \to \lambda$ as $n \to \infty$ (constant sequence).

2. If there exists $n_0 \in \mathbb{N}$ such that $x_n = \lambda$ for all $n \geqslant n_0$, then $x_n \to \lambda$ as $n \to \infty$ (eventually constant sequence).

Exercise. The sequence $(x_n)_{n=1}^{\infty}$ is convergent if and only if for some (or, equivalently, for all) $n_0 \in \mathbb{N}$ the sequence $(x_{n_0+n})_{n=1}^{\infty}$ is convergent. Thus, for convergence only the tail of the sequence matters. It follows for sequences (x_n) and (y_n) that if there exist $n_0, n_1 \in \mathbb{N}$ such that $x_{n_0+n} = y_{n_1+n}$ for all $n \in \mathbb{N}$ (*i.e.,* if the sequences have the same tail), then (x_n) is convergent if and only if (y_n) is convergent. Moreover, $\lim x_n = \lim y_n$.

Theorem 2. In $\mathbb{R}$ we have $\frac{1}{n} \to 0$ as $n \to \infty$.

Remark. In $\mathbb{Q}$ we also have $\frac{1}{n} \to 0$ as $n \to \infty$, but the proof is entirely algebraic. (In an ordered field F the definition of $x_n \to \lambda$ is that for all $\varepsilon \in F$ with $\varepsilon > 0$ there eixts $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ if $n \geqslant N$ then $|x_n - \lambda| < \varepsilon$.)

Lemma 3. Assume that $x_n \to \lambda$ and $y_n \to \mu$ as $n \to \infty$. Then

(i) $x_n + y_n \to \lambda + \mu$ as $n \to \infty$,

(ii) $x_n \cdot y_n \to \lambda \cdot \mu$ as $n \to \infty$,

(iii) if $y_n \neq 0$ for all n and $\mu \neq 0$, then $\dfrac{x_n}{y_n} \to \dfrac{\lambda}{\mu}$ as $n \to \infty$.

Remarks. 1. If for some real constant $C > 0$ and for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_n - \lambda| < C\varepsilon$ for all $n \geqslant N$, then $x_n \to \lambda$ as $n \to \infty$.

2. If $y_n \to \mu$ as $n \to \infty$ and if $y_n \neq 0$ for all n, one cannot deduce that $\mu \neq 0$. However, the converse almost holds. If $y_n \to \mu$ as $n \to \infty$ and if $\mu \neq 0$, then there exists $n_0 \in \mathbb{N}$ such that $y_n \neq 0$ for all $n \geqslant n_0$. If in addition $x_n \to \lambda$ as $n \to \infty$, then the sequence $\left(\frac{x_n}{y_n}\right)_{n \geqslant n_0}$ converges to $\frac{\lambda}{\mu}$.

3. If $x_n \to \lambda$ as $n \to \infty$ and if c is a constant, then $c \cdot x_n \to c \cdot \lambda$ as $n \to \infty$.

Example. $\dfrac{2n^2 - 7n}{n^2 + 90n - 999} \to 2$ as $n \to \infty$.

Lemma 4. Let (x_n) and (y_n) be convergent real sequences. Then

(i) if $x_n \leqslant y_n$ for all $n \in \mathbb{N}$, then $\lim x_n \leqslant \lim y_n$.

(ii) Assume that (z_n) is another real sequence and that $x_n \leqslant z_n \leqslant y_n$ for all n. If $\lim x_n = \lim y_n$, then (z_n) is also convergent and $\lim z_n = \lim x_n = \lim y_n$. ("Sandwich lemma")

Remarks. 1. If for some constant C we have $x_n \leqslant C$ for all $n \in \mathbb{N}$, then $\lim x_n \leqslant C$. It follows that, if (x_n) is a sequence in the *closed* interval $[a, b] = \{t \in \mathbb{R} : a \leqslant t \leqslant b\}$, then $\lim x_n \in [a, b]$.

2. One cannot replace $\leqslant$ with $<$, *i.e.*, if $x_n < y_n$ for all n, then that does not imply that $\lim x_n < \lim y_n$. *E.g.*, consider $x_n = 0$ and $y_n = \frac{1}{n}$ for all n.

3. Suppose there exists $n_0 \in \mathbb{N}$ such that $x_n \leqslant y_n$ for all $n \geqslant n_0$. Then we can still conclude that $\lim x_n \leqslant \lim y_n$. There is a similar variant of (ii).

4. Assume that (z_n) is a complex sequence converging to $\lambda \in \mathbb{C}$. Then the real sequence $(|z_n|)$ converges to $|\lambda|$. It follows that, if for some constant $C \geqslant 0$ we have $|z_n| \leqslant C$ for all $n \in \mathbb{N}$, then $|\lambda| \leqslant C$.

Examples. 1. $x_n = \dfrac{1}{n\sqrt{1+n}} \to 0$ as $n \to \infty$.

2. $x_n = \dfrac{1}{\sqrt{6n^2 - n + 3}} \to 0$ as $n \to \infty$.

Proposition 5.

(i) For all $x \in \mathbb{R}$ there exists a sequence (q_n) in $\mathbb{Q}$ such that $q_n \to x$ as $n \to \infty$.

(ii) For all $x < y$ in $\mathbb{R}$ there exists $q \in \mathbb{Q}$ such that $x < q < y$ and there exists $r \in \mathbb{R} \setminus \mathbb{Q}$ such that $x < r < y$.

Remarks. 1. For all $x \in \mathbb{R}$ there exists a unique $n \in \mathbb{Z}$ such that $n \leqslant x < n+1$. This n is called the *integer part of* x and is denoted by $\lfloor x \rfloor$. The proof of Proposition 5(i) also shows that $\frac{\lfloor nx \rfloor}{n} \to x$ as $n \to \infty$.

2. Given a sequence (x_n), we have $x_n \to \lambda$ if and only if $|x_n - \lambda| \to 0$.

3. For all $x \in \mathbb{R}$ there exists a sequence (r_n) in $\mathbb{R} \setminus \mathbb{Q}$ such that $r_n \to x$.

Definitions. Let (x_n) be a sequence in $\mathbb{R}$ (or in any ordered field F). We say that (x_n) is

- *increasing* if $x_n \leqslant x_{n+1}$ for all n (equivalently, $x_m \leqslant x_n$ for all $m < n$),
- *strictly increasing* if $x_n < x_{n+1}$ for all n (equivalently, $x_m < x_n$ for all $m < n$),
- *decreasing* if $x_n \geqslant x_{n+1}$ for all n (equivalently, $x_m \geqslant x_n$ for all $m < n$),
- *strictly decreasing* if $x_n > x_{n+1}$ for all n (equivalently, $x_m > x_n$ for all $m < n$),
- *bounded above* if $\{x_n : n \in \mathbb{N}\}$ is bounded above, *i.e.*, there exists C such that $x_n \leqslant C$ for all $n \in \mathbb{N}$,
- *bounded below* if $\{x_n : n \in \mathbb{N}\}$ is bounded below, *i.e.*, there exists C such that $x_n \geqslant C$ for all $n \in \mathbb{N}$,
- *bounded* if $\{x_n : n \in \mathbb{N}\}$ is bounded, *i.e.*, there exists C such that $|x_n| \leqslant C$ for all $n \in \mathbb{N}$.

Note the the definition of being bounded makes sense also in $\mathbb{C}$.

Remark. The n^{th} term of a sequence $(x_i)_{i\in\mathbb{N}}$ is the number x_n together with its position n in the sequence. So a sequence always has infinitely many terms even if the set $\{x_i : i \in \mathbb{N}\}$ is finite.

Lemma 6. Every convergent sequence is bounded.

Remark. The converse is false. *E.g.,* $0, 1, 0, 1, 0, 1, \dots$.

Theorem 7.

(i) A sequence (x_n) in $\mathbb{R}$ that is bounded above converges to $\sup\{x_n : n \in \mathbb{N}\}$.

(ii) A sequence (x_n) in $\mathbb{R}$ that is bounded below converges to $\inf\{x_n : n \in \mathbb{N}\}$.

Definition. A *subsequence* of a sequence (x_n) is a sequence (y_n) such that there exist $k_1 < k_2 < k_3 < \dots$ in $\mathbb{N}$ such that $y_n = x_{k_n}$ for all $n \in \mathbb{N}$.

More formally, thinking of x as a function $x \colon \mathbb{N} \to S$ (for some set S), a subsequence of x is a composite $y = x \circ k$, where $k \colon \mathbb{N} \to \mathbb{N}$ is a strictly increasing function (*i.e.,* $\forall\, m, n \in \mathbb{N}\ m < n \implies k(m) < k(n)$).

Note. Given a sequence (x_n), if (y_n) is a subsequence of (x_n) and (z_n) is a subsequence of (y_n), then (z_n) is a subsequence of (x_n). ("A subsequence of a subsequence is a subsequence.") Indeed, we have $y = x \circ k$ and $z = y \circ \ell$ for strictly increasing functions $k, \ell \colon \mathbb{N} \to \mathbb{N}$. Hence $z = x \circ (k \circ \ell)$ and $k \circ \ell$ is strictly increasing ($m < n \implies \ell(m) < \ell(n) \implies k(\ell(m)) < k(\ell(n))$). With subscript notation we have

$$z_n = y_{\ell_n} = x_{k_{\ell_n}} \qquad \text{for all } n \in \mathbb{N}\ .$$

Proposition 8. Let (x_n) be a convergent sequence with limit λ. Then every subsequence (y_n) of (x_n) converges to λ.

Examples.

1. A second proof that $\frac{1}{n} \to 0$ as $n \to \infty$.

2. If $x \in \mathbb{R}$ and $0 \leqslant x < 1$, then $x^n \to 0$ as $n \to \infty$.

3. If $z \in \mathbb{C}$ and $|z| < 1$, then $z^n \to 0$ as $n \to \infty$.

4. If $x \in \mathbb{R}$ and $x > 0$, then $x^{1/n} \to 1$ as $n \to \infty$.

5. $n^{1/n} \to 1$ as $n \to \infty$.

6. If $x_n = \left(1 + \frac{1}{n}\right)^n$ for $n \in \mathbb{N}$, then (x_n) is convergent. We define

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

Note. Given a sequence (x_n) and a constant λ, by considering (1) on page 5, we have $x_n \not\to \lambda$ as $n \to \infty$ if and only if

$$(\exists\, \varepsilon > 0)\ (\forall\, N \in \mathbb{N})\ (\exists\, n \in \mathbb{N})\ (n \geqslant N \wedge |x_n - \lambda| \geqslant \varepsilon)\ .$$

This is equivalent to saying that for some $\varepsilon > 0$ the set $\{n \in \mathbb{N} : |x_n - \lambda| \geqslant \varepsilon\}$ is an infinite subset of $\mathbb{N}$; enumerating this set in increasing order as $k_1 < k_2 < \dots$ yields a subsequence (x_{k_n}) of (x_n) satisfying $|x_{k_n} - \lambda| \geqslant \varepsilon$ for all $n \in \mathbb{N}$.

Definition. A sequence (x_n) is said to be *divergent* if it is not convergent. This is equivalent to

$$(\forall\,\lambda)\ (\exists\,\varepsilon>0)\ (\forall\,N\in\mathbb{N})\ (\exists\,n\in\mathbb{N})\ (n\geqslant N\wedge|x_n-\lambda|\geqslant\varepsilon)\ .$$

Example. Consider the sequence (x_n) defined by $x_n=1$ if n is odd, and $x_n=0$ if n is even. Then the subsequences (x_{2n}) and (x_{2n-1}) converge to 0, 1, respectively. It follows from Proposition 8 that (x_n) is divergent.

Exercise. Prove directly from definition that the sequence (x_n) is divergent.

