Contents

Chapter 1

The Real Number System

1.1 Definition and Properties of Real Numbers

We will give an axiomatic definition of the real numbers. We will define the set of real numbers, denoted by $\mathbb{R}$, as a set with two operations $+$ and $\cdot$, called addition and multiplication, as well as an ordering $<$, which satisfy the laws of addition (A), the laws of multiplication (M), the distributive law (D), the order laws (O) and the Dedekind Completeness Axiom (C). These laws and axioms will be given and discussed below.

Additional Note. An alternative way to define the real numbers is to use Peano's axiom to define the natural numbers, and then construct the integers and rational numbers in turn. Then the axioms (A), (M), (D) and (O) become properties of rational numbers, so that we call them laws.

Addition and multiplication are maps which assign to every two elements in $a, b \in \mathbb{R}$ an element in $\mathbb{R}$ which is denoted by $a+b$ and $a \cdot b$ (in general, written ab), respectively. We require that these operations satisfy the following axioms.

A. Axioms of addition

(A1) **Associative Law:** $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{R}$.

(A2) **Commutative Law:** $a + b = b + a$ for all $a, b \in \mathbb{R}$.

(A3) **Zero:** There is a real number 0 such that $a + 0 = a$ for all $a \in \mathbb{R}$.

(A4) **Additive inverse:** For each $a \in \mathbb{R}$ there is $-a \in \mathbb{R}$ such that $a + (-a) = 0$.

Notation. For $a, b \in \mathbb{R}$ one writes $a - b := a + (-b)$.

Additional Notes. 1. Any set with the operation $+$ satisfying (A1), (A3), (A4) is called a group. If also (A2) is satisfied, the group is called an Abelian (or commutative) group.
2. You will encounter detailed discussions of (Abelian) groups in algebra courses.
3. The additive inverse of a number is also called the negation of a number.

Theorem 1.1 (Basic group properties)**.**
(a) *The number* 0 *is unique.*
(b) *For all* $a \in \mathbb{R}$*, the number* $-a$ *is unique.*
(c) *For all* $a, b \in \mathbb{R}$*, the equation* $a + x = b$ *has a unique solution. This solution is* $x = b - a$.
(d) $\forall a \in \mathbb{R},\ -(-a) = a$.
(e) $\forall a, b \in \mathbb{R},\ -(a + b) = -a - b$.
(f) $-0 = 0$.

Proof. (a) Let $0, 0' \in \mathbb{R}$ such that $a + 0 = a$ and $a + 0' = a$ for all $a \in \mathbb{R}$. We must show that $0 = 0'$:

$$\begin{aligned} 0 &= 0 + 0' \\ &= 0' + 0 \qquad \text{by (A2)} \\ &= 0' \end{aligned}$$

(b) Let $a \in \mathbb{R}$ and $a', a'' \in \mathbb{R}$ such that $a + a' = 0$ and $a + a'' = 0$. We must show that $a' = a''$:

$$\begin{aligned} a' &= a' + 0 && \text{by (A3)} \\ &= a' + (a + a'') \\ &= (a' + a) + a'' && \text{by (A1)} \\ &= (a + a') + a'' && \text{by (A2)} \\ &= 0 + a'' \\ &= a'' + 0 && \text{by (A2)} \\ &= a'' && \text{by (A3)} \end{aligned}$$

(c) First we show that $x = b - a$ is a solution. So let $x = b - a$.
Then

$$\begin{aligned} a + x = a + (b - a) &= a + (b + (-a)) \\ &= a + ((-a) + b) && \text{by (A2)} \\ &= (a + (-a)) + b && \text{by (A1)} \\ &= 0 + b && \text{by (A4)} \\ &= b + 0 && \text{by (A2)} \\ &= b && \text{by (A3)} \end{aligned}$$

To show that the solution is unique let $x \in \mathbb{R}$ such that $a + x = b$.
Then

$$\begin{aligned} x = x + 0 &= x + (a + (-a)) && \text{by (A3), (A4)} \\ &= (x + a) + (-a) && \text{by (A1)} \\ &= (a + x) + (-a) && \text{by (A2)} \\ &= b - a && \because a + x = b \end{aligned}$$

This shows that the solution is unique.
(d) Note that

$$-a + (-(-a)) = 0 \quad \text{by (A4).}$$

On the other hand

$$-a + a = a + (-a) = 0 \quad \text{by (A2), (A4).}$$

By part (b), it follows that

$$-(-a) = a.$$

Alternatively:

$$\begin{aligned} -(-a) = -(-a) + 0 &= -(-a) + (a + (-a)) && \text{by (A3), (A4)} \\ &= (a + (-a)) + (-(-a)) && \text{by (A2)} \\ &= a + ((-a) + (-(-a))) && \text{by (A1)} \\ &= a + 0 && \text{by (A4)} \\ &= a && \text{by (A3)} \end{aligned}$$

(e)

$$\begin{aligned} (a + b) + (-a - b) &= (a + b) + ((-a) + (-b)) \\ &= ((b + a) + (-a)) + (-b) && \text{by (A1), (A2)} \\ &= (b + (a + (-a))) + (-b) && \text{by (A1)} \\ &= (b + 0) + (-b) && \text{by (A4)} \\ &= b + (-b) && \text{by (A3)} \\ &= 0 && \text{by (A4)} \end{aligned}$$

By part (b), $-(a + b) = -a - b$.
(f) $0 + 0 = 0$ by (A3), so that $-0 = 0$ by (A4) and (b). $\square$

M. Axioms of multiplication

(M1) **Associative Law:** $a(bc) = (ab)c$ for all $a, b, c \in \mathbb{R}$.

(M2) **Commutative Law:** $ab = ba$ for all $a, b \in \mathbb{R}$.

(M3) **One:** There is a real number 1 such that $1 \neq 0$ and $a \cdot 1 = a$ for all $a \in \mathbb{R}$.

(M4) **Multiplicative inverse:** For each $a \in \mathbb{R}$ with $a \neq 0$ there is $a^{-1} \in \mathbb{R}$ such that $aa^{-1} = 1$.

D. The distributive law axiom

(D) **Distributive Law:** $a(b + c) = ab + ac$ for all $a, b, c \in \mathbb{R}$.

Additional Notes. 1. Any set with the operations $+$, $\cdot$ satisfying the axioms (A1)–(A4), (M1)–(M4) and (D) is called a field.
2. You will encounter detailed discussions of fields in algebra courses.
3. The multiplicative inverse of a number is also briefly called the inverse of a number.
4. The set of nonzero real numbers, $\mathbb{R} \setminus \{0\}$, is an Abelian group with respect to multiplication.

Theorem 1.2 (Basic field properties: Distributive laws)**.**
(a) $\forall a, b, c \in \mathbb{R},\ (a + b)c = ac + bc.$
(b) $\forall a \in \mathbb{R},\ a \cdot 0 = 0.$
(c) $\forall a, b \in \mathbb{R},\ ab = 0 \Leftrightarrow a = 0$ *or* $b = 0.$
(d) $\forall a, b \in \mathbb{R},\ (-a)b = -(ab).$
(e) $\forall a \in \mathbb{R},\ (-1)a = -a.$
(f) $\forall a, b \in \mathbb{R},\ (-a)(-b) = ab.$