Definition. A real sequence (x_n) *tends to* ∞ (written $x_n\to\infty$ as $n\to\infty$) if

$$(\forall\,K\in\mathbb{R})\ (\exists\,N\in\mathbb{N})\ (\forall\,n\in\mathbb{N})\ (n\geqslant N\implies x_n>K)\ .$$

A real sequence (x_n) *tends to* $-\infty$ (written $x_n\to-\infty$ as $n\to\infty$) if

$$(\forall\,K\in\mathbb{R})\ (\exists\,N\in\mathbb{N})\ (\forall\,n\in\mathbb{N})\ (n\geqslant N\implies x_n<K)\ .$$

Note. If the real sequence (x_n) tends to ∞ or $-\infty$, then we write $\lim\limits_{n\to\infty}x_n=\infty$ or $\lim\limits_{n\to\infty}x_n=-\infty$, respectively. In either case, (x_n) is not bounded, and hence (x_n) is divergent by Lemma 6, so $\lim\limits_{n\to\infty}x_n$ does not exist.

Exercise. Prove that $n\to\infty$ as $n\to\infty$.

Definitions. Given a sequence (x_n), the *series* $\sum\limits_{n=1}^{\infty}x_n$ (or $\sum\limits_{n}x_n$, or simply $\sum x_n$) is the sequence (s_n) given by $s_n=\sum\limits_{k=1}^{n}x_k$. We say

- s_n is the n^{th} *partial sum* of the series $\sum x_k$, and
- x_n is the n^{th} *term* of the series $\sum x_k$.

We say the series $\sum x_n$ is *convergent* if the sequence (s_n) is convergent. In this case we denote the limit $\lim\limits_{n\to\infty}s_n$ by $\sum\limits_{n=1}^{\infty}x_n$ and say that the sum $\sum\limits_{n=1}^{\infty}x_n$ *exists*. We say the series $\sum x_n$ is *divergent* if the sequence (s_n) is divergent.

Note. Writing $\sum\limits_{n=1}^{\infty}x_n=\lambda$ implicitly implies that $\sum x_n$ is convergent *and* that $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}x_k=\lambda$.

More definitions. Given a real sequence (x_n), we say that the series $\sum x_n$ *tends to* ∞ if $s_n=\sum\limits_{k=1}^{n}x_k\to\infty$ as $n\to\infty$. In this case we write $\sum\limits_{n=1}^{\infty}x_n=\infty$. Similarly, we say the series $\sum x_n$ *tends to* $-\infty$ if $s_n\to-\infty$ as $n\to\infty$. In this case we write $\sum\limits_{n=1}^{\infty}x_n=-\infty$. In either case, the series is divergent, and thus the sum $\sum\limits_{n=1}^{\infty}x_n$ does not exist.

Examples.
1. If $z \in \mathbb{C}$ and $|z| < 1$, then $\sum z^n$ converges and $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$. This is called the *geometric series.*
2. The series $\sum \frac{1}{\sqrt{n}}$ is divergent.
3. The series $\sum \frac{1}{n}$, called the *harmonic series*, is divergent.
4. The series $\sum (-1)^{n-1}$ is divergent.
5. $\sum_{n=2}^{\infty} \frac{1}{n(n-1)} = 2.$

Lemma 9. Given sequences (x_n) and (y_n),

(i) if $\sum x_n$ and $\sum y_n$ are convergent, then $\sum (x_n + y_n)$ is convergent and $\sum_{n=1}^{\infty} (x_n + y_n) = \sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} y_n$,

(ii) if $\sum x_n$ is convergent, then $\sum c x_n$ convergent and $\sum_{n=1}^{\infty} c x_n = c \sum_{n=1}^{\infty} x_n$ for any constant c,

(iii) if $\sum x_n$ is convergent and there exists $n_0 \in \mathbb{N}$ such that $y_n = x_n$ for all $n \geqslant n_0$, then $\sum y_n$ is convergent.

Example. The series $\sum \left(\frac{1}{n} - \frac{1}{2^n} \right)$ is divergent.

Lemma 10. If the series $\sum x_n$ is convergent, then $x_n \to 0$ as $n \to \infty$.

Example. If $z \in \mathbb{C}$ and $|z| \geqslant 1$, then $\sum z^n$ is divergent.

Note. $\frac{1}{n} \to 0$ as $n \to \infty$, yet $\sum \frac{1}{n}$ is divergent.

Remark. If P is a property of sequences (*e.g.,* being bounded above), then we say a series has property P if its sequence of partial sums has property P.

Proposition 11. Let (x_n) be a real sequence with $x_n \geqslant 0$ for all n. Then $\sum x_n$ is convergent if and only if $\sum x_n$ is bounded above. It follows that $\sum x_n$ is divergent if and only if $\sum x_n$ is not bounded above, and moreover, in this case $\sum_{n=1}^{\infty} x_n = \infty$.

Theorem 12 (Comparison Test). Let (x_n) and (y_n) be real sequences such that for some $n_0 \in \mathbb{N}$ we have $0 \leqslant x_n \leqslant y_n$ for all $n \geqslant n_0$. Then

(i) $\sum y_n$ is convergent $\implies$ $\sum x_n$ is convergent, and

(ii) $\sum x_n$ is divergent $\implies$ $\sum y_n$ is divergent.

Examples.
1. $\sum \frac{1}{n^n}$ is convergent.
2. $\sum \frac{1}{\sqrt{n^2+1}}$ is divergent.

Exercise. Assume $x_n > 0$ and $y_n > 0$ for all n, and that $\frac{x_n}{y_n} \to \ell \neq 0$ as $n \to \infty$ (*i.e.,* x_n and y_n have the same order of magnitude). Then $\sum x_n$ converges if and only if $\sum y_n$ converges.

Definition. A series $\sum x_n$ is said to *converge absolutely* (or said to be *absolutely convergent*) if $\sum |x_n|$ is convergent.

Theorem 13. Absolute convergence implies convergence. Moreover, if $\sum x_n$ converges absolutely, then $\left| \sum\limits_{n=1}^{\infty} x_n \right| \leqslant \sum\limits_{n=1}^{\infty} |x_n|$.

Note. The converse is false. *E.g.,* consider the sequence $1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \ldots$.

Theorem 14 (Cauchy Condensation Test). Assume (x_n) is a real sequence such that $x_n \geqslant 0$ and $x_n \geqslant x_{n+1}$ for all n. Then $\sum x_n$ converges if and only if $\sum 2^n \cdot x_{2^n}$ converges.

Example. Given $\alpha \in \mathbb{R}$, the series $\sum \frac{1}{n^\alpha}$ converges if and only if $\alpha > 1$.

Theorem 15 (Ratio Test). Assume that $x_n \neq 0$ for all n, and that $\left|\frac{x_{n+1}}{x_n}\right| \to \ell$ as $n \to \infty$. Then

(i) if $\ell < 1$, then $\sum x_n$ converges absolutely,

(ii) if $\ell > 1$, then $\sum x_n$ diverges.

Examples.
1. $\sum \frac{1}{n!}$ is convergent.

2. $\sum \frac{n^2}{2^n}$ is convergent. It follows that $\frac{n^2}{2^n} \to 0$ as $n \to \infty$. More generally, given $\alpha, \delta \in \mathbb{R}$, if $\delta > 0$, then $\frac{n^\alpha}{(1+\delta)^n} \to 0$ as $n \to \infty$. ("Exponential growth beats power growth.")

3. $\sum \frac{1}{n}$ diverges and $\sum \frac{1}{n^2}$ converges, yet in both cases the limit $\ell = 1$.

4. $\sum_n \frac{z^n}{n}$ converges absolutely for $|z| < 1$, and diverges for $|z| > 1$. How about $|z| = 1$?

Theorem 16 (Dirichlet's Test). Let (x_n) be a real sequence such that $x_n \geqslant x_{n+1}$ for all n and that $x_n \to 0$ as $n \to \infty$. Let (z_n) be a (real or complex) sequence such that $\sum z_n$ is bounded. Then $\sum x_n z_n$ converges.

Note. The conditions that $x_n \geqslant x_{n+1}$ for all n and that $x_n \to 0$ as $n \to \infty$ imply that $x_n \geqslant 0$ for all n.

Example. Let $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in $\mathbb{C}$. Then $\sum_n \frac{z^n}{n}$ converges for all $z \in \mathbb{T} \setminus \{1\}$ and diverges for $z = 1$.

Theorem 17 (Alternating Series Test). Let (x_n) be a real sequence such that $x_n \geqslant x_{n+1}$ for all n and that $x_n \to 0$ as $n \to \infty$. Then $\sum (-1)^{n+1} x_n$ converges.

Example. $\sum_n \frac{(-1)^n}{n^\alpha}$ is convergent for $\alpha > 0$ (even absolutely convergent for $\alpha > 1$), and divergent for $\alpha \leqslant 0$.

Exercise. This could have been done straight after the example of geometric series. It fully recovers our old notion of what real numbers are. So given $x \in \mathbb{R}$, let $d_0 = \lfloor x \rfloor$ (the integer part of x which we proved exists using the lub axiom), and for $n \in \mathbb{N}$ let $d_n = \lfloor 10^n x \rfloor - 10\lfloor 10^{n-1}x \rfloor$. Prove that $d_n \in \{0, 1, 2, \dots, 9\}$ for all $n \in \mathbb{N}$ and that $x = d_0 + \sum_{n=1}^{\infty} 10^{-n} d_n$. Thus x has decimal expansion $d_0.d_1d_2d_3\dots$. Show also that the decimal expansion obtained this way does not have a tail of 9s, *i.e.,* there is no $n_0 \in \mathbb{N}$ such that $d_n = 9$ for all $n > n_0$. Finally, show that decimal expansions with no tail of 9s are unique. More precisely, if $x = d_0 + \sum_{n=1}^{\infty} 10^{-n} d_n = d'_0 + \sum_{n=1}^{\infty} 10^{-n} d'_n$ are two decimal expansions of x, *i.e.,* that $d_0, d'_0 \in \mathbb{Z}$ and $d_n, d'_n \in \{0, 1, 2, \dots, 9\}$ for all $n \in \mathbb{N}$. Then

either $d_n = d'_n$ for all $n \geqslant 0$,

or $\exists\ n \in \mathbb{N}$ such that $d_k = 0$ and $d'_k = 9$ for all $k \geqslant n$, and $d_{n-1} = d'_{n-1} + 1$,

or $\exists\ n \in \mathbb{N}$ such that $d_k = 9$ and $d'_k = 0$ for all $k \geqslant n$, and $d_{n-1} = d'_{n-1} - 1$.

Definition. A sequence (x_n) is called a *Cauchy sequence* if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_m - x_n| < \varepsilon$ for all $m, n \geqslant N$. With symbols:

$$(\forall\, \varepsilon > 0)\ (\exists\, N \in \mathbb{N})\ (\forall\, m, n \in \mathbb{N})\ (m, n \geqslant N \implies |x_m - x_n| < \varepsilon)$$

Proposition 18. For a sequence we have

$$\text{“convergent”} \implies \text{“Cauchy”} \implies \text{“bounded”}.$$

Note. “Bounded” $\not\Longrightarrow$ “Cauchy” in general. *E.g.,* $1, 0, 1, 0, 1, 0, \dots$.

Lemma 19. If (x_n) is a Cauchy sequence and has a convergent subsequence, then (x_n) is convergent.

Theorem 20 (General Principle of Convergence). Every Cauchy sequence in $\mathbb{R}$ or $\mathbb{C}$ is convergent.

Theorem 21 (Bolzano–Weierstrass). Every bounded sequence in $\mathbb{R}$ or $\mathbb{C}$ has a convergent subsequence.

Remark. The proof of Theorem 20 above used series. An alternative proof is given by Proposition 18 plus Theorem 21 plus Lemma 19.

Lemma 22. Every bounded sequence has a Cauchy subsequence.

Remark. The proof of Theorem 21 above used the “lion hunt” argument. An alternative proof uses Lemma 22 plus Theorem 20.

Lemma 23. Every real sequence has a monotonic (*i.e.,* either increasing or decreasing) subsequence.

Remark. Yet another proof of Theorem 21 uses Lemma 23 plus Theorem 7.

3 Continuous Functions

We will consider functions defined on subsets of $\mathbb{R}$ or $\mathbb{C}$ taking values in $\mathbb{R}$ or $\mathbb{C}$, *i.e.,* functions $f\colon A \to \mathbb{R}$ or $f\colon A \to \mathbb{C}$ where $A \subset \mathbb{R}$ or $A \subset \mathbb{C}$.

Sometimes we will have $A = \mathbb{R}$ or $A = \mathbb{C}$ or, more generally, A will be an interval in $\mathbb{R}$ or a disc in $\mathbb{C}$. Some of our definitions and results will, however, work for arbitrary sets.

An *interval* is a subset I of $\mathbb{R}$ such that $y \in I$ whenever $x, z \in I$ and $x < y < z$. It is easy to check that an interval has one of the following forms:

$$\begin{array}{rlrlr}
(a,b) &= \{t \in \mathbb{R} : a < t < b\} & (-\infty,b) &= \{t \in \mathbb{R} : t < b\} & (-\infty,\infty) = \mathbb{R} \\
[a,b) &= \{t \in \mathbb{R} : a \leqslant t < b\} & (-\infty,b] &= \{t \in \mathbb{R} : t \leqslant b\} & \\
(a,b] &= \{t \in \mathbb{R} : a < t \leqslant b\} & (a,\infty) &= \{t \in \mathbb{R} : a < t\} & \\
[a,b] &= \{t \in \mathbb{R} : a \leqslant t \leqslant b\} & [a,\infty) &= \{t \in \mathbb{R} : a \leqslant t\} &
\end{array}$$

where $a, b \in \mathbb{R}$. A *closed interval* is an interval of the form $[a,b]$, $(-\infty,b]$, $[a,\infty)$ or $(-\infty,\infty)$. An *open interval* is an interval of the form (a,b), $(-\infty,b)$, (a,∞) or $(-\infty,\infty)$. Note that $(-\infty,\infty)$ is both an open and a closed interval.

In $\mathbb{C}$ there are two types of discs. Given $a \in \mathbb{C}$ and $r \in \mathbb{R}$ with $r > 0$ we have

$D(a,r) = \{z \in \mathbb{C} : |z-a| < r\}$ is the *open disc of centre a and radius r*,

$\overline{D}(a,r) = \{z \in \mathbb{C} : |z-a| \leqslant r\}$ is the *closed disc of centre a and radius r.*

[In the *Metric & Topological Spaces* course next term you will define what it means for a subset of $\mathbb{R}$ or $\mathbb{C}$ to be open or to be closed. Fortunately, open intervals and open discs are open, closed intervals and closed discs are closed.]

Note. We can also define discs in $\mathbb{R}$, where they become intervals. For $a, r \in \mathbb{R}$ with $r > 0$

the set $\{x \in \mathbb{R} : |x-a| < r\}$ is the open interval $(a-r, a+r)$, and

the set $\{x \in \mathbb{R} : |x-a| \leqslant r\}$ is the closed interval $[a-r, a+r]$.

Note. As in the next definition, when the field is not specified, the statement applies to both real- and complex-valued functions on subsets of either $\mathbb{R}$ or $\mathbb{C}$.

Definition. Let f be a function on a set A. Given $a \in A$, we say f *is continuous at a* if for all real $\varepsilon > 0$, there exists a real $\delta > 0$ such that for all $x \in A$ if $|x-a| < \delta$, then $|f(x) - f(a)| < \varepsilon$. With symbols:

$$(\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (|x-a| < \delta \implies |f(x) - f(a)| < \varepsilon)$$

Remark. Informally, this means that if $x \approx a$ then $f(x) \approx f(a)$. More formally, we write

$f(x) = f(a) + \eta(x)$ where η is the function on A defined by $\eta(x) = f(x) - f(a)$.

We think of approximating f near a by the constant function $x \mapsto f(a)$. In this context $\eta(x)$ is the error term at x. Continuity at a means that the error term can be made arbitrarily small if x is sufficiently close to a.

Definition. Let f be a function on a set A. We say f *is continuous* if f is continuous at a for every $a \in A$. With symbols:

$$(\forall\, a \in A)\ (\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon)$$

Examples.
1. $f(x) = c$ for all $x \in A$ (constant function).

2. $f(x) = x$ for all $x \in A$ (identity function on A).

3. $f(x) = x^2$ for all $x \in A$. Note that here δ depends on a as well as ε.

4. $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$ is continuous. Similarly, $f\colon \mathbb{C} \to \mathbb{R}$ given by $f(x) = |x|$ is continuous.

5. $\mathrm{Re}, \mathrm{Im}\colon \mathbb{C} \to \mathbb{R}$ are continuous.

Theorem 1. Let f be a function on a set A, and let $a \in A$. TFAE.

(i) f is continuous at a.

(ii) For every sequence (x_n) in A, if $x_n \to a$ then $f(x_n) \to f(a)$.

Note. Part (ii) above says that taking limit commutes with a continuous function: $f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n)$.

Given functions f and g on a set A, we define new functions $f + g$ and $f \cdot g$ on A pointwise by $(f + g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(x)g(x)$. If $g(x) \neq 0$ for all $x \in A$, we also define f/g on A by $(f/g)(x) = f(x)/g(x)$.

Proposition 2. Let f and g be functions on a set A. Given $a \in A$, if f and g are continuous at a, then so are $f + g$, $f \cdot g$ and, when defined, f/g. It follows that if f and g are continuous, then so are $f + g$, $f \cdot g$ and, when defined, f/g.

Exercise. In lectures we proved this using Theorem 1 and Lemma 2.3. Give a direct proof using the ε-δ definition of continuity.

Examples.
1. *Polynomial functions* are continuous. These are functions of the form $p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$, where n is a non-negative integer and $a_0, a_1, \ldots, a_n$ are fixed constants called the *coefficients* of p. We talk about real or complex polymonials depending on whether the coefficients are real or complex, respectively.

2. $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$ is continuous.

3. Given polynomial functions p and q, let $A = \{x : q(x) \neq 0\}$. Then p/q is defined on A and is continuous. Functions of this form are called *rational functions*.

Composition. Suppose f is a function on a set A, g is a function on a set B, and $B \supset f(A) = \mathrm{im}\, f = \{y : \exists\, x \in A\ f(x) = y\}$. Then the *composite* $g \circ f$ is the function on A defined by $(g \circ f)(x) = g(f(x))$.

Proposition 3. Let A, B, f, g be as above. Let $a \in A$ and $b = f(a)$ (so $b \in B$). If f is continuous at a and g is continuous at b, then $g \circ f$ is continuous at a. It follows that if f and g are continuous, then so is $g \circ f$.

Exercise. In lectures we gave a direct proof using the ε-δ definition. Give an other proof using Theorem 1.

Examples.

1. Given a function $f\colon A \to \mathbb{C}$ and $a \in A$, if f is continuous at a, then so are the functions $\mathrm{Re}(f), \mathrm{Im}(f), |f|\colon A \to \mathbb{R}$.

2. Given a function f on a set A and given $B \subset A$, the *restriction of f to B* is the function $f{\restriction}_B$ on B defined by $f{\restriction}_B\,(x) = f(x)$ for all $x \in B$. Note that $f{\restriction}_B = f \circ j$ where j is the identity function on B. We shall say that f *is continuous on* B if f is continuous at b for every $b \in B$. (So "f continuous on A" is the same as "f continuous".) How is this related to the continuity of $f{\restriction}_B$?

It follows from Proposition 3 that if $b \in B$ and f is continuous at b, then $f{\restriction}_B$ is also continuous at b. This is also clear from the comparison of the definitions. "f continuous at b" means

$$(\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (|x - b| < \delta \implies |f(x) - f(a)| < \varepsilon)$$

whereas "$f{\restriction}_B$ continuous at b" means

$$(\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in B)\ (|x - b| < \delta \implies |f(x) - f(a)| < \varepsilon)$$

It follows that if f is continuous on B, then $f{\restriction}_B$ is continuous. The converse is false in general as the next two examples demonstrate.

(i) Let $A = \mathbb{R}$, let $f\colon A \to \mathbb{R}$ be given by

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

and let $B = [0, \infty)$. Then $f{\restriction}_B$ is the constant one function, which is continuous. So in particular $f{\restriction}_B$ is continuous at 0, whereas f is not.

(ii) Let $A = \mathbb{R}$, let $f\colon A \to \mathbb{R}$ be given by

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

and let $B = \mathbb{Q}$. Then $f{\restriction}_B$ is the constant zero function, which is continuous. However, f is not continuous at any point.

Remark. This is one final comment on the above. Given sets $B \subset A$ and a point $b \in B$, assume that there exists $r > 0$ such that $\{x \in A : |x - b| < r\} \subset B$. Then if f is a function on A such that $f{\restriction}_B$ is continuous at b, then f is continuous at b. It follows that if

$$(\dagger) \qquad (\forall\ x \in B)\ (\exists\ r > 0)\ (\{y \in A : |y - x| < r\} \subset B)$$

then if $f{\restriction}_B$ is continuous, then f is continuous at every point of B (*i.e.*, f is continuous on B). If ($\dagger$) above holds, then we say B is an *open subset* of A. (See the *Metric & Topological Spaces* course next term.)

Definition. Let f be a function on a set A. Given $B \subset A$, we say f *is bounded on* B if the set $\{f(x) : x \in B\}$ is bounded, *i.e.,* if there exists $C \in \mathbb{R}$ such that $|f(x)| \leqslant C$ for all $x \in B$. We say f *is bounded* if f is bounded on A.

Theorem 4. A continuous real-valued function on a closed bounded interval is bounded and attains its bounds. *I.e.,* if $f\colon [a,b] \to \mathbb{R}$ is continuous, then there exists $C \in \mathbb{R}$ such that $|f(x)| \leqslant C$ for all $x \in [a,b]$, and there exist $y, z \in [a,b]$ such that $f(y) = \inf\{f(x) : x \in [a,b]\}$ and $f(z) = \sup\{f(x) : x \in [a,b]\}$.

Note. If $f\colon [a,b] \to \mathbb{C}$ is continuous, then so is $|f|\colon [a,b] \to \mathbb{R}$. So by the above result f is bounded.

Examples.
1. $f\colon (0,1] \to \mathbb{R}$, $f(x) = \frac{1}{x}$, is continuous but not bounded.

2. $f\colon [0,\infty) \to \mathbb{R}$, $f(x) = x$, is continuous but not bounded.

Theorem 5 (Intermediate Value Theorem, IVT). Let I be an interval and $f\colon I \to \mathbb{R}$ be continuous. Then for all $x, y \in I$ and $c \in \mathbb{R}$, if $f(x) < c < f(y)$, then there exists $z \in \mathbb{R}$, strictly between x and y, such that $f(z) = c$. It follows that $J = f(I)$ is an interval.

Examples.
1. Every real polynomial of odd degree has a real root.

2. The Möbius map $f(z) = \frac{z+1}{z-1}$ maps the Möbius lines $\mathbb{R} \cup \{\infty\}$ and $\mathrm{i}\mathbb{R} \cup \{\infty\}$ to the Möbius lines $\mathbb{R} \cup \{\infty\}$ and $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, respectively, which divide $\mathbb{C} \cup \{\infty\}$ into four regions. Theorem 5 shows why f maps each quadrant onto one of these regions.

3. In *Numbers & Sets* you constructed functions $f\colon \mathbb{R} \to \mathbb{R}$ that take every value on every interval. So the converse of Theorem 5 is completely false.

Definition. Let $f\colon A \to \mathbb{R}$ be a function defined on a subset A of $\mathbb{R}$. We say that

f *is increasing* if for all $x, y \in A$, if $x \leqslant y$ then $f(x) \leqslant f(y)$;

f *is strictly increasing* if for all $x, y \in A$, if $x < y$ then $f(x) < f(y)$;

f *is decreasing* if for all $x, y \in A$, if $x \leqslant y$ then $f(x) \geqslant f(y)$;

f *is strictly decreasing* if for all $x, y \in A$, if $x < y$ then $f(x) > f(y)$.

Theorem 6. Let $f\colon I \to \mathbb{R}$ be a strictly increasing (or strictly decreasing) function on an interval I. Let $J = f(I)$ be the image of f. Then $f\colon I \to J$ is a bijection and $f^{-1}\colon J \to I$ is continuous.