Proof. (a) By (M2) and (D),

$$(a + b)c = c(a + b) = ca + cb = ac + bc.$$

(b) By (A3) and (D),

$$a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0,$$

and by (A3), (A2), $a \cdot 0 = 0 + a \cdot 0$. Then Theorem 1.1,(c) gives $a \cdot 0 = 0$.
(c) If $b = 0$, then $ab = a \cdot 0 = 0$ by (b).
If $a = 0$, then $ab = ba = b \cdot 0 = 0$ by (M2) and (b).
Now assume that $ab = 0$. If $b = 0$, the property "$a = 0$ or $b = 0$" follows. So now assume $b \neq 0$. Then

$$a = a \cdot 1 = a(bb^{-1}) = (ab)b^{-1} = 0 \cdot b^{-1} = 0$$

by (M3), (M4), (M1), (M2) and (b).
(d) Using field laws, we get

$$ab + (-a)b = (a + (-a))b = 0 \cdot b = 0,$$

and from Theorem 1.1(b), $(-a)b = -ab$.
(e) is a special case of (d).
(f) From (d) and other laws and rules (which one's?) we find

$$(-a)(-b) = -[a(-b)] = -[(-b)a] = -[-ba] = ba = ab. \qquad \square$$

Theorem 1.3 (Basic field properties: multiplication)**.**
(a) *The number* 1 *is unique.*
(b) *For all $a \in \mathbb{R}$ with $a \neq 0$, the number a^{-1} is unique.*
(c) *For all $a, b \in \mathbb{R}$ with $a \neq 0$, the equation $ax = b$ has a unique solution. This solution is $x = a^{-1}b$.*
(d) $\forall a \in \mathbb{R} \setminus \{0\},\ (a^{-1})^{-1} = a.$
(e) $\forall a, b \in \mathbb{R} \setminus \{0\},\ (ab)^{-1} = a^{-1}b^{-1}.$
(f) $\forall a \in \mathbb{R} \setminus \{0\},\ (-a)^{-1} = -a^{-1}.$
(g) $1^{-1} = 1.$

Proof. See tutorials. The proofs are similar to those of Theorem 1.1. □

Notation. $a^{-1}b$ is also written as $\frac{b}{a}$, In particular, $a^{-1} = \frac{1}{a}$.

Tutorial 1.1.1. 1. Prove Theorem 1.3.

Next we give the axioms for the set of positive real numbers. It is convenient to use the notation $a > 0$ for positive numbers a.

O. The order axioms

(O1) **Trichotomy:** For each $a \in \mathbb{R}$, exactly one of the following statements is true: $a > 0$ or $a = 0$ or $-a > 0$.

(O2) If $a > 0$ and $b > 0$, then $a + b > 0$.

(O3) If $a > 0$ and $b > 0$, then $ab > 0$.

The definition of positivity of real numbers gives rise to an order relation for real numbers:

Definition. Let $a, b \in \mathbb{R}$. Then a is called larger than b, written $a > b$, if $a - b > 0$.

Notes. 1. Since $a - 0 = a$, the notation $a > 0$ is consistent.
2. It is convenient to introduce the following notations:

$$\begin{aligned} &a \geq b \Leftrightarrow a > b \text{ or } a = b,\\ &a < b \Leftrightarrow b > a,\\ &a \leq b \Leftrightarrow a < b \text{ or } a = b. \end{aligned}$$

3. We will define general powers later. Below we use the notation $a^2 = a \cdot a$.

Theorem 1.4 (Basic order properties)**.** *Let $a, b, c, d \in \mathbb{R}$. Then*
(a) $a < 0 \Leftrightarrow -a > 0$.
(b) $a < b$ *and* $b < c \Rightarrow a < c$.
(c) $a < b \Rightarrow a + c < b + c$.
(d) $a < b$ *and* $c < d \Rightarrow a + c < b + d$.
(e) $a < b$ *and* $c > 0 \Rightarrow ca < cb$.
(f) $0 \leq a < b$ *and* $0 \leq c < d \Rightarrow ac < bd$.
(g) $a < b$ *and* $c < 0 \Rightarrow ca > cb$.
(h) $a \neq 0 \Rightarrow a^2 > 0$.
(i) $a > 0 \Rightarrow a^{-1} > 0$ *and* $a < 0 \Rightarrow a^{-1} < 0$.
(j) $0 < a < b \Rightarrow b^{-1} < a^{-1}$.
(k) $1 > 0$.

Proof. (a)

$$\begin{aligned} a < 0 &\Leftrightarrow 0 > a && \text{by definition of } <\\ &\Leftrightarrow 0 - a > 0 && \text{by definition of } 0 > a\\ &\Leftrightarrow -a > 0 && \because 0 - a = -a + 0 = -a \end{aligned}$$

(b)

$$\begin{aligned} a < b \text{ and } b < c &\Rightarrow b - a > 0 \text{ and } c - b > 0 && \text{by definition}\\ &\Rightarrow (b - a) + (c - b) > 0 && \text{by (O2)}\\ &\Rightarrow c - a > 0 && \text{by (A1)-(A4)}\\ &\Rightarrow a < c && \text{by definition} \end{aligned}$$

(h) Since $a \neq 0$, either $a > 0$ or $a < 0$.
If $a > 0$, then $a^2 = aa > 0$ by (O3).
If $a < 0$, then $-a > 0$ by (a) and $a^2 = (-a)(-a)$ by Theorem 1.2, (f).
Hence $a^2 = (-a)(-a) > 0$ by (O3).
(i) Since $a \neq 0$, a^{-1} exists with $aa^{-1} = 1$ by (M4). Then $a^{-1} \neq 0$ by Theorem 1.2, (c). Hence $(a^{-1})^2 > 0$ by (h).
Thus, if $a > 0$,
$a^{-1} = a(a^{-1})^2 > 0$ by (O3).
Similarly, use (g) if $a < 0$.
(j) By (i), $a^{-1} > 0$ and $b^{-1} > 0$. Hence $a^{-1}b^{-1} > 0$ by (O3). Then

$$\begin{aligned} b^{-1} &= a(a^{-1}b^{-1}) < b(a^{-1}b^{-1}) \qquad \text{by (e)} \\ &= (bb^{-1})a^{-1} = a^{-1}. \end{aligned}$$

(k) $1 = 1 \cdot 1 = 1^2 > 0$ by (h). (c)–(g): See tutorials. □

Note. There is still one axiom missing, the axiom of Dedekind completeness. However, we will postpone the formulation of this axiom to the next section since we need some further definitions.

Tutorial 1.1.2. 1. Prove Theorem 1.4, (c)–(g).

2. **The absolute value function.** Define the following function on $\mathbb{R}$:

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

Prove the following statements for $x, y \in \mathbb{R}$:
(a) $|x| \geq 0$,
(b) $|xy| = |x|\,|y|$,
(c) $|y| < x \Leftrightarrow -x < y < x$,
(d) $|x + y| \leq |x| + |y|$.

3. Let $x, y, z \in \mathbb{R}$. Which of the following statements are **true** and which are **false**?
(a) $x \leq y \Rightarrow xz \leq yz$,
(b) $0 < x \leq y \Rightarrow \frac{1}{y} \leq \frac{1}{x}$,
(c) $x < y < 0 \Rightarrow \frac{1}{y} < \frac{1}{x}$,
(d) $x^2 < 1 \Rightarrow x < 1$,
(e) $x^2 < 1 \Rightarrow -1 < x < 1$,
(f) $x^2 > 1 \Rightarrow x > 1$.

4. In each of the following questions fill in the $\square$ with < or >.
(a) $a \geq 3 \Rightarrow \frac{a-2}{7} \square \frac{a}{7}$,
(b) $a \geq 1 \Rightarrow \frac{3}{a+1} \square \frac{3}{a}$,
(c) $a > 1 \Rightarrow \frac{9}{a} \square \frac{10}{a-1}$,
(d) $a > 1 \Rightarrow \frac{1}{a^2} \square \frac{1}{a}$,
(e) $a \geq 2 \Rightarrow \frac{1}{a^2-1} \square \frac{1}{a}$,
(f) $a > 3 \Rightarrow \frac{-3}{a} \square \frac{-2}{a-1}$.

5. Let $x \geq 0$ and $y \geq 0$. Show that $x < y \Leftrightarrow x^2 < y^2$.

1.2 Bounded Sets, Suprema and Infima

Definition 1.1. Let S be a nonempty subset of $\mathbb{R}$. Then
1. If there is a number $A \in \mathbb{R}$ such that $x \leq A$ for all $x \in S$, then A is said to be an **upper bound of** S, and S is said to be **bounded above**.
2. If there is a number $a \in \mathbb{R}$ such that $x \geq a$ for all $x \in S$, then a is said to be a **lower bound of** S, and S is said to be **bounded below**.
3. If S is bounded above and bounded below, then S is called **bounded**.