Note. f is not assumed continuous, so J need not be an interval.

Application. Let $I = [0,\infty)$, $n \in \mathbb{N}$ and $f\colon I \to \mathbb{R}$ be defined by $f(x) = x^n$. Then $f(I) = I$ and $f\colon I \to I$ is a continuous bijection with continuous inverse.

This shows that for all $x \in [0,\infty)$ there is a unique $y \in [0,\infty)$, denoted by $x^{1/n}$, such that $y^n = x$. Moreover, the map $x \mapsto x^{1/n}\colon [0,\infty) \to \mathbb{R}$ is continuous.

Given $q = \frac{r}{s} \in \mathbb{Q}$ with $r, s \in \mathbb{N}$, we define $x^q = \left(x^r\right)^{1/s}$ and $x^0 = 1$ for $x \in [0, \infty)$, and define $x^{-q} = \frac{1}{x^q}$ for $x \in (0, \infty)$. These powers are well-defined and have the usual properties. We will return to this when we define x^α for arbitrary real α.

Note. It now follows that if $x_n \geqslant 0$ for all $n \in \mathbb{N}$, and $x_n \to \lambda$ as $n \to \infty$, then, for example, $\sqrt{x_n} \to \sqrt{\lambda}$, etc.

Limits. Let f be a function defined near a point a, *i.e.,* there exists $r > 0$ such that f is defined on some set A that contains the set $\{x : 0 < |x - a| < r\}$. We say $f(x)$ *tends to* λ *as* x *tends to* a, and write $f(x) \to \lambda$ as $x \to a$, if

$$(\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (0 < |x - a| < \delta \implies |f(x) - \lambda| < \varepsilon)$$

Informally, if $x \approx a$ and $x \neq a$, then $f(x) \approx \lambda$.

Note. 1. f need not be defined at a, and if it is, the value $f(a)$ is irrelevant.

2. Let $g(x) = f(x)$ for $x \neq a$, and $g(a) = \lambda$. Then $f(x) \to \lambda$ as $x \to a$ if and only if g is continuous at a.

Example. Let $f(x) = x + 1$ for $x \neq 0$ and $f(0) = 0$. Then $f(x) \to 1$ as $x \to 0$.

Properties. 1. $f(x) \to \lambda$ as $x \to a$ if and only if for every sequence (x_n) in $A \setminus \{a\}$, if $x_n \to a$ then $f(x_n) \to \lambda$.

2. If $f(x) \to \lambda$ and $f(x) \to \mu$ as $x \to a$, then $\lambda = \mu$. We write $\lim\limits_{x \to a} f(x) = \lambda$.

3. Let f and g be functions defined near a. Assume that $f(x) \to \lambda$ and $g(x) \to \mu$ as $x \to a$. Then $(f + g)(x) \to \lambda + \mu$, $(fg)(x) \to \lambda\mu$ and, when defined, $(f/g)(x) \to \lambda/\mu$ as $x \to a$.

Note. Composition, however, does not quite work. If $f(x) \to b$ as $x \to a$, and $g(y) \to \lambda$ as $y \to b$, then it need not follow that $g(f(x)) \to \lambda$ as $x \to a$. For example, define $f(x) = g(x) = 0$ for all $x \neq 0$, $f(0) = 0$, and $g(0) = 1$. Then $f(x) \to 0$ as $x \to 0$, $g(y) \to 0$ as $y \to 0$, but $g(f(x)) \not\to 0$ as $x \to 0$.

Definition. Let f be a real-valued function on a set A that contains the set $\{x : 0 < |x - a| < r\}$ for some $r > 0$. We say $f(x)$ *tends to* ∞ *as* x *tends to* a, and write $f(x) \to \infty$ as $x \to a$, if

$$(\forall\, K \in \mathbb{R})\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (0 < |x - a| < \delta \implies f(x) > K)$$

We say $f(x)$ *tends to* $-\infty$ *as* x *tends to* a, and write $f(x) \to -\infty$ as $x \to a$, if

$$(\forall\, K \in \mathbb{R})\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (0 < |x - a| < \delta \implies f(x) < K)$$

Note. Property 1 above remains valid with $\lambda = \infty$ or $\lambda = -\infty$, and Property 2 above remains valid for arbitrary $\lambda, \mu \in \mathbb{R} \cup \{\infty, -\infty\}$.

One-sided limits. Let f be a function defined on a subset A of $\mathbb{R}$. If A contains the interval $(a, a + r)$ for some $r > 0$, we say $f(x)$ *tends to* λ *as* x *tends to* a *from above*, and write $f(x) \to \lambda$ as $x \to a^+$, if

$$(\forall\, \varepsilon > 0)\ (\exists\, \delta > 0)\ (\forall\, x \in A)\ (a < x < a + \delta \implies |f(x) - \lambda| < \varepsilon)$$

If A contains the interval $(a-r,a)$ for some $r>0$, we say $f(x)$ *tends to* λ *as* x *tends to* a *from below*, and write $f(x)\to\lambda$ as $x\to a^-$, if

$$(\forall\,\varepsilon>0)\ (\exists\,\delta>0)\ (\forall\,x\in A)\ (a-\delta<x<a \implies |f(x)-\lambda|<\varepsilon)$$

These satisfy properties analogous to the ones above. In particular, limits at a from above and below are unique when they exist and are denoted by $\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$, respectively.

Note. If f is defined on a subset A of $\mathbb{R}$ that contains the set $(a-r,a)\cup(a,a+r)$ for some $r>0$, then $\lambda=\lim\limits_{x\to a} f(x)$ if and only if $\lambda=\lim\limits_{x\to a^+} f(x)=\lim\limits_{x\to a^-} f(x)$.

Exercise. Formulate definitions of $f(x)\to\infty$ or $-\infty$ as $x\to a^+$ or a^-. Formulate and prove properties analogous to Properties 1 and 2 above.

Limits at $\pm\infty$. Let f be a function on a subset A of $\mathbb{R}$. If A contains (r,∞) for some $r\in\mathbb{R}$, we say $f(x)$ *tends to* λ *as* x *tends to* ∞, and write $f(x)\to\lambda$ as $x\to\infty$, if

$$(\forall\,\varepsilon>0)\ (\exists\,\rho\in\mathbb{R})\ (\forall\,x\in A)\ (\rho<x \implies |f(x)-\lambda|<\varepsilon)$$

This satisfies properties analogous to Properties 1,2 and 3 of limits above. In particular, we write $\lim\limits_{x\to\infty} f(x)$ for the unique λ if it exists.

Exercise. Formulate definitions of $f(x)\to\infty$ or $-\infty$ as $x\to\infty$, and of $f(x)\to\lambda$ or ∞ or $-\infty$ as $x\to-\infty$. Formulate appropriate analogues of the properties above and prove they are satisfied.

4 Differentiation

Let $f\colon A\to\mathbb{R}$ be a real-valued function on a subset A of $\mathbb{R}$. Given $c\in A$, we say f *is differentiable at* c if the limit $\lim\limits_{x\to c}\frac{f(x)-f(c)}{x-c}$ exists. We denote this limit by $f'(c)$, called *the derivative of* f *at* c.

Remarks. 1. It is implicit in this definition that the domain A of the function f contains the interval $(c-r,c+r)$ for some $r>0$, and hence the definition makes sense. (See the definition of limits at the end of the last chapter.)

2. If f is differentiable at c, then $\lim\limits_{x\to c}\frac{f(x)-f(c)}{x-c}=\lim\limits_{h\to 0}\frac{f(c+h)-f(c)}{h}$.

Assume that B is a subset of A, and for all $c\in B$ there exists $r>0$ such that $(c-r,c+r)\subset B$. We say f *is differentiable on* B if f is differentiable at c for all $c\in B$. In that case the *derivative of* f *on* B is the function $f'\colon B\to\mathbb{R}$ that maps $c\mapsto f'(c)$.

Note. 1. Typically, B will be an open interval (a,b), (a,∞), $(-\infty,b)$ or $(-\infty,\infty)=\mathbb{R}$. Note that if, say, $c\in(a,b)$, then $(c-r,c+r)\subset(a,b)$ for $r=\min(c-a,b-c)$.

2. When $B=A$, we might simply say f *is differentiable* or the *derivative of* f, instead of f is differentiable on A or the derivative of f on A.

Interpretation. Geometrically, the ratio $\frac{f(x)-f(c)}{x-c}$ is the slope of the chord joining the points $(c, f(c))$ and $(x, f(x))$ on the graph of f. In the limit as $x \to c$, this becomes $f'(c)$, the slope of the tangent to the curve $y = f(x)$ at the point $(c, f(c))$.

Another very useful way of thinking about differentiation is in terms of approximation. Define a function

$$\varepsilon(h) = \begin{cases} \frac{f(c+h)-f(c)}{h} & \text{if } h \neq 0 \text{ and } c+h \in A \text{ (in particular if } 0 < |h| < r) \\ 0 & \text{if } h = 0 \ . \end{cases}$$

If f is differentiable at c, then we have

$$f(c+h) = f(c) + f'(c) \cdot h + h\varepsilon(h) \quad \text{where } \varepsilon(h) \to 0 \text{ as } h \to 0 \ .$$

The condition that $\lim_{h \to 0} \varepsilon(h) = 0$ is equivalent to ε being continuous at 0.

Conversely, assume there exist $\lambda \in \mathbb{R}$ such that

$$f(c+h) = f(c) + \lambda h + h\varepsilon(h)$$

where ε is a function such that $\varepsilon(h) \to 0$ as $h \to 0$ (equivalently, $\varepsilon(0) = 0$ and ε is continuous at 0). Then f is differentiable at c and $f'(c) = \lambda$. The map $h \mapsto f(c) + \lambda h$ is a linear function approximating f near c, and $h\varepsilon(h)$ is the error term. We compare this to continuity.

f continuous at c: $\quad f(c+h) = f(c) + \eta(h), \quad \eta(h) \to 0$ as $h \to 0$.

f differentiable at c: $\quad f(c+h) = f(c) + \lambda h + h\varepsilon(h), \quad \varepsilon(h) \to 0$ as $h \to 0$.

Proposition 1. If f is differentiable at c, then f is continuous at c.

Notation. Let η and φ be functions defined on the set $(-r, r) \setminus \{0\}$ for some $r > 0$, and assume $\varphi(h) \neq 0$ for all h. Then we write

$$\eta = \mathrm{o}(\varphi) \qquad \text{if} \qquad \frac{\eta(h)}{\varphi(h)} \to 0 \text{ as } h \to 0 \ .$$

Thus, for example, $\eta = \mathrm{o}(1)$ means $\eta(h) \to 0$ as $h \to 0$, and $\eta = \mathrm{o}(h)$ means $\eta(h)/h \to 0$ as $h \to 0$. We then have

f continuous at c: $\quad f(c+h) = f(c) + \mathrm{o}(1)$.

f differentiable at c: $\quad f(c+h) = f(c) + \lambda h + \mathrm{o}(h)$.

Examples.

1. $f(x) = \lambda$ for all $x \in \mathbb{R}$ (constant function). Then $f'(x) = 0$ for all x.

2. $f(x) = x$ for all $x \in \mathbb{R}$ (identity). Then $f'(x) = 1$ for all x.

3. $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = 1/x$. Then $f'(x) = -1/x^2$ for all x.

4. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$. Then $f'(x) = 1$ for $x > 0$, $f'(x) = -1$ for $x < 0$, and f is not differentiable at 0.

Proposition 2. Assume f and g are differentiable at c. Then so are $f + g$, fg and, if $g(c) \neq 0$, f/g. Moreover, we have

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

(ii) $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$ (product rule)

(iii) $\left(\frac{f}{g}\right)'(c) = \dfrac{f'(c)g(c) - f(c)g'(c)}{g(c)^2}$ (quotient rule).

Remarks. 1. Part (ii) implies that for $\lambda \in \mathbb{R}$ we have $(\lambda f)'(c) = \lambda f'(c)$.

2. In part (iii) we showed that if $g(c) \neq 0$ and if g is continuous at c, then there exists $\delta > 0$ such that $g(x) \neq 0$ for all $x \in (c-\delta, c+\delta)$.