Note. Recall the definition of intervals. [Note for lecturers: We do not need to write this down. Students have seen this in first year. Just refer students to the notes in case they have forgotten the definitions.] Let a, b. Then

$$(a,b) = \{x \in \mathbb{R} : a < x < b\}, \quad (a,b] = \{x \in \mathbb{R} : a < x \leq b\}, \quad [a,b) = \{x \in \mathbb{R} : a \leq x < b\},$$
$$[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}, \quad (-\infty,b) = \{x \in \mathbb{R} : x < b\}, \quad (-\infty,b] = \{x \in \mathbb{R} : x \leq b\},$$
$$(a,\infty) = \{x \in \mathbb{R} : a < x\}, \quad [a,\infty) = \{x \in \mathbb{R} : a \leq x\}.$$

Example 1.2.1. Let $S = (-\infty, 1)$. Does S have upper and lower bounds?

Solution. Since $x < 1$ for all $x \in S$, 1 is an upper bound of S, and S is therefore bounded above. But also 2 and 7.3 are upper bounds of S, for example.
Assume that S has a lower bound m. Then $m \leq 0$ since $0 \in S$, and $m - 1 < m \leq 0 < 1$ shows that $m - 1 \in S$. So $m \leq m - 1$ since m is a lower bound of S. But this is false, so that m cannot be a lower bound of S. Therefore S has no lower bound, and S is therefore not bounded below.

Definition 1.2. Let S be a nonempty subset of $\mathbb{R}$.
1. If S has an upper bound M which is an element of S, then M is called the **greatest element** or **maximum** of S, and we write $M = \max S$.
2. If S has a lower bound m which is an element of S, then m is called the **least element** or **minimum** of S, and we write $m = \min S$.

From the definition, we immediately obtain

Proposition 1.5. *Let S be a nonempty subset of $\mathbb{R}$. Then*
1. $M = \max S \Leftrightarrow M \in S$ and $x \leq M$ for all $x \in S$,
2. $m = \min S \Leftrightarrow m \in S$ and $x \geq m$ for all $x \in S$.

We have already implicitly used in our notation that maximum and minimum are unique if they exist. The next result states this formally.

Proposition 1.6. *Let S be a nonempty subset of $\mathbb{R}$. If maximum or minimum of S exist, then they are unique.*

Proof. Assume that S has maxima M_1 and M_2. We must show $M_1 = M_2$. Since $M_1 \in S$ and M_2 is an upper bound of S, we have $M_1 \leq M_2$. Since $M_2 \in S$ and M_1 is an upper bound of S, we have $M_2 \leq M_1$. Hence $M_1 = M_2$.
A similar proof holds for the minimum. $\square$

Example 1.2.2. Let $a, b \in \mathbb{R}$, $a < b$, and $S = [a, b)$. Then $\min S = a$, but S has no maximum.

Solution. Clearly, $a \in S$ and $a \leq x$ for all $x \in S$, so that $a = \min S$.
Assume S has a maximum M. Since $M \in S$, $a \leq M < b$. Put $c = \dfrac{M+b}{2}$. Then

$$a \leq M = \frac{2M}{2} = \frac{(1+1)M}{2} = \frac{M+M}{2} < \frac{M+b}{2} < \frac{b+b}{2} = b.$$

Therefore

$$a \leq M < c < b,$$

which shows $c \in S$ and $M < c$, contradicting that $M = \max S$ is an upper bound of S.

Definition 1.3. Let S be a nonempty subset of $\mathbb{R}$.
A real number M is said to be the **supremum** or **least upper bound** of S, if
(a) M is an upper bound of S, and
(b) if L is any upper bound of S, then $M \leq L$.
The supremum of S is denoted by $\sup S$.

Definition 1.4. Let S be a nonempty subset of $\mathbb{R}$.
A real number m is said to be the **infimum** or **greatest lower bound** of S, if
(a) m is a lower bound of S, and
(b) if l is any lower bound of S, then $m \geq l$.
The infimum of S is denoted by $\inf S$.

Note. Let S be a nonempty subset of $\mathbb{R}$.
1. By definition, if S has a supremum, then $\sup S$ is the minimum of the (nonempty) set of the upper bounds of S. Hence $\sup S$ is unique by Proposition 1.6.
2. Similarly, $\inf S$ is unique if it exists.

Proposition 1.7. *Let S be a nonempty subset of* $\mathbb{R}$.
1. If $\max S$ *exists, then* $\sup S$ *exists, and* $\sup S = \max S$.
2. If $\min S$ *exists, then* $\inf S$ *exists, and* $\inf S = \min S$.

Proof. 1. $\max S$ is an upper bound of S by definition. Since $\max S \in S$, $\max S \leq L$ for any upper bound L of S. Hence $\max S = \sup S$.
The proof of 2. is similar. □

Example 1.2.3. Let $a < b$ and put $S = [a, b)$. Find $\inf S$ and $\sup S$ if they exist.

Solution. By Proposition 1.7 and Example 1.2.2, $\inf S = \min S = a$ exists.
From $x < b$ for all $x \in S$ we have that b is an upper bound of S. If there were an an upper bound L of S with $L < b$, it would follow as in the solution to Example 1.2.2 that there is $c \in S$ with $L < c$, which contradicts the fact that L is an upper bound of S.
Hence $\sup S = b$.

Note. It is crucial for Analysis that sufficiently may subsets have a supremum and/or infimum. For example, with the currently formulated axioms, we cannot decide if the sets
$S_1 = \{x \in \mathbb{R} : x > 0 \text{ and } x^2 < 2\}$ or $S_2 = \{x \in \mathbb{R} : x > 0 \text{ and } x^2 \leq 2\}$
have suprema.
Note that the naïve approach putting $S_2 = (0, \sqrt{2}]$ is invalid as we do not know if $\sqrt{2}$, i. e., a real number $x > 0$ with $x^2 = 2$, exists.

Thus we have the last axiom of the real number system.

C. The Dedekind completeness axiom

(C) Every nonempty subset of $\mathbb{R}$ which is bounded above has a supremum.

Additional Note. It can be shown that the real number system is uniquely determined by the axioms (A1)-(A4), (M1)-(M4), (D), (O1)-(O3), (C).

Theorem 1.8 (Positive square root)**.** *Let $a \geq 0$. Then there is a unique $x \geq 0$ such that $x^2 = a$. We write $x = \sqrt{a} = a^{\frac{1}{2}}$.*

Proof. If $a = 0$, we have $0^2 = 0$ and $x^2 > 0$ if $x > 0$, so that $\sqrt{0} = 0$ is the unique number $x \geq 0$ such that $x^2 = 0$.
Now let $a > 0$. Note that $x \geq 0$ and $x^2 = a > 0$ gives $x > 0$. For the uniqueness proof let $x_1 > 0$ and $x_2 > 0$ such that $x_1^2 = a$ and $x_2^2 = a$. Then

$$0 = a - a = x_1^2 - x_2^2 = (x_1 + x_2)(x_1 - x_2).$$

From $x_1 + x_2 > 0$ it follows that $x_1 - x_2 = 0$, i. e., the uniqueness $x_1 = x_2$.

For the existence of the square root define

$$S_a = \{x \in \mathbb{R} : 0 < x,\ x^2 < a\}.$$

First we are going to show that $S_a \neq \emptyset$ and that S_a is bounded.
If $a < 1$, then $a^2 < a \cdot 1 = a$, so that $a \in S_a$.
If $a \geq 1$, then $(\frac{1}{2})^2 < 1^2 = 1 \leq a$, so that $\frac{1}{2} \in S_a$.
For $x > a + 1$ we have

$$x^2 > (a+1)^2 = a^2 + 2a + 1 > 2a > a,$$

so that any $x > a + 1$ does not belong to S_a. Thus $x \leq a + 1$ for all $x \in S_a$. We have shown that S_a is bounded above with upper bound $a + 1$.
By the Dedekind completeness axiom there exists $M_a = \sup S_a$. Note that $M_a > 0$ since S_a has positive elements. To complete the proof we will show that $M_a^2 = a$.
By proof by contradiction, assume that $M_a^2 \neq a$.
Case I: $M_a^2 < a$. Put

$$\varepsilon = \min\left\{\frac{a - M_a^2}{4M_a}, M_a\right\}.$$

Then $\varepsilon > 0$ and

$$\begin{aligned}
(M_a + \varepsilon)^2 - a &= M_a^2 + 2M_a\varepsilon + \varepsilon^2 - a \\
&= M_a^2 - a + (2M_a + \varepsilon)\varepsilon \\
&\leq M_a^2 - a + 3M_a\varepsilon \\
&\leq M_a^2 - a + 3M_a\frac{a - M_a^2}{4M_a} \\
&= \frac{1}{4}(M_a^2 - a) < 0.
\end{aligned}$$

Thus

$$(M_a + \varepsilon)^2 < a,$$

giving $M_a + \varepsilon \in S_a$, contradicting the fact that M_a is an upper bound of S_a.
Case II: $M_a^2 > a$. Put

$$\varepsilon = \frac{M_a^2 - a}{2M_a}.$$

Then $0 < \varepsilon < \frac{1}{2}M_a$ and

$$\begin{aligned}
(M_a - \varepsilon)^2 - a &= M_a^2 - 2M_a\varepsilon + \varepsilon^2 - a \\
&> M_a^2 - a - 2M_a\varepsilon \\
&= M_a^2 - a - 2M_a\frac{M_a^2 - a}{2M_a} \\
&= 0.
\end{aligned}$$

Hence for all $x \geq M_a - \varepsilon > \frac{1}{2}M_a > 0$,

$$x^2 \geq (M_a - \varepsilon)^2 > a$$

so that any $x \geq M_a - \varepsilon$ does not belong to S_a. Hence $M_a - \varepsilon$ is an upper bound of S_a, contradicting the fact that M_a is the least upper bound of S_a.
So $M_a^2 \neq a$ is impossible, and $M_a^2 = a$ follows. □

The completeness axiom can be thought of as ensuring that there are no 'gaps' on the real line.

Theorem 1.9 (Characterizations of the supremum)**.** *Let S be a nonempty subset of $\mathbb{R}$. Let $M \in \mathbb{R}$. The following are equivalent:*
(a) $M = \sup S$*;*
(b) *M is an upper bound of S, and for each $\varepsilon > 0$, there is $s \in S$ such that $M - \varepsilon < s$;*
(c) *M is an upper bound of S, and for each $x < M$, there exists $s \in S$ such that $x < s$.*

Proof. (a) $\Rightarrow$ (b): Since (a) holds, M is the least upper bound and thus an upper bound.
Let $\varepsilon > 0$. Then $M - \varepsilon$ is not an upper bound of S since M is the least upper bound of S. Hence $x \leq M - \varepsilon$ does not hold for all $x \in S$, and therefore there must be $s \in S$ such that $M - \varepsilon < s$.
(b) $\Rightarrow$ (c): Let $x < M$. Then $\varepsilon = M - x > 0$, and by assumption there is $s \in S$ with $M - \varepsilon < s$. Then $x = M - \varepsilon < s$.
(c) $\Rightarrow$ (a): By assumption, M is an upper bound of S. Suppose that M is not the least upper bound. Then there is an upper bound L of S with $L < M$. By assumption (c), there is $s \in S$ such that $L < s$, contradicting that L is an upper bound of S. So M must be indeed the least upper bound of S. □

There is an apparent asymmetry in the Dedekind completion. However, there is a version for infima, which is obtained by reflection. To this end, if S is a (nonempty) subset of $\mathbb{R}$ set

$$-S = \{-x : x \in S\}.$$

Note that because of $-(-x) = x$ this can also can be written as

$$-S = \{x \in \mathbb{R} : -x \in S\}.$$

Since $x \leq y \Leftrightarrow -x \geq -y$ it is easy to see that the following properties hold:

Proposition 1.10. *Let S be be nonempty subset of $\mathbb{R}$. Then*
(a) *$-S$ is bounded below if and only if S is bounded above,* $\inf(-S) = -\sup S$*, and if* $\max S$ *exists, then* $\min(-S)$ *exists and* $\min(-S) = -\max S$.
(b) *$-S$ is bounded above if and only if S is bounded below,* $\sup(-S) = -\inf S$*, and if* $\min S$ *exists, then* $\max(-S)$ *exists and* $\max(-S) = -\min S$.
(c) *$-S$ is bounded if and only if S is bounded.*

Thus the Dedekind completeness axiom immediately gives

Theorem 1.11. *Every nonempty subset of $\mathbb{R}$ which is bounded below has an infimum.*

Theorem 1.12 (Characterizations of the infimum)**.** *Let S be a nonempty subset of $\mathbb{R}$. Let $m \in \mathbb{R}$. The following are equivalent:*
(a) $m = \inf S$*;*
(b) *m is a lower bound of S, and for each $\varepsilon > 0$, there is $s \in S$ such that $s < m + \varepsilon$;*
(c) *m is a lower bound of S, and for each $x > m$, there exists $s \in S$ such that $s < x$.*

Theorem 1.13 (Dedekind cut)**.** *Let A and B be nonempty subsets of $\mathbb{R}$ such that*
(i) $A \cap B = \emptyset$,
(ii) $A \cup B = \mathbb{R}$,
(iii) $\forall\, a \in A\, \forall\, b \in B,\ a \leq b$.
Then there is $c \in \mathbb{R}$ such that $a \leq c \leq b$ for all $a \in A$ and $b \in B$.

Proof. A is nonempty and bounded above (any $b \in B$ is an upper bound of A), so $c = \sup A$ exists by the Dedekind completeness axiom, and $c \geq a$ for all $a \in A$ by definition of upper bound. (iii) says that each $b \in B$ is an upper bound of A, and hence $c \leq b$ since c is the least upper bound of A. □

Note. The above theorem says that Dedekind completeness implies the Dedekind cut property. Conversely, if the ordered field axioms and the Dedekind cut property are satisfied, then the Dedekind completeness axiom holds (see tutorials).

Tutorial 1.2.1. 1. In each of the following cases, state if the given set is bounded above or not. If a set is bounded above, give two different upper bounds for the set, give the supremum of the set and state if the set has a maximum or not.

(a) $(-3, 2)$	(b) $(1, \infty)$	(c) $[10, 11]$
(d) $\{5, 4\}$	(e) $\{10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -5\}$	
(f) $(-\infty, 2]$	(g) $\{x \in \mathbb{R} : x^2 < 3\}$	(h) $\{x \in \mathbb{R} : x^2 \leq 3\}$

2. For each of the sets in Q. 1, state if the given set is bounded below or not. If a set is bounded below, give two different lower bounds for the set, give the infimum of the set and state if the set has a minumum or not.

3. Let S be a nonempty subset of $\mathbb{R}$.
(a) If a is the greatest element of S, then what is $\sup S$?
(b) If $\sup S = a$, then what are the upper bounds of S? .
(c) If $\sup S = a$, does S have a maximum?
(d) If $\sup S = a$ and $a \in S$, does S have a maximum?

4. Prove Proposition 1.10.

5. Prove Theorem 1.11.

6. Prove Theorem 1.12.

7. Let S and T be non-empty subsets of $\mathbb{R}$ which are bounded above. Use Theorem 1.9 to prove that $\sup(S+T) = \sup S + \sup T$.
Step 1: Let $K = \sup S$ and $M = \sup T$. Show that $K + M$ is an upper bound of $S + T$.
Step 2: Let $\varepsilon > 0$. Then $\frac{\varepsilon}{2} > 0$ as well. Since $\sup S = K$, by Theorem 1.9 there is $s \in S$ such that $K - \frac{\varepsilon}{2} < s$. Also, since $\sup T = M$, there is $t \in T$ such that $M - \frac{\varepsilon}{2} < t$. Show that there exists $u \in S + T$ such that $K + M - \varepsilon < u$.
Step 3: Conclude, using Theorem 1.9, that $\sup(S + T) = K + M = \sup S + \sup T$.

8. **Some simple questions.** Decide which of the following statements are **true** and which are **false**.