Examples. 1. All real polynomials are differentiable on $\mathbb{R}$. For example, if $f(x) = x^2$, then $f = gg$, where $g(x) = x$. Hence, $f'(x) = 1x + x1 = 2x$. By induction, it follows that for $n \in \mathbb{N}$, if $f(x) = x^n$, then $f'(x) = nx^{n-1}$. In general, if $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$, then f is differentiable on $\mathbb{R}$ and $f'(x) = a_1 + 2a_2 x + \cdots + na_n x^{n-1}$.

2. The rational function p/q is differentiable on the set $\{x \in \mathbb{R} : q(x) \neq 0\}$.

Theorem 3 (Chain Rule). If f is differentiable at a, and g is differentiable at $b = f(a)$, then $g \circ f$ is differentiable at a, and $(g \circ f)(a) = g'(f(a))f'(a)$.

Note. The assumptions imply that $g \circ f$ is defined on the interval $(a-s, a+s)$ for some $s > 0$.

Examples. 1. If f is differentiable at c, and $f(c) \neq 0$, then $|f|$ is differentiable at c with $f'(c) = \operatorname{sign}(f(c))f'(c)$. Note however that if $f(c) = 0$ then that does not imply that f is not differentiable at c. Consider $f(x) = 0$ for all x.

2. Assume g is differentiable at c and $g(c) \neq 0$. Then $1/g = h \circ g$, where $h(t) = 1/t$ for all $t \in \mathbb{R} \setminus \{0\}$, and hence $1/g$ is differentiable at c and $(1/g)'(c) = h'(g(c))g'(c) = -g'(c)/g(c)^2$.

Theorem 4 (Rolle's Theorem). Assume that the function $f\colon [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b)$, then there exists $c \in (a,b)$ such that $f'(c) = 0$.

Theorem 5 (Mean Value Theorem, MVT). Assume that the function $f\colon [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b). Then there exists $c \in (a,b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Note. Geometrically, the MVT says that there is a tangent to the curve $y = f(x)$ that is parallel with the chord connecting $(a, f(a))$ and $(b, f(b))$.

Corollary 6. Let I be an interval, and $J = I \setminus \{\text{endpoints of } I\}$. Assume that the function $f\colon I \to \mathbb{R}$ is continuous on I and differentiable on J. Then

(i) $f' = 0$ on $J \iff f$ constant on I,

(ii) $f' \geqslant 0$ on $J \iff f$ increasing on I,

(iii) $f' \leqslant 0$ on $J \iff f$ decreasing on I,

(iv) $f' > 0$ on $J \implies f$ strictly increasing on I,

(v) $f' < 0$ on $J \implies f$ strictly decreasing on I.

Remarks. 1. It is important that the domain here is an interval. For example, if $f = 0$ on $(-\infty, 0)$ and $f = 1$ on $(0, \infty)$, then $f' = 0$ on $\mathbb{R} \setminus \{0\}$, yet f is not constant.

2. f strictly increasing on $I \;\not\Longrightarrow\; f' > 0$ on J. *E.g.*, $f(x) = x^3$, $I = J = \mathbb{R}$.

Remark. Assume that f is a bijection between open intervals such that both f and f^{-1} are differentiable. Then it follows from the Chain Rule that $f'(x) \neq 0$ for all x, and $\big(f^{-1}\big)'(y) = 1/f'\big(f^{-1}(y)\big)$ for all y.

Lemma 7. Let I be an interval and $f\colon I \to \mathbb{R}$ be a continuous and injective function. Then f is strictly monotonic.

Theorem 8 (Inverse Function Theorem, IFT). Let I be an open interval and $f\colon I \to \mathbb{R}$ be a differentiable function such that $f'(x) \neq 0$ for all $x \in I$. Then $J = f(I)$ is an open interval, $f\colon I \to J$ is a bijection, $g = f^{-1}\colon J \to I$ is differentiable, and $g'(y) = 1/f'(g(y))$.

Application. Let $n \in \mathbb{N}$. Then $f\colon (0, \infty) \to \mathbb{R}$ given by $f(x) = x^{1/n}$ is differentiable with $f'(x) = \frac{1}{n} x^{\frac{1}{n}-1}$ for all x.

More generally, one can show that for $q \in \mathbb{Q}$, the function $f\colon (0, \infty) \to \mathbb{R}$ given by $f(x) = x^q$ is differentiable with $f'(x) = qx^{q-1}$ for all x.

Complex Differentiation. Let $f\colon U \to \mathbb{C}$ be a function on a set $U \subset \mathbb{C}$. Given $a \in U$, we say f *is complex differentiable at* a if $\lim\limits_{h\to 0} \frac{f(a+h)-f(a)}{h}$ exists. We denote this limit by $f'(a)$ and call it the *derivative of* f *at* a. We say f *is complex differentiable on* U if f is complex differentiable at a for all $a \in U$. The *derivative of* f *on* U is the function $f'\colon U \to \mathbb{C}$, $a \mapsto f'(a)$.

Note. For the definition of $f'(a)$ to make sense, we need that $D(a, r) \subset U$ for some $r > 0$ (at least according to the way we defined limits). Similarly, for complex differentiability on U, we need U to be an *open set*, which means that for all $a \in U$ there exists $r > 0$ such that $D(a, r) \subset U$. In all our examples U will be either $\mathbb{C}$ or an open disc $D(w, R)$ for some $w \in \mathbb{C}$ and $R > 0$, which are open.

Remarks. Note that f is complex differentiable at a if and only if there exist $\lambda \in \mathbb{C}$ and a function ε such that $f(a+h) = f(a) + \lambda h + h\varepsilon(h)$ and $\varepsilon(h) \to 0$ as $h \to 0$ (or, equivalently, $\varepsilon(0) = 0$ and ε is continuous at 0), and then $f'(a) = \lambda$. It is easy to check that Propositions 1 and 2 and Theorem 3 remain valid with nearly identical proofs.

Lemma 9. Let $\sum\limits_{n=0}^{\infty} x_n$ and $\sum\limits_{n=0}^{\infty} y_n$ be absolutely convergent series (real or complex). For $n \in \mathbb{N}$ let $z_n = \sum\limits_{j=0}^{n} x_j y_{n-j}$. Then $\sum\limits_{n=0}^{\infty} z_n$ is absolutely convergent, and moreover $\sum\limits_{n=0}^{\infty} z_n = \Big(\sum\limits_{n=0}^{\infty} x_n\Big)\Big(\sum\limits_{n=0}^{\infty} y_n\Big)$.

Definition. The *exponential function* is the function exp defined as follows.

$$\exp\colon \mathbb{C} \to \mathbb{C}\ , \quad \exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}\ .$$

It follows from the Ratio Test that the series converges absolutely for all $z \in \mathbb{C}$.

Theorem 10. Properties of exp.

(i) $\exp(z + w) = \exp(z)\exp(w)$ for all $z, w \in \mathbb{C}$

(ii) $\exp(0) = 1$, and $\exp(z) \neq 0$ for all $z \in \mathbb{C}$

(iii) $\overline{\exp z} = \exp(\bar{z})$

(iv) $\exp(x) \in \mathbb{R}$ for all $x \in \mathbb{R}$

(v) $|\exp(ix)| = 1$ for all $x \in \mathbb{R}$

(vi) exp is complex differentiable at 0 with $\exp'(0) = 1$

(vii) exp is complex differentiable on $\mathbb{C}$ with $\exp' = \exp$.

Note. The restriction of exp to $\mathbb{R}$ is a real-valued function $\exp\colon \mathbb{R} \to \mathbb{R}$.

Theorem 11. The real exponential function is a strictly increasing differentiable bijection $\mathbb{R} \to (0, \infty)$.

Definition. The *(natural) logarithm* is the function $\log\colon (0, \infty) \to \mathbb{R}$ defined as the inverse of the real exponential function.

Theorem 12. The natural logarithm is a strictly increasing differentiable bijection $(0, \infty) \to \mathbb{R}$ satisfying

(i) $\log 1 = 0$, (ii) $\log(xy) = \log x + \log y$, (iii) $\log'(y) = \dfrac{1}{y}$.

Definition (exponentiation.). For a real number $x > 0$ and for $\alpha \in \mathbb{C}$ we define $x^\alpha = \exp(\alpha \log x)$. Note that if $\alpha \in \mathbb{R}$ then x^α is real and positive.

Theorem 13. Properties of exponentiation.

(i) $x^0 = 1$, $1^\alpha = 1$, $x^{\alpha+\beta} = x^\alpha x^\beta$ and $(xy)^\alpha = x^\alpha y^\alpha$ for all $x, y > 0$, $\alpha, \beta \in \mathbb{C}$.

(ii) $\big(x^\alpha\big)^\beta = x^{\alpha\beta}$ for all $x > 0$, $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{C}$.

(iii) For $x > 0$ and $\alpha \in \mathbb{Q}$ the definition of x^α agrees with the previous definition.

(iv) The map $x \mapsto x^\alpha\colon (0, \infty) \to (0, \infty)$ is a strictly increasing, differentiable bijection if $\alpha > 0$, and a strictly decreasing, differentiable bijection if $\alpha < 0$.

(v) Let $\alpha \in \mathbb{C}$. Then $\dfrac{x^\alpha}{\exp x} \to 0$ as $x \to \infty$.

If $\mathrm{Re}(\alpha) > 0$, then $\dfrac{\log x}{x^\alpha} \to 0$ as $x \to \infty$ and $x^\alpha \log x \to 0$ as $x \to 0^+$.

Exercise. Recall that $\mathrm{e} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$. Prove that $\mathrm{e} = \exp(1)$ and that

$$\mathrm{e}^z = \exp(z) = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n \qquad \text{for all } z \in \mathbb{C} .$$

Trigonometric and hyperbolic functions. We define functions sin, cos, sinh, cosh: $\mathbb{C} \to \mathbb{C}$ by letting

$$\sin z = \frac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}} \qquad \sinh z = \frac{\mathrm{e}^{z} - \mathrm{e}^{-z}}{2}$$

$$\cos z = \frac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2} \qquad \cosh z = \frac{\mathrm{e}^{z} + \mathrm{e}^{-z}}{2} .$$

By the definition of exp, we have

$$\begin{aligned}
\sin z &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} &&= z - \frac{z^3}{6} + \frac{z^5}{120} - \dots \\
\cos z &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} &&= 1 - \frac{z^2}{2} + \frac{z^4}{24} - \dots \\
\sinh z &= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} z^{2n+1} &&= z + \frac{z^3}{6} + \frac{z^5}{120} + \dots \\
\cosh z &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} z^{2n} &&= 1 + \frac{z^2}{2} + \frac{z^4}{24} + \dots
\end{aligned}$$

Proposition 14.

(i) If f is any of these functions, then $\overline{f(z)} = f(\bar{z})$ for all $z \in \mathbb{C}$. It follows that $f(x) \in \mathbb{R}$ for all $x \in \mathbb{R}$.

sin and sinh are odd functions ($f(-z) = -f(z)$ for all z).

cos and cosh are even functions ($f(-z) = f(z)$ for all z).

$\sin(0) = \sinh(0) = 0$ and $\cos(0) = \cosh(0) = 1$.

(ii) For all $z, w \in \mathbb{C}$ we have

$$\begin{aligned}
\sin(z+w) &= \sin z \cdot \cos w + \cos z \cdot \sin w & \sin(2z) &= 2\sin z \cdot \cos z \\
\cos(z+w) &= \cos z \cdot \cos w - \sin z \cdot \sin w & \cos(2z) &= \cos^2 z - \sin^2 z \\
\sinh(z+w) &= \sinh z \cdot \cosh w + \cosh z \cdot \sinh w & \sinh(2z) &= 2\sinh z \cdot \cosh z \\
\cosh(z+w) &= \cosh z \cdot \cosh w + \sinh z \cdot \sinh w & \cosh(2z) &= \cosh^2 z + \sinh^2 z
\end{aligned}$$

(iii) $1 = \cos^2 z + \sin^2 z = \cosh^2 z - \sinh^2 z$,

$\mathrm{e}^{\mathrm{i}z} = \cos z + \mathrm{i}\sin z$ and $\mathrm{e}^z = \cosh z + \sinh z$.