(a) $\frac{1}{2} \in \{0, 1\}$	(b) $3 \in (0, 3)$	(c) $17 \in [0, 17]$
(d) $17 \in (-3, 18)$	(e) $17 \in [16, 18]$	(f) $2 \in \{1, 3, 5, 7\}$
(g) $2.5 \in \{x \in \mathbb{R} : x^2 \geq 4\}$	(h) $-1 \in \{x \in \mathbb{R} : 2x + 7 < 5\}$	

9*. Assume that the Dedekind cut property, Theorem 1.13 as well as the ordered field axioms are satisfied. Show that the Dedekind completeness holds.

1.3 Natural Numbers

Naïvely, the natural numbers are $0, 1, 2, 3, \ldots$, where the dots stand for the numbers which are obtained by adding one to the previous number. However, our axioms of the real number system do not tell us that this is possible. Hence we need another axiom to characterize the natural numbers.

P. Peano's axioms of the natural numbers
There is a unique subset $\mathbb{N}$ of $\mathbb{R}$, called the set of natural numbers, satisfying the following properties:

(P1) $0 \in \mathbb{N}$,

(P2) $a \in \mathbb{N} \Rightarrow a + 1 \in \mathbb{N}$,

(P3) $\forall\, a \in \mathbb{N}, a + 1 \neq 0$,

(P4) If $S \subset \mathbb{N}$ such that $0 \in S$ and $a + 1 \in S$ for all $a \in S$, then $S = \mathbb{N}$.

Additional Notes. 1. Here we define natural numbers to start at 0. This is common if one uses the axiomatic definition of the natural numbers. You may start the natural numbers with 1, if you are consistent with this.
2. If one defines the natural numbers without reference to real number, there is no arithmetic on $\mathbb{N}$ needed. Then the number $a + 1$ is called the successor of a, and axioms (P2) and (P3) can be phrased as follows: Every natural number has a successor, and 0 is not the successor of a natural number.
3. The natural numbers are used to "count". In everyday life, counting refers to concrete objects ('There are 220 students in the class', 'I have got no pen with me', 'I ate three rolls this morning'). However, they can also be defined as cardinalities of abstract (finite) sets. Here the cardinality $\#S$ of a (finite) set S is the number of elements in the set. The empty set $\emptyset$ has no elements, so $\#\emptyset = 0$ serves as the definition of the number 0. To get the next

number, consider the set which consists of the empty set. It has exactly one element, namely the empty set. So $\#\{\emptyset\} = 1$. Next $\#\{\emptyset, \{\emptyset\}\} = 2$, and so on.

Recall that the integers $\mathbb{Z}$ and the rational numbers $\mathbb{Q}$ are defined by

$$\mathbb{Z} = \{x \in \mathbb{R} : x \in \mathbb{N} \text{ or } -x \in \mathbb{N}\},$$
$$\mathbb{Q} = \left\{ x \in \mathbb{R} : x = \frac{p}{q},\ p, q \in \mathbb{Z},\ q \neq 0 \right\}.$$

As a consequence of (P4) one has

The Principle of Mathematical Induction
Suppose we have a statement $A(n)$ associated with each $n \in \mathbb{N}$. If
(i) the statement $A(0)$ is true, and
(ii) whenever the statement $A(n)$ is true, also the statement $A(n+1)$ is true,
then the statement $A(n)$ is true for all $n \in \mathbb{N}$.

Note. The statement $A(0)$ is called the induction base. The induction base does not need to be taken at 0, one can take any integer which is suitable for the statements to be proved.

Theorem 1.14. 1. *All nonzero natural numbers n satisfy $n \geq 1$.*
2. $\forall\, n \in \mathbb{N} \setminus \{0\}\ \exists\, m \in \mathbb{N},\ n = m + 1.$
3. $\forall\, m, n \in \mathbb{N},\ m + n \in \mathbb{N}.$
4. *for all $m, n \in \mathbb{N}$ with $m \geq n$ there is $k \in \mathbb{N}$ such that $m = n + k$.*

Proof. A selection of these proofs may be presented in class.
1. For each $n \in \mathbb{N}$ our statement $A(n)$ is: if $n \neq 0$, then $n \geq 1$.
We have to show that this statement is true for all $n \in \mathbb{N}$.
Induction base: $A(0)$ is true since the assumption $0 \neq 0$ does not hold.
Induction step: Now let $n \in \mathbb{N}$ and assume that $A(n)$ holds. We must show that $A(n+1)$ holds.
If $n = 0$ we prove $A(1)$ directly: $0 + 1 = 1 \geq 1$.
Now let $n \neq 0$. Then $n \geq 1$ since $A(n)$ is true. Since $1 > 0$ by Theorem 1.4, (k), we therefore have $n+1 \geq 1+1 \geq 1$.
Hence, by the principle of mathematical induction, $A(n)$ is true for all $n \in \mathbb{N}$.
2. Let $S = \{0\} \cup \{n \in \mathbb{N} : \exists\, m \in \mathbb{N},\ n = m + 1\}$.
Then $S \subset \mathbb{N}$, $0 \in S$, and if $n \in S$ then $n + 1 \in S$ since $n \in \mathbb{N}$.
Hence $S = \mathbb{N}$ by (P4).
3. For $n \in \mathbb{N}$ let $A(n)$ be the statement:
For all $m \in \mathbb{N}$, $m + n \in \mathbb{N}$.
$A(0)$ is trivially true.
Now assume that $A(n)$ is true. Then, for $m \in \mathbb{N}$,

$$m + (n + 1) = (m + 1) + n \in \mathbb{N}$$

by induction hypothesis. Hence $A(n + 1)$ holds.
By the principle of mathematical induction, $A(n)$ holds for all $n \in \mathbb{N}$. The proof is complete.
4. For $n \in \mathbb{N}$ let $A(n)$ be the statement:
For all $m \in \mathbb{N}$ with $m \geq n$ there exists $k \in \mathbb{N}$ such that $m = n + k$.
$A(0)$ is true with $k = m$ for all m.
Now assume $A(n)$ is true and let $m \in \mathbb{N}$ with $m \geq n + 1$. By induction hypothesis, since $m \geq n + 1 \geq n$ there is $k \in \mathbb{N}$ such that $m = n + k$.
If $k = 0$, then $m = n < n + 1$, which is impossible.
Hence $k \neq 0$, and so $k = l + 1$ for some $l \in \mathbb{N}$ by part 2. It follows that

$$m = n + l + 1 = (n + 1) + l,$$

and therefore $A(n + 1)$ is true.
By the principle of mathematical induction, $A(n)$ hold for all $n \in \mathbb{N}$. The proof is complete. □

Finally, we state and prove two important principles which are consequences of the Dedekind completeness and Peano's axioms.

Theorem 1.15 (Well-ordering principle)**.** *Every nonempty subset of* $\mathbb{N}$ *has a smallest element.*

Proof. Let $S \subset \mathbb{N}$, $S \neq \emptyset$. Assume S has no minimum. Let

$$T = \{n \in \mathbb{N} : \forall\, m \in S,\ n < m\}.$$

Note that $n \in T$ and $m \in S$ gives $n < m$, so that $n \neq m$. Hence $T \cap S = \emptyset$.
Assume that $0 \in S$. Since $0 \le n$ for all $n \in \mathbb{N}$ by Theorem 1.14, 1, and since $S \subset \mathbb{N}$, we have $0 \le m$ for all $m \in S$. But S has no minimum, and therefore $0 \notin S$. Thus $0 < 1 \le m$ for all $m \in S$.
Hence $0 \in T$.
Now let $n \in T$. We want to show that $n + 1 \in T$.
By definition of T, we have $n < m$ for all $m \in S$. By Theorem 1.14, 4, For each such m there is $k \in \mathbb{N}$ such that $m = n + k$. Since $n < m$, it follows that $k \neq 0$, and hence $k \ge 1$ by Theorem 1.14, 1. Thus

$$m = n + k \ge n + 1$$

for all $m \in S$. Since we assume that S has no minimum, it follows that $n + 1 \notin S$ and hence $n + 1 \neq m$ for all $m \in S$. Altogether $n + 1 < m$ for all $m \in S$. This shows that $n + 1 \in T$.
By (P4), $T = \mathbb{N}$. Hence

$$S = \mathbb{N} \cap S = T \cap S = \emptyset,$$

a contradiction. □

Note. One can show that if $S \subset \mathbb{Z}$, $S \neq \emptyset$, and S is bounded below, then S has a minimum.

Theorem 1.16 (The Archimedean principle)**.** *For each* $x \in \mathbb{R}$ *there is* $n \in \mathbb{N}$ *such that* $n > x$.

Proof. Assume the Archimedian principle is false. Then there is $x \in \mathbb{R}$ such that $n \le x$ for all $n \in \mathbb{N}$. That means, $\mathbb{N}$ is bounded above and therefore has a supremum M.
By Theorem 1.9 there is $n \in \mathbb{N}$ such that $M - 1 < n$. Then

$$n + 1 > (M - 1) + 1 = M$$

and $n + 1 \in \mathbb{N}$ contradict the fact that M is an upper bound of $\mathbb{N}$. Hence the Archimedian principle must be true. □

Definition 1.5. A subset S of $\mathbb{R}$ is said to be **dense** in $\mathbb{R}$ if for all $x, y \in \mathbb{R}$ with $x < y$ there is $s \in S$ such that $x < s < y$.

Real numbers which are not rational numbers are called **irrational** numbers.

Note. 1. $\sqrt{2}$ is irrational (see tutorials).
2. $\mathbb{Q}$ is an ordered field, i. e., $\mathbb{Q}$ satisfies all axioms of $\mathbb{R}$ except the Dedekind completeness axiom (see tutorials).
3. If a is rational and b is irrational, then $a + b$ is irrational (see tutorials).

Theorem 1.17. *The set of rational numbers as well as the set of irrational numbers are dense in* $\mathbb{R}$.

Proof. Let $x, y \in \mathbb{R}$, $x < y$. By the Archimedian principle, there is a natural number n_0 such that

$$n_0 > \frac{2}{y - x} > 0.$$

Let

$$S = \{n \in \mathbb{Z} : n > n_0 x\}.$$

By the Archimedean principle, $S \neq \emptyset$. Also, S is bounded below. So S has a minimum m. Then $m - 1 \notin S$, so that $m - 1 \le n_0 x$.
From $1 < 2 < 4$ it follows that $1 < \sqrt{2} < 2$. Then

$$x = \frac{n_0 x}{n_0} < \frac{m}{n_0} < \frac{m - 1 + \sqrt{2}}{n_0} < \frac{m + 1}{n_0} \le \frac{n_0 x + 2}{n_0} = x + \frac{2}{n_0} < x + 2\frac{y - x}{2} = y.$$

Then

$$u = \frac{m}{n_0} \in \mathbb{Q}, \quad v = \frac{m - 1 + \sqrt{2}}{n_0} \in \mathbb{R} \setminus \mathbb{Q}$$

and, by above,

$$x < u < v < y,$$

so that we have have a rational and an irrational number between each pair of real numbers x, y with $x < y$. Hence both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$. $\square$

Example 1.3.1. Let $S = \{x \in \mathbb{Q} : x < \sqrt{2}\}$ and $T = \{x \in \mathbb{Q} : x \le \sqrt{2}\}$. Then $S = T$, and S and T have no supremum in $\mathbb{Q}$.

Example 1.3.2 (Bernoulli's inequality)**.** For all $x \in \mathbb{R}$ with $x \ge -1$ and all $n \in \mathbb{N}$, $(1 + x)^n \ge 1 + nx$.

Solution. For $n = 0$, $(1 + x)^n = (1 + x)^0 = 1 = 1 + 0 \cdot x = 1 + nx$, so that the inequality holds for $n = 0$.
Let $n \in \mathbb{N}$ and assume the Bernoulli inequality holds for this n. Then

$$(1 + x)^{n+1} = (1 + x)(1 + x)^n \ge (1 + x)(1 + nx) = 1 + x + nx + nx^2 = 1 + (n + 1)x + nx^2 \ge 1 + (n + 1)x.$$

The first inequality follows from $1 + x \ge 0$ and the induction hypothesis, the last from $nx^2 \ge 0$. By the principle of induction, the Bernoulli inequality holds for all $n \in \mathbb{N}$.

Note. We will make use of more rules of real numbers without proof. See the tutorials for the power rules, for example.

Tutorial 1.3.1. 1. Show that if $S \subset \mathbb{Z}$, $S \ne \emptyset$, and S is bounded below, then S has a minimum.

2. Show that $\sqrt{2}$ is irrational. **Hint.** Assume that $\sqrt{2} = \frac{p}{q}$ with positive integers p and q. You may assume that p and q do not have common factors, i. e., there are no positive integers k, p_1, q_1 with $k \ne 1$ such that $p = p_1 k$ and $q = q_1 k$.

3. Show that the rational numbers satisfy the axioms (A1)–(A4), (M1)–(M4), (D), and (O1)–(O3).

4. Let $a, b \in \mathbb{Q}$ with $b \ne 0$ and $r \in \mathbb{R} \setminus \mathbb{Q}$. Show that $a + br \in \mathbb{R} \setminus \mathbb{Q}$.

5. For $x \in \mathbb{R}$ and $n \in \mathbb{N}$, x^n is defined inductively by
(i) $x^0 = 1$,
(ii) $x^{n+1} = xx^n$ for $n \in \mathbb{N}$.
Show that (a) $x^n x^m = x^{n+m}$, (b) $(x^n)^m = x^{nm}$, (c) $x^n y^n = (xy)^n$.

Chapter 2

Sequences

Students have seen sequences and learnt rules for limits and how to apply these rules to find limits. In this chapter, students will learn the proper definitions of limits and the proofs of the rules for limits.

2.1 Sequences

Definition 2.1. A **(real) sequence** is an ordered list of infinitely many real numbers

$$a_1, a_2, a_3, a_4, \ldots .$$

We usually denote a sequence by $(a_n) = a_1, a_2, a_3, \ldots$.

The number a_n is called the n-th term of the sequence. The subscript n is called the index.

Note. The index of the sequence does not have to start at 1, and therefore one also writes $(a_n)_{n=n_0}^{\infty}$, where n_0 can be any integer, e. g., $(2n-3)_{n=1}^{\infty}$, $(2n-3)_{n=0}^{\infty}$, $(2n-3)_{n=-5}^{\infty}$.

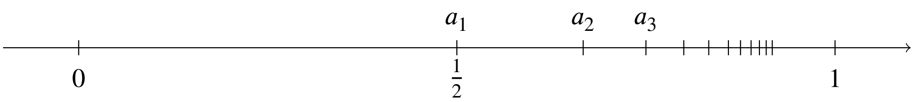
Recall the following note and sketch from First Year Calculus. [Lecturers: do not write this down in class.]

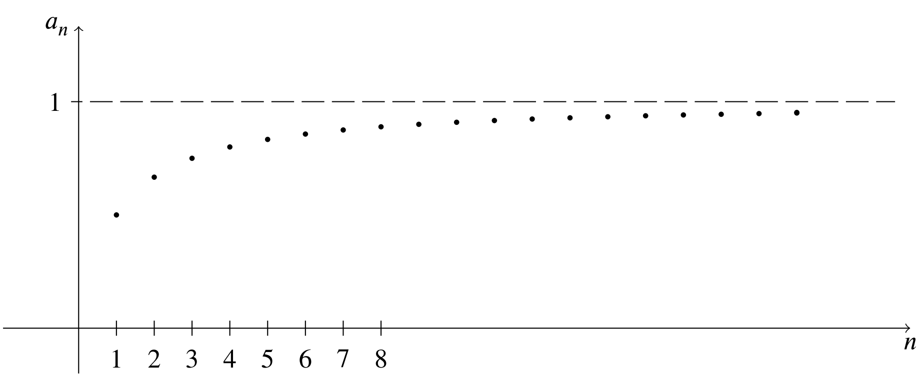
Note. (i) A sequence can be identified with a function on the (positive) integers:

$$a_n = f(n).$$

(ii) A sequence can be plotted as points of the real axis or as the graph of the function, see (i).

Below is a plot for the sequence $\left(\frac{n}{n+1}\right)_{n=1}^{\infty}$.





In first year calculus students have seen the following informal definition. [Lecturers: You do not need to write this down. In any case, emphasize that this notion has to be formalized.]