$\cosh z = \cos(\mathrm{i}z)$ and $\sinh z = -\mathrm{i}\sin(\mathrm{i}z)$.

(iv) The functions sin, cos, sinh and cosh are complex differentiable on $\mathbb{C}$ with derivatives: $\sin' = \cos$, $\cos' = -\sin$, $\sinh' = \cosh$ and $\cosh' = \sinh$.

Real trigonometric functions. It follows from above that cos and sin are real-valued on $\mathbb{R}$, and moreover since $\sin^2 x + \cos^2 x = 1$ for all x, they take values in the interval $[-1, 1]$. Thus, they define functions $\cos, \sin\colon \mathbb{R} \to [-1, 1]$.

Definition of $\boldsymbol{\pi}$. We let $\pi/2$ be the smallest positive root of $\cos x = 0$.

Remark. Standard properties of sin follow easily: $\sin\left(\frac{\pi}{2}\right) = 1$; $\sin\pi = 0$; sin is strictly increasing on $[0, \pi/2]$, symmetric about the line $x = \pi/2$ and about the point $(\pi, 0)$ and is periodic with period 2π. We have $\sin x = 0$ if and only if $x \in \pi\mathbb{Z}$. Standard properties of cos now follow. In particular, $\cos x = 1$ if and only if $x \in 2\pi\mathbb{Z}$.

Theorem 15 (Periodicity of exp). Given $z, w \in \mathbb{C}$, we have

(i) $\exp z = 1$ if and only if $z \in 2\pi\mathrm{i}\mathbb{Z}$, and

(ii) $\exp z = \exp w$ if and only if $z - w \in 2\pi\mathrm{i}\mathbb{Z}$.

Remark. Other properties of $\sin, \cos$ follow. For example, given $x, y \in \mathbb{R}$, we have $\sin x = \sin y$ if and only if either $x = y + 2\pi n$ for some $n \in \mathbb{Z}$, or $x = \pi - y + 2\pi n$ for some $n \in \mathbb{Z}$.

Real hyperbolic functions. By Proposition 14, the restriction to $\mathbb{R}$ of the hyperbolic functions give functions $\sinh, \cosh\colon \mathbb{R} \to \mathbb{R}$.

The function cosh is a strictly increasing bijection from $[0, \infty)$ onto $[1, \infty)$, and an even function on $\mathbb{R}$, whereas sinh is an odd function and a strictly increasing bijection $\mathbb{R} \to \mathbb{R}$. Moreover, $\sinh x < \cosh x$ for all $x \in \mathbb{R}$, and $\frac{\sinh x}{\cosh x} \to 1$ as $x \to \infty$.

Further trigonometric and hyperbolic functions. Tangent is defined as $\tan z = \frac{\sin z}{\cos z}$ on the set $\mathbb{C} \setminus \{\frac{\pi}{2} + \pi n : n \in \mathbb{Z}\}$. It is complex differentiable by the quotient rule with $\tan' z = \sec^2 z = \frac{1}{\cos^2 z}$. The restriction to the real interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ is a strictly increasing differentiable bijection $\tan\colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$.

Hyperbolic tangent defined as $\tanh z = \frac{\sinh z}{\cosh z}$ for $z \in \mathbb{C} \setminus \{\frac{\pi\mathrm{i}}{2} + \pi\mathrm{i}n : n \in \mathbb{Z}\}$. On $\mathbb{R}$ this gives a strictly increasing differentible bijection $\tanh\colon \mathbb{R} \to (-1, 1)$.

Higher derivatives. A function f is *twice differentiable at a* if f is defined and differentiable on an open interval I with $a \in I$ and whose derivative $f'\colon I \to \mathbb{R}$ is differentiable at a. We write $f''(a)$ for the derivative of f' at a, and call it the *second derivative of f at a*. A function f on an open interval I is said to be *twice differentiable on I* if f is differentiable on I, and its derivative $f'\colon I \to \mathbb{R}$ is also differentiable on I. In that case f is twice differentiable at a for every $a \in I$. We then denote by $f''\colon I \to \mathbb{R}$ the map $a \mapsto f''(a)$, and call it the *second derivative of f on I*. Note that $f'' = (f')'$ is the derivative of f' on I.

Remark. In the definitions above we could replace open intervals with more general open sets.

Example. The derivative of a real polynomial on $\mathbb{R}$ is again a real polynomial on $\mathbb{R}$, and hence differentiable on $\mathbb{R}$. Thus, real polynomials are twice differentiable. Similarly, the derivative of a rational function is again a rational function with the same domain, and hence rational functions are also twice differentiable.

Definition. For each $n \in \mathbb{N}$, we recursively define the following four concepts: *f is n-times differentiable at a, the n^{th} derivative $f^{(n)}(a)$ of f at a, f is n-times differentiable on I, the n^{th} derivative $f^{(n)}\colon I \to \mathbb{R}$ of f on I.*

In the case $n = 1$, being n-times differentiable simply means being differentiable, and the n^{th} derivative is just the derivative. Thus, $f^{(1)}(a) = f'(a)$ and $f^{(1)} = f'$. Now let $n > 1$. We then say *f is n-times differentiable at a* if f is defined and is $(n-1)$-times differentiable on an open interval I with $a \in I$ and whose $(n-1)^{\text{th}}$ derivative $f^{(n-1)}\colon I \to \mathbb{R}$ is differentiable at a. We write $f^{(n)}(a)$ for the derivative of $f^{(n-1)}$ at a, and call it the *n^{th} derivative of f at a*. A function f on an open interval I is said to be *n-times differentiable on I* if f is $(n-1)$-times differentiable on I and its $(n-1)^{\text{th}}$ derivative $f^{(n-1)}\colon I \to \mathbb{R}$ is also differentiable on I. In that case f is n-times differentiable at a for every $a \in I$. We then denote by $f^{(n)}\colon I \to \mathbb{R}$ the map $a \mapsto f^{(n)}(a)$, and call it the *n^{th} derivative of f on I*. Note that $f^{(n)} = \big(f^{(n-1)}\big)'$ on I.

Remark. Given $m, n \in \mathbb{N}$, a function f is $(m+n)$-times differentiable at a if and only if f is m-times differentiable on an open interval containing a and $f^{(m)}$ is n-times differentiable at a. In that case $f^{(m+n)}(a) = \big(f^{(m)}\big)^{(n)}(a)$.

Terminology. For an integer $n \geqslant 0$, we say a function *f is a C^n-function* or that *f is n-times continuously differentiable* if f is n-times differentiable and $f^{(n)}$ is continuous. By convention, a C^0-function is simply a continuous function. We say *f is a C^∞-function* or a *smooth function* or that *f is infinitely differentiable* if f is a C^n-function for all n.

Examples. Polynomials, rational functions, exp, log, the trigonometric and hyperbolic functions are all smooth. Sums, products, quotients and composites of smooth functions are smooth.

Remark. In the above definitions of higher derivatives we could replace "interval" by "disc" and "real-valued" by "comlex-valued" to obtain analogous concepts of higher complex differentiability. It is an absolutely amazing fact (see the Part IB *Complex Analysis* course) that if U is an open subset of $\mathbb{C}$, and $f\colon U \to \mathbb{C}$ is complex differentiable, then automatically f is n-times complex differentiable for all n! For the rest of the chapter we work with real functions.

Motivation. If f is a polynomial of degree at most n, then we have $f(a+h) = f(a) + \sum\limits_{k=1}^{n} \frac{f^{(k)}(a)}{k!} h^k$ for all $a, h \in \mathbb{R}$. Consider in particular the case $n = 0$ and $n = 1$, and compare with the following.

$$\begin{aligned} f \text{ is continuous at } a &\iff f(a+h) = f(a) + \mathrm{o}(1) \\ f \text{ is differentiable at } a &\iff f(a+h) = f(a) + f'(a)h + \mathrm{o}(h) \end{aligned}$$

Theorem 16 (Taylor's theorem with Peano's form of the remainder). Assume that f is n-times differentiable at a for some $n \in \mathbb{N}$. Then

$$f(a+h) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} h^k + \mathrm{o}(h^n) \tag{2}$$

for all h in some open interval containing 0.

Definition. Assume f is a C^∞-function on some open interval containing a. The polynomial $\sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!} h^k$ is *the n^{th}-order Taylor polynomial of f at a*, and the series $\sum\limits_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} h^n$ is called *the Taylor series of f at a*.
(Here we use the usual convention that $f^{(0)} = f$.)

Question. Let f be as in the previous definition. Then (2) remains valid for all $n \in \mathbb{N}$. It is then natural to ask if there exists some $\delta > 0$ such that the Taylor series of f at a converges to $f(a+h)$ whenever $|h| < \delta$.

Example. Let $f(x) = \exp x$, $x \in \mathbb{R}$. Then $f^{(n)}(0) = f(0) = 1$ for all $n \in \mathbb{N}$. Thus, the Taylor series of exp at 0 is $\sum\limits_{n=0}^{\infty} \frac{h^n}{n!}$, which does converge to exp on the whole of $\mathbb{R}$.

Remark. The answer to the above question is 'no' in general. Let us define *the n^{th} remainder term R_n of f at a* by

$$f(a+h) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} h^k + R_n(h) \ .$$

Theorem 16 implies that $\lim\limits_{h \to 0} R_n(h) = 0$ for *fixed* $n \in \mathbb{N}$. For the Taylor series to converge to f, we would need $\lim\limits_{n \to \infty} R_n(h) = 0$ for each *fixed* h in some open interval containing 0. The latter does not hold if *e.g.,* $R_n(h) = nh^{n+1}$, even though in this case $R_n(h) = \mathrm{o}(h^n)$ for each fixed $n \in \mathbb{N}$.

Theorem 17 (Taylor's theorem with Lagrange's form of the remainder). Let $n \in \mathbb{N}$, $a, r \in \mathbb{R}$ with $r > 0$. Assume that f is n-times differentiable on the open interval $(a-r, a+r)$. Then for every h with $0 < |h| < r$ there exists $\theta \in (0,1)$ such that

$$f(a+h) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(n)}(a+\theta h)}{n!} h^n \ . \tag{3}$$

Remarks. 1. The case $n = 1$ is the Mean Value Theorem. So we can think of the above as a generalized Mean Value Theorem. Accordingly, we shall prove it via a generalized Rolle's theorem.

2. Here $R_n(h) = \frac{f^{(n)}(a+\theta h) - f^{(n)}(a)}{n!} h^n$. One cannot deduce Theorem 16 from this without some further assumptions (*e.g.,* continuity of $f^{(n)}$ at a is enough).

Application: the Binomial Theorem. Fix $\alpha \in \mathbb{R}$. Define $f \colon (-1, \infty) \to \mathbb{R}$ by $f(x) = (1+x)^\alpha$. Then the Taylor series of f at 0 converges to f on $(-1,1)$:

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \frac{\alpha^{\underline{n}}}{n!} x^n \qquad \text{whenever } |x| < 1 \ ,$$

where $\alpha^{\underline{0}} = 1$ and $\alpha^{\underline{n}} = \alpha(\alpha-1)\cdots(\alpha-n+1)$ for $n \geqslant 1$. Here we prove this only in the range $-\frac{1}{2} < x < 1$ and postpone the full proof until the last chapter.

5 Power Series

A series of the form $\sum\limits_{n=0}^{\infty} a_n(z-a)^n$ is called a *power series about* a, where $a, z \in \mathbb{C}$ and $a_n \in \mathbb{C}$ for all n. We think of a and $(a_n)_{n=0}^{\infty}$ as fixed, and think of z as a variable. Set $D = \{z \in \mathbb{C} : \sum a_n(z-a)^n \text{ converges}\}$, and define $f\colon D \to \mathbb{C}$ by $f(z) = \sum\limits_{n=0}^{\infty} a_n(z-a)^n$. Note that $a \in D$ always holds.

Examples. 1. $\sum\limits_{n=0}^{\infty} \frac{z^n}{n!}$ is a power series about 0, $D = \mathbb{C}$, $f = \exp$.