Informal Definition. A sequence $\{a_n\}$ is said to **converge** if there is a number L such that a_n is as close to L as we like for all sufficiently large n.

We have to formalize "as close to" and "for all sufficiently large": "as close to" means that the distance does not exceed any given (small) positive number. The distance between two numbers x and y is $y - x$ if $x < y$ and $x - y$ if $x > y$. Hence the distance between x and y is $|y - x|$.

Therefore we have the following formal definition:

Definition 2.2. Let (a_n) be a sequence and $L \in \mathbb{R}$.
1. The statement 'a_n tends to L as n tends to infinity', which we write as '$a_n \to L$ as $n \to \infty$', is defined by: $\forall \varepsilon > 0 \ \exists K \in \mathbb{R} \ \forall n \in \mathbb{N}, n \geq K, |a_n - L| < \varepsilon$.
2. If $a_n \to L$ as $n \to \infty$, we say that (a_n) converges to L, and we also write $\lim\limits_{n\to\infty} a_n = L$.
3. The sequence (a_n) is said to be **convergent** if it converges to some real number. Otherwise, the sequence (a_n) is said to be **divergent**.

Note. 1. It is useful to note that $|a_n - L| < \varepsilon$ is equivalent to $L - \varepsilon < a_n < L + \varepsilon$.
2. The number K depends on ε. We may write K_ε to emphasize the dependence on ε.
3. If one wants to prove convergence of a sequence from first principles, then one has to 'guess' a limit L and then prove that it is indeed the limit.
4. The 'first few terms' do not matter for convergence and the limit. That is, the sequence $(a_n)_{n=n_0}^{\infty}$ converges if and only if the sequence $(a_n)_{n=m_0}^{\infty}$ converges, and their limits coincide.
5. By the Archimedean principle, there are only finitely many natural numbers such that $n < K$, and one may also assume $K \in \mathbb{N}$, without loss of generality.

Example 2.1.1. Prove that the sequence $(a_n) = \left(\frac{n}{n+1}\right)$ converges and find its limit.

Solution. We first have to guess the limit. From the rules used in first year, we guess that the limit should be $L = 1$. Now let $\varepsilon > 0$. We must find K_ε such that

$$n \geq K_\varepsilon \Rightarrow |a_n - 1| < \varepsilon.$$

For this we first simplify:

$$|a_n - 1| = \left| \frac{n}{n+1} - 1 \right| = \frac{1}{n+1}.$$

Hence $|a_n - 1| < \varepsilon$ provided that $\dfrac{1}{n+1} < \varepsilon$, which can be written as $n + 1 > \frac{1}{\varepsilon}$ or $n > \frac{1}{\varepsilon} - 1$. Hence if we take $K_\varepsilon = \frac{1}{\varepsilon} - 1$, then $|a_n - 1| < \varepsilon$ for all $n > K_\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have shown that $\frac{n}{n+1} \to 1$ as $n \to \infty$.

Lemma 2.1. *If $L, M \in \mathbb{R}$ such that $|L - M| < \varepsilon$ for all $\varepsilon > 0$, then $L = M$.*

Proof. Assume $L \neq M$. Then either $L < M$ or $L > M$. But

$$L < M \Rightarrow |L - M| = M - L > 0,$$
$$M < L \Rightarrow |L - M| = L - M > 0,$$

so that $|L - M| > 0$. Then

$$0 < \frac{|L - M|}{2} < |L - M|,$$

which contradicts $|L - M| < \varepsilon$ for $\varepsilon = \frac{|L-M|}{2}$. Hence the assumption $L \neq M$ must be false, and $L = M$ follows. ☐

Theorem 2.2. *If the sequence (a_n) converges, then its limit is unique.*

Proof. Assume that (a_n) converges to L and M. Let $\varepsilon > 0$. Then there are numbers k_L and k_M such that
(i) $|a_n - L| < \frac{\varepsilon}{2}$ if $n \geq k_L$,
(ii) $|a_n - M| < \frac{\varepsilon}{2}$ if $n \geq k_M$.
Put $K = \max\{k_L, k_M\}$. For positive integers $n \geq K$ we have $n \geq k_L$ and $n \geq k_M$ and therefore

$$|L - M| = |(L - a_n) + (a_n - M)| \leq |a_n - L| + |a_n - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

By Lemma 2.1, $L = M$. $\square$

We are now going to prove the rules you have learnt in first year.

Theorem 2.3 (Limit Laws)**.** *Let $c \in \mathbb{R}$ and suppose that $\lim\limits_{n\to\infty} a_n = L$ and $\lim\limits_{n\to\infty} b_n = M$ both exist. Then*

(a) $\lim\limits_{n\to\infty} c = c$.
(b) $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n = L + M$.
(c) $\lim\limits_{n\to\infty} (ca_n) = c \lim\limits_{n\to\infty} a_n = cL$.
(d) $\lim\limits_{n\to\infty} (a_n b_n) = \left(\lim\limits_{n\to\infty} a_n\right)\left(\lim\limits_{n\to\infty} b_n\right) = LM$.
(e) *If* $M \neq 0$, $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n} = \dfrac{L}{M}$.
(f) *If* $L \neq 0$ *and* $M = 0$, *then* $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n}$ *does not exist.*
(g) *If* $k \in \mathbb{Z}^+$, $\lim\limits_{n\to\infty} a_n^k = \left(\lim\limits_{n\to\infty} a_n\right)^k = L^k$.
(h) *If* $k \in \mathbb{Z}^+$, $\lim\limits_{n\to\infty} \sqrt[k]{a_n} = \sqrt[k]{\lim\limits_{n\to\infty} a_n} = \sqrt[k]{L}$. *If k is even, we assume* $a_n \geq 0$ *and* $L \geq 0$.
(i) *If* $\lim\limits_{n\to\infty} |a_n| = 0$, *then* $\lim\limits_{n\to\infty} a_n = 0$.

Proof. (a) For all $\varepsilon > 0$ and all indices n, with $a_n = c$, $|a_n - c| = |c - c| = 0 < \varepsilon$.
(b) Let $\varepsilon > 0$. Then there are numbers k_L and k_M such that
(i) $|a_n - L| < \frac{\varepsilon}{2}$ if $n \geq k_L$,
(ii) $|b_n - M| < \frac{\varepsilon}{2}$ if $n \geq k_M$.
Put $K = \max\{k_L, k_M\}$. For positive integers $n \geq K$ we have $n \geq k_L$ and $n \geq k_M$ and therefore

$$|(a_n + b_n) - (L + M)| = |(a_n - L) + (b_n - M)| \leq |a_n - L| + |b_n - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $(a_n + b_n)$ converges to $L + M$.
(c) Let $\varepsilon > 0$. Then there is a number K such that
$|a_n - L| < \frac{\varepsilon}{1+|c|}$ if $n \geq K$.
For positive integers $n \geq K$ we have

$$|ca_n - cL| = |c(a_n - L)| = |c|\,|a_n - L| < |c|\frac{\varepsilon}{1 + |c|} < \varepsilon.$$

(d) We consider 2 cases: one special case, to which then the general case is reduced.
Case I: $L = M = 0$. Let $\varepsilon > 0$ and put $\varepsilon' = \min\{1, \varepsilon\}$. Then $0 < \varepsilon' \leq \varepsilon$, and there are numbers k_L and k_M such that
(i) $|a_n - 0| < \varepsilon'$ if $n \geq k_L$,
(ii) $|b_n - 0| < \varepsilon'$ if $n \geq k_M$.
Put $K = \max\{k_L, k_M\}$. For positive integers $n \geq K$ we have $n \geq k_L$ and $n \geq k_M$ and therefore

$$|a_n b_n| = |a_n|\,|b_n| < \varepsilon'^2 \leq \varepsilon' \leq \varepsilon.$$