2. $\sum\limits_{n=1}^{\infty} \left(-\frac{1}{n}\right)(1-z)^n$ is a power series about 1. Indeed, this is the same as the series $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(z-1)^n$. So here $a = 1$, $a_n = \frac{(-1)^{n-1}}{n}$ for $n \geqslant 1$, and $a_0 = 0$. It follows from the Ratio Test and Dirichlet's Test that $D = \overline{D}(1,1) \setminus \{0\}$.

3. $\sum\limits_{n=0}^{\infty} n^n z^n$ is a power series with $D = \{0\}$.

Lemma 1. Assume that the power series $\sum_{n=0}^{\infty} a_n z^n$ converges if $z = w$ for some $w \in \mathbb{C}$. Then the power series converges absolutely for all z with $|z| < |w|$.

Conventions. Let A be a non-empty subset of $\mathbb{R}$. We know that if A is bounded above then A has a supremum which is unique and denoted by $\sup A$. If A is not bounded above we shall write $\sup A = \infty$. We also declare that $x < \infty$ is true for all $x \in \mathbb{R}$, and $\infty < x$ is true for no $x \in \mathbb{R}$.

Definition. The *radius of convergence* R *of a power series* $\sum\limits_{n=0}^{\infty} a_n(z-a)^n$ is defined by $R = \sup\{|z-a| : \sum a_n(z-a)^n \text{ converges}\}$.

Examples. The power series $\sum_{n=0}^{\infty} \frac{z^n}{n!}$, $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(z-1)^n$ and $\sum_{n=0}^{\infty} n^n z^n$ have radius of convergence $R = \infty$, $R = 1$ and $R = 0$, respectively.

Theorem 2. If $\sum\limits_{n=0}^{\infty} a_n(z-a)^n$ has radius of convergence R, then the power series converges absolutely whenever $|z-a| < R$ and diverges if $|z-a| > R$.

Examples. The power series $\sum z^n$, $\sum z^n/n$ and $\sum z^n/n^2$ all have radius of convergence $R = 1$. Thus, for each of these series, the set D of values z at which the series converges is of the form $D = D(0,1) \cup E$ for some $E \subset \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. Indeed we have $E = \emptyset$, $E = \mathbb{T} \setminus \{1\}$ and $E = \mathbb{T}$, respectively.

Theorem 3. Let R be the radius of convergence of the power $\sum\limits_{n=0}^{\infty} a_n(z-a)^n$. Then $f\colon D(a,R) \to \mathbb{C}$, $f(z) = \sum\limits_{n=0}^{\infty} a_n(z-a)^n$, is complex differentiable and $f'(z) = \sum\limits_{n=1}^{\infty} na_n(z-a)^{n-1}$. It follows that f is k-times complex differentiable for all $k \in \mathbb{N}$, and $f^{(k)}(z) = \sum\limits_{n=k}^{\infty} n^{\underline{k}} a_n(z-a)^{n-k}$.

6 Integration

Let $[a, b]$ be a closed and bounded interval in $\mathbb{R}$, and let $f\colon [a, b] \to \mathbb{R}$ be a bounded function (*i.e.,* there exists $C \in \mathbb{R}$ such that $|f(t)| \leqslant C$ for all $t \in [a, b]$). A *dissection of* $[a, b]$ is a finite sequence $\mathcal{D}$ of the form $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ where $n \in \mathbb{N}$. For each $k = 1, \dots, n$ we set

$$m_k(f) = \inf\{f(t) : t \in [x_{k-1}, x_k]\} \quad \text{and} \quad M_k(f) = \sup\{f(t) : t \in [x_{k-1}, x_k]\} .$$

We then define *the lower (Riemann) sum $L_{\mathcal{D}}(f)$ of f with respect to $\mathcal{D}$* by

$$L_{\mathcal{D}}(f) = \sum_{k=1}^{n} m_k(f) \cdot (x_k - x_{k-1}) ,$$

and *the upper (Riemann) sum $U_{\mathcal{D}}(f)$ of f with respect to $\mathcal{D}$* by

$$U_{\mathcal{D}}(f) = \sum_{k=1}^{n} M_k(f) \cdot (x_k - x_{k-1}) .$$

Note. $L_{\mathcal{D}}(f) \leqslant U_{\mathcal{D}}(f)$.

Given dissections $\mathcal{D}$ and $\mathcal{D}'$ of $[a, b]$, we say *$\mathcal{D}'$ is a refinement of $\mathcal{D}$*, and we write $\mathcal{D} \leqslant \mathcal{D}'$, if $\mathcal{D}'$ is obtained from $\mathcal{D}$ by the addition of extra dissection points. Formally, $\mathcal{D} \leqslant \mathcal{D}'$ if $\mathcal{D}$ is a subsequence of $\mathcal{D}'$. Thus, if $\mathcal{D}'$ is of the form $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ for some $n \in \mathbb{N}$, then $\mathcal{D}$ is of the form $a = x_{k_0} < x_{k_1} < x_{k_2} < \cdots < x_{k_m} = b$ for some $1 \leqslant m \leqslant n$ and $0 = k_0 < k_1 < \cdots < k_m = n$.

Lemma 1. If $\mathcal{D} \leqslant \mathcal{D}'$, then $L_{\mathcal{D}}(f) \leqslant L_{\mathcal{D}'}(f) \leqslant U_{\mathcal{D}'}(f) \leqslant U_{\mathcal{D}}(f)$.

Corollary 2. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two arbitrary dissections of $[a, b]$. Then $L_{\mathcal{D}_1}(f) \leqslant U_{\mathcal{D}_2}(f)$.

Notation. We denote by $\mathcal{D}_1 \vee \mathcal{D}_2$ the *common refinement* of $\mathcal{D}_1$ and $\mathcal{D}_2$.

Definitions. Let $f\colon [a, b] \to \mathbb{R}$ be a bounded function on a closed and bounded interval $[a, b]$. The *lower (Riemann) integral of f over* $[a, b]$ is defined as

$$\underline{\int_a^b} f = \sup_{\mathcal{D}} L_{\mathcal{D}}(f)$$

and the *upper (Riemann) integral of f over* $[a, b]$ is defined as

$$\overline{\int_a^b} f = \inf_{\mathcal{D}} U_{\mathcal{D}}(f)$$

where the sup and inf above are taken over all dissections $\mathcal{D}$ of $[a, b]$. By Corollary 2, these are well defined and satisfy

$$\underline{\int_a^b} f \leqslant \overline{\int_a^b} f .$$

We say f *is Riemann integrable over* $[a,b]$ if $\underline{\int_a^b} f = \overline{\int_a^b} f$ and the common value is *the Riemann integral of* f *over* $[a,b]$ and is denoted by

$$\int_a^b f \qquad \text{or} \qquad \int_a^b f(t)\,\mathrm{d}t\ .$$

Examples.

1. The constant function $f(t) = c$ is integrable over $[a,b]$ with $\int_a^b f = (b-a)c$.

2. The function $f(t) = t$ for all t is integrable over $[0,1]$ with $\int_0^1 t\,\mathrm{d}t = \frac{1}{2}$.

3. Let $f(t) = 1$ if $t \in \mathbb{Q}$ and $f(t) = 0$ otherwise. Then $\underline{\int_0^1} f = 0 \neq 1 = \overline{\int_0^1} f$. Thus f is not integrable over $[0,1]$.

4. Let $f(t) = \frac{1}{t}$ for $t \neq 0$ and $f(0) = 0$. Then f is not integrable over $[0,1]$ because f is not bounded on $[0,1]$.

Proposition 3. Assume f is Riemann integrable over $[a,b]$. Then

$$(b-a)\inf_{[a,b]} f \leqslant \int_a^b f \leqslant (b-a)\sup_{[a,b]} f \qquad \text{and} \qquad \left|\int_a^b f\right| \leqslant (b-a)\sup_{[a,b]}|f|\ .$$

Theorem 4 (Riemann's criterion). Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over $[a,b]$ if and only if for all $\varepsilon > 0$ there exists a dissection $\mathcal{D}$ of $[a,b]$ such that $U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f) < \varepsilon$.

Corollary 5. Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over $[a,b]$ if and only if there exists a sequence $(\mathcal{D}_n)$ of dissections of $[a,b]$ such that $U_{\mathcal{D}_n}(f) - L_{\mathcal{D}_n}(f) \to 0$ as $n \to \infty$.

Remarks. 1. Assume that $(\mathcal{D}_n)$ is a sequence of dissections of $[a,b]$ such that $U_{\mathcal{D}_n}(f) - L_{\mathcal{D}_n}(f) \to 0$ as $n \to \infty$. Then both $L_{\mathcal{D}_n}(f) \to \int_a^b f$ and $U_{\mathcal{D}_n}(f) \to \int_a^b f$ as $n \to \infty$. It follows that if $\mathcal{D}_n$ is the sequence $a = x_0^{(n)} < x_1^{(n)} < \dots < x_{m_n}^{(n)} = b$ and $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$ for $1 \leqslant k \leqslant m_n$, then the *Riemann sum*

$$\sum_{k=1}^{m_n} f(\xi_k^{(n)})(x_k^{(n)} - x_{k-1}^{(n)}) \quad \text{converges to} \quad \int_a^b f(t)\,\mathrm{d}t \quad \text{as} \quad n \to \infty\ .$$

2. Given a dissection $\mathcal{D}$: $a = x_0 < \dots < x_n = b$ of $[a,b]$, we define the *mesh* $|\mathcal{D}|$ *of* $\mathcal{D}$ by $|\mathcal{D}| = \max_{1 \leqslant k \leqslant n}(x_k - x_{k-1})$. A theorem of Darboux says that if f is integrable, and $(\mathcal{D}_n)$ is a sequence of dissections of $[a,b]$ such that $|\mathcal{D}_n| \to 0$ as $n \to \infty$, then $U_{\mathcal{D}_n}(f) - L_{\mathcal{D}_n}(f) \to 0$ as $n \to \infty$. (See Examples Sheet 4.)

3. If f is integrable then f cannot oscillate too much too often. Indeed, given $\varepsilon > 0$, choose a dissection $\mathcal{D}$ such that

$$U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f) = \sum_{k=1}^{n}(M_k(f) - m_k(f)) \cdot (x_k - x_{k-1}) < \varepsilon^2\ .$$

Then the total length of intervals $[x_{k-1}, x_k]$ where $M_k(f) - m_k(f) > \varepsilon$ is less than ε.

Lemma 6. Let $f, g\colon [a,b] \to \mathbb{R}$ be bounded functions. Assume that there exists $K \geqslant 0$ such that $|f(x) - f(y)| \leqslant K|g(x) - g(y)|$ for all $x, y \in [a,b]$. Then if g is integrable over $[a,b]$, then so is f.

Remark. The condition says that up to a constant K, the function f never oscillates more than g.

Exercise. Given an arbitrary set A and a function $h\colon A \to \mathbb{R}$, show that

$$\sup_A h - \inf_A h = \sup\{|h(s) - h(t)| : s, t \in A\} .$$

This is used in the proof of the above lemma.

Theorem 7. Let $f, g\colon [a,b] \to \mathbb{R}$ be functions that are integrable over $[a,b]$. Let $\lambda, \mu \in \mathbb{R}$.

(i) (Linearity) $\lambda f + \mu g$ is integrable and $\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g$.

(ii) (Monotonicity) If $f \leqslant g$ on $[a,b]$, then $\int_a^b f \leqslant \int_a^b g$.

(iii) $|f|$ is integrable and $\left| \int_a^b f \right| \leqslant \int_a^b |f|$.

(iv) $\min(f,g)$ and $\max(f,g)$ are integrable.

(v) (Cauchy–Schwarz) fg is integrable and $\left| \int_a^b fg \right| \leqslant \left(\int_a^b f^2 \right)^{1/2} \cdot \left(\int_a^b g^2 \right)^{1/2}$.

Proposition 8.

(i) Assume that $h(x) = 0$ for all but finitely many $x \in [a,b]$. Then h is integrable over $[a,b]$ and $\int_a^b h(t)\,\mathrm{d}t = 0$.

(ii) Given functions $f, g\colon [a,b] \to \mathbb{R}$, if f is integrable and $g(x) = f(x)$ for all but finitely many $x \in [a,b]$, then g is integrable and $\int_a^b g = \int_a^b f$.

Theorem 9. If f is continuous on $[a,b]$, then f is integrable over $[a,b]$.

Theorem 10. If f is monotonic on $[a,b]$, then f is integrable over $[a,b]$.

Proposition 11. Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function.

(i) Given $a < c < b$, if f is integrable over $[a,c]$ and $[c,b]$, then f is integrable over $[a,b]$ and $\int_a^b f = \int_a^c f + \int_c^b f$.

(ii) Given $a \leqslant c < d \leqslant b$, if f is integrable over $[a,b]$, then f is integrable over $[c,d]$.

Corollary 12. Let $a = c_0 < c_1 < \cdots < c_k = b$ be a dissection of $[a,b]$ and f be a bounded function on $[a,b]$. Then f is integrable over $[a,b]$ if and only if f is integrable over $[c_{j-1}, c_j]$ for each $j = 1, 2, \ldots, k$, and in that case $\int_a^b f = \sum_{j=1}^k \int_{c_{j-1}}^{c_j} f$.

Corollary 13. A piecewise monotonic function on $[a, b]$ is integrable over $[a, b]$.

Theorem 14. Let f be a bounded function on $[a, b]$ that is continuous at all but finitely many $x \in [a, b]$. Then f is integrable over $[a, b]$.

Definition. Given $a < b$ and an integrable function f on $[a, b]$, we define $\int_b^a f = -\int_a^b f$. We also set $\int_a^a f = 0$ (although this already follows from the definition of Riemann integral).

Note. 1. If f is defined and is integrable on a closed and bounded interval containing a and b, then

$$\left| \int_a^b f \right| \leqslant |b - a| \cdot \sup|f| \ .$$

2. If f is defined and is integrable on a closed and bounded interval containing a, b and c, then

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

Definition. Let $f\colon [a, b] \to \mathbb{R}$ be an integrable function, and let $c \in [a, b]$. The function $F\colon [a, b] \to \mathbb{R}$ defined by $F(x) = \int_c^x f(t)\,\mathrm{d}t$, is called *an indefinite integral of* f.

Note. $F(y) - F(x) = \int_x^y f(t)\,\mathrm{d}t$ for all $x, y \in [a, b]$.

Theorem 15. Let f be an integrable function over $[a, b]$ and F be an indefinite integral of f. Then F is continuous on $[a, b]$. In fact, there exists $K \geqslant 0$ such that $|F(y) - F(x)| \leqslant K|y - x|$ for all $x, y \in [a, b]$.

Theorem 16 (Fundamental theorem of calculus). Let f be an integrable function over $[a, b]$ and F be an indefinite integral of f. Given $x \in [a, b]$, if f is continuous at x, then F is differentiable at x and $F'(x) = f(x)$.

Remark. The derivative of F at x is interpreted in a one-sided sense at the endpoints. Thus,

$$F'(a) = \lim_{x \to a^+} \frac{F(x) - F(a)}{x - a} \qquad \text{and} \qquad F'(b) = \lim_{x \to b^-} \frac{F(x) - F(b)}{x - b}$$

when these limits exist.

Definition. Let I be an interval. Given functions $f, F\colon I \to \mathbb{R}$, we say F *is an antiderivative of* f *on* I if F is differentiable on I (in the one-sided sense at any endpoint) and $F'(x) = f(x)$ for all $x \in I$.

Corollary 17. A continuous function $f\colon [a, b] \to \mathbb{R}$ has an antiderivative on $[a, b]$. If G is any such antiderivative, then $\int_a^b f(t)\,\mathrm{d}t = G(b) - G(a)$.

Note. It follows from the above result, that given a continuous function $f\colon [a, b] \to \mathbb{R}$ and $y_0 \in \mathbb{R}$, the initial value problem $\frac{\mathrm{d}y}{\mathrm{d}t} = f$, $y(a) = y_0$, has a unique solution on $[a, b]$.

Theorem 18. If f is integrable over $[a, b]$ and has an antiderivative F on $[a, b]$, then $\int_a^b f = F(b) - F(a)$.

Remark. This generalizes Corollary 17. However, the assumption on the existence of an antiderivative here is necessary as not every integrable function has an antiderivative.

Corollary 19 (Integration by parts). Let $f, g\colon [a, b] \to \mathbb{R}$ be integrable over $[a, b]$ with antiderivatives F and G, respectively, on $[a, b]$. Then

$$\int_a^b f \cdot G = F(b)G(b) - F(a)G(a) - \int_a^b F \cdot g \ .$$

Corollary 20 (Integration by substitution). Let $\varphi\colon [a, b] \to \mathbb{R}$ be a continuously differentiable function, and let f be a continuous function on a closed bounded interval that contains the image of φ. Then

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y = \int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x \ .$$

Theorem 21 (Taylor's theorem with the integral form of the remainder). Let I be an open interval, let $n \in \mathbb{N}$ and let $f\colon I \to \mathbb{R}$ be an n-times continuously differentiable function on I. Given $a \in I$, for every h with $a+h \in I$ we have

$$f(a+h) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}h^k + \frac{1}{(n-1)!}\int_0^h (h-t)^{n-1} f^{(n)}(a+t)\,\mathrm{d}t \ .$$

Remark. We can write the above expression as

$$f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k - \frac{f^{(n)}(a)}{n!}h^n + \frac{1}{(n-1)!}\int_0^h (h-t)^{n-1} f^{(n)}(a+t)\,\mathrm{d}t \ .$$

Note that

$$\frac{1}{(n-1)!}\int_0^h (h-t)^{n-1}\,\mathrm{d}t = \frac{h^n}{n!} \ .$$

It follows that the n^{th} remainder term R_n of f at a is given by

$$R_n(h) = \frac{1}{(n-1)!}\int_0^h (h-t)^{n-1}(f^{(n)}(a+t) - f^{(n)}(a))\,\mathrm{d}t \ .$$

Theorem 22 (Binomial Theorem). Let $\alpha \in \mathbb{R}$. Then

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad \text{whenever } -1 < x < 1 \ .$$

where $\binom{\alpha}{n} = \frac{\alpha^{\underline{n}}}{n!}$ for all $n \geqslant 0$.

Improper integrals. Assume that we are given $a < b$ in $\mathbb{R}$ and a function $f\colon (a,b] \to \mathbb{R}$ that is integrable over $[c,b]$ for all c with $a < c < b$. Note that in particular f is bounded on $[c,b]$ for every c with $a < c < b$, but f need not be bounded on $(a,b]$. We say that *the integral* $\int_a^b f$ *converges* or that *the integral* $\int_a^b f$ *exists in the improper sense* if the limit $\lim_{c\to a^+} \int_c^b f$ exists. In that case we set

$$\int_a^b f = \lim_{c\to a^+} \int_c^b f$$

and call this value the *improper integral of* f *over* $(a,b]$.

Example. $\int\limits_0^1 x^\alpha \,\mathrm{d}x$ converges if and only if $\alpha > -1$ in which case $\int\limits_0^1 x^\alpha \,\mathrm{d}x = \frac{1}{\alpha+1}$.

Exercise. If $f\colon (a,b] \to \mathbb{R}$ is bounded on $(a,b]$ and integrable over $[c,b]$ for all c with $a < c < b$, then $\int_a^b f$ exists in the improper sense. Moreover, for any $\lambda \in \mathbb{R}$, the function

$$g(x) = \begin{cases} \lambda & \text{if } x = a \ , \\ f(x) & \text{if } a < x \leqslant b \end{cases}$$

is integrable over $[a,b]$ and $\int_a^b g = \int_a^b f$.

More improper integrals. Improper integrals over intervals of the form $[a,b)$ with $a < b$ in $\mathbb{R}$ are defined in the obvious way. If $f\colon (a,b) \to \mathbb{R}$ is integrable over $[c,d]$ whenever $a < c < d < b$, then we say that the integral $\int_a^b f$ converges or that the integral $\int_a^b f$ exists in the improper sense if $\int_c^d f$ converges as $c \to a^+$ and $d \to b^-$. In that case we set

$$\int_a^b f = \lim_{\substack{c\to a^+\\ d\to b^-}} \int_c^d f$$

and call this value the improper integral of f over (a,b).

Note. The existence of this double limit means the following. Assume that g is a (real or complex) function defined at each pair (c,d) with $a < c < d < b$. Given $\lambda \in \mathbb{R}$, we say $g(c,d)$ converges to λ as $c \to a^+$ and $d \to b^-$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $a < c < a+\delta$, $b-\delta < d < b$ and $c < d$, then $|g(c,d) - \lambda| < \varepsilon$. When such λ exists it is unique and is denoted by

$$\lim_{\substack{c\to a^+\\ d\to b^-}} g(c,d) \ .$$

Exercise. Assume that the function $f\colon (a,b) \to \mathbb{R}$ is integrable over $[c,d]$ whenever $a < c < d < b$. Show that $\int_a^b f$ converges if and only for *some* $c \in (a,b)$ (equivalently, for *all* $c \in (a,b)$) the integrals $\int_a^c f$ and $\int_c^b f$ exist in the improper sense. Moreover in this case we have $\int_a^b f = \int_a^c f + \int_c^b f$.

Yet more improper integrals. The following is a list of further types of improper integrals. The formulation of the exact definitions and conditions is left as an exercise:

- The improper integral of f over $[a,\infty)$ is $\displaystyle\int_a^\infty f = \lim_{b\to\infty}\int_a^b f$.

- The improper integral of f over $(-\infty,\infty)$ is $\displaystyle\int_{-\infty}^\infty f = \lim_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b f$.

- The improper integral of f over (a,∞) is $\displaystyle\int_a^\infty f = \lim_{\substack{b\to a^-\\ c\to\infty}}\int_b^c f$.

Moreover, the third integral converges if and only for *some* $c > a$ (equivalently, for *all* $c > a$) the integrals $\int_a^c f$ and $\int_c^\infty f$ exist in the improper sense and in that case $\int_a^\infty f = \int_a^c f + \int_c^\infty f$. A similar result holds for the second integral on the list.

Remark. There is some analogy between the improper integral $\int_a^\infty f$ for a function $f\colon [a,\infty)\to\mathbb{R}$ on the one hand, and a series $\sum_{n=1}^\infty a_n$ on the other hand. Indeed, the two are related via the integral test for convergence (see Examples Sheet 4). Also, there is a comparison test for the improper integral (see Theorem 24 below). However, not everything works. *E.g.,* the existence of $\int_1^\infty f$ does not imply that $f(t)\to 0$ as $t\to\infty$ (again, see Examples Sheet 4).

Lemma 23. We are given a function F defined on (a,∞) for some $a\in\mathbb{R}$. If the limit $\lim_{t\to\infty} F(t)$ exists, then for every sequence such that $x_n\to\infty$, the sequence $(F(x_n))$ converges to $\lim_{t\to\infty} F(t)$. Conversely, if $(F(x_n))$ is convergent for every sequence such that $x_n\to\infty$, then $\lim_{t\to\infty} F(t)$ exists.

Remarks. 1. In the converse statement, we do not assume that the limit $\lim_{n\to\infty} F(x_n)$ is always the same. However, that is true, and indeed the first step of our proof was to show that.

2. Similar results hold for other types of limits, *e.g.,* for $\lim_{x\to a^+} F(x)$ where F is defined on $(a,b]$ for some $b > a$.

Theorem 24 (A comparison test for improper integrals). Let $a\in\mathbb{R}$ and let $f,g\colon [a,\infty)\to\mathbb{R}$ be functions that are integrable over $[a,b]$ for all $b > a$. Assume that $|f(x)|\leqslant g(x)$ for all $x\geqslant a$. Then if $\int_a^\infty g$ exists in the improper sense then so does $\int_a^\infty f$, and moreover $|\int_a^\infty f|\leqslant\int_a^\infty g$.

Exercise. Formulate and prove comparison tests for other types of improper integrals.