Case II: L and M are arbitrary. Then

$$a_n b_n = (a_n - L)(b_n - M) + L(b_n - M) + a_n M.$$

By (a), (b), (c), $(a_n - L) \to 0$ and $(b_n - M) \to 0$ as $n \to \infty$, and by (b), (c), and Case I, it follows that $(a_n b_n)$ converges with

$$\lim_{n\to\infty} a_n b_n = 0 + L \cdot 0 + LM = LM.$$

(e) First consider $a_n = 1$. Then

$$\left|\frac{1}{b_n} - \frac{1}{M}\right| = \frac{|M - b_n|}{|b_n M|}.$$

We must assure that b_n stays away from 0, and hence we estimate

$$|M| = |b_n + M - b_n| \le |b_n| + |b_n - M|,$$

which gives

$$|b_n| \ge |M| - |b_n - M|.$$

Since $b_n \to M \neq 0$, there is k_1 such that $|b_n - M| < \frac{|M|}{2}$ for $n \ge k_1$, and hence

$$|b_n| \ge |M| - \frac{|M|}{2} = \frac{|M|}{2}.$$

If we therefore let $\varepsilon > 0$ and choose K such that

$$|b_n - M| < \min\left\{\frac{|M|}{2}, \frac{M^2}{2}\varepsilon\right\}$$

for $n \ge K$, it follows that

$$\left|\frac{1}{b_n} - \frac{1}{M}\right| = \frac{|M - b_n|}{|b_n|\,|M|} \le |M - b_n|\frac{2}{|M|^2} < \frac{|M|^2}{2}\varepsilon\,\frac{2}{|M|^2} = \varepsilon.$$

The general case now follows with (d):

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} \frac{1}{b_n} = L \cdot \frac{1}{M}.$$

(f) Assume that $P = \lim_{n\to\infty} \frac{a_n}{b_n}$ exists. Then, by (d),

$$L = \lim_{n\to\infty} a_n = \lim_{n\to\infty}\left(\frac{a_n}{b_n}\, b_n\right) = \lim_{n\to\infty}\frac{a_n}{b_n} \cdot \lim_{n\to\infty} b_n = P \cdot M = P \cdot 0 = 0,$$

which contradicts $L \neq 0$.
(g) Follows from (d) by induction.
(h) We will only prove the case $k = 2$ as we have not yet established the existence of the k-th root.
If $L = 0$, let $\varepsilon > 0$ and choose K such that $a_n < \varepsilon^2$ for $n \ge K$. Then $\sqrt{a_n} < \varepsilon$ for these n, and $\lim_{n\to\infty} \sqrt{a_n} = 0 = \sqrt{L}$.
If $L > 0$, then

$$|\sqrt{a_n} - \sqrt{L}| = \left|\frac{a_n - L}{\sqrt{a_n} + \sqrt{L}}\right| = \frac{|a_n - L|}{\sqrt{a_n} + \sqrt{L}} \le \frac{1}{\sqrt{L}}|a_n - L|,$$

and choosing K such that $|a_n - L| < \sqrt{L}\varepsilon$ for $n \ge K$, it follows that $|\sqrt{a_n} - \sqrt{L}| < \varepsilon$ for $n \ge K$.
(i) is trivial. □

Theorem 2.4 (**Sandwich Theorem**). *If $a_n \le b_n \le c_n$ and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$, then $\lim_{n\to\infty} b_n = L$.*

Proof. Let $\varepsilon > 0$ and choose k_1 and k_2 such that $|a_n - L| < \varepsilon$ if $n \ge k_1$ and $|c_n - L| < \varepsilon$ if $n \ge k_2$. In particular, for $n \ge K = \max\{k_1, k_2\}$,

$$L - \varepsilon < a_n < L + \varepsilon, \quad L - \varepsilon < c_n < L + \varepsilon$$

gives

$$L - \varepsilon < a_n \le b_n \le c_n < L + \varepsilon.$$

Hence $|b_n - L| < \varepsilon$ if $n \ge K$. □

Theorem 2.5. *Let $x \in \mathbb{R}$. Then x^n converges if and only if $-1 < x \leq 1$.*

Proof. For $x = 1$, $1^n = 1$, so that $\lim_{n\to\infty} 1^n = 1$ by Theorem 2.3 (a).
Let $0 < x < 1$ and let $\varepsilon > 0$. Then $0 < 1 < \frac{1}{x}$ and therefore

$$y := \frac{1}{x} - 1 > 0.$$

Then

$$\frac{1}{x^n} = \left(\frac{1}{x}\right)^n = (1+y)^n \geq 1 + ny$$

by Bernoulli's inequality. Put $K = \dfrac{1-\varepsilon}{y\varepsilon}$. Then, for $n > K$,

$$0 < x^n \leq \frac{1}{1+ny} < \frac{1}{1+\frac{1-\varepsilon}{\varepsilon}} = \varepsilon.$$

Hence $x^n \to 0$ as $n \to \infty$. Now let $x > 1$. If x^n would converge to some L as $n \to \infty$, then

$$1 = \lim_{n\to\infty} 1^n = \lim \left(\frac{1}{x}\right)^n \cdot \lim_{n\to\infty} x^n = 0 \cdot L = 0,$$

which is impossible.
The remaining cases are left as an exercise. □

Theorem 2.6. *If $r > 0$, then*

$$\lim_{n\to\infty} \frac{1}{n^r} = 0.$$

Proof. We are only going to prove the case that r is a positive integer.
Let $r = 1$ and choose $\varepsilon > 0$. Let $K = \frac{2}{\varepsilon}$. Then, for $n \geq K$,

$$0 < \frac{1}{n} \leq \frac{1}{K} = \frac{\varepsilon}{2} < \varepsilon,$$

whence $\left|\frac{1}{n}\right| < \varepsilon$.
Now assume the statement holds for an integer $r > 0$. Then

$$\lim_{n\to\infty} \frac{1}{n^{r+1}} = \lim_{n\to\infty} \frac{1}{n} \cdot \lim_{n\to\infty} \frac{1}{n^r} = 0 \cdot 0 = 0.$$

So the result follows for all positive integers by induction. □

Tutorial 2.1.1. 1. (a) Prove, using the definition of convergence, that the sequence $\left(\frac{n}{n+1}\right)$ does not converge to 2.
(b) Prove, using the definition of convergence, that the sequence $((-1)^n)$ does not converge to any L.

2. (a) Prove that if $\lim_{n\to\infty} a_n = L$, then $\lim_{n\to\infty} |a_n| = |L|$. (Hint: use (and prove) the inequality $||x| - |y|| \leq |x - y|$.)
(b) Give an example to show that the converse to part (a) is not true.

3. **Contractive maps.** Suppose that for some $c \in \mathbb{R}$ with $0 < c < 1$, we have $|a_{n+1} - L| \leq c|a_n - L|$ for all $n \in \mathbb{N}$.
(a) Use induction to prove that $|a_n - L| \leq c^n|a_0 - L|$ for all $n \in \mathbb{N}$.
(b) Use the Sandwich Theorem and the fact that $\lim_{n\to\infty} c^n = 0$ to prove that $\lim_{n\to\infty} a_n = L$.

4. **Recursive algorithm for finding $\sqrt{a}$.** Let $a > 1$ and define

$$a_0 = a \quad \text{and} \quad a_n = \frac{1}{2}\left(a_{n-1} + \frac{a}{a_{n-1}}\right) \text{ for } n \geq 1.$$

(a) Prove that $0 < a_n - \sqrt{a} = \frac{1}{2a_{n-1}}(a_{n-1} - \sqrt{a})^2$ for $n \geq 1$.
(b) Use (a) to prove that $0 \leq a_n - \sqrt{a} \leq \frac{1}{2}(a_{n-1} - \sqrt{a})$ for $n \geq 1$.
(c) Deduce that $\lim_{n\to\infty} a_n = \sqrt{a}$.
(d) Apply four steps of the recursive algorithm with $a = 3$ to approximate $\sqrt{3}$.

2.2 Bounded and Monotonic Sequences

Definition 2.3. 1. A sequence (a_n) is said to be **bounded above** if there is a number $M \in \mathbb{R}$ such that $a_n \leq M$ for all indices n.
2. A sequence (a_n) is said to be **bounded below** if there is a number $m \in \mathbb{R}$ such that $a_n \geq m$ for all indices n.
3. A sequence (a_n) is said to be **bounded** if it is a bounded above and bounded below.

Note. A sequence $(a_n)_{n=n_0}^{\infty}$ is bounded (above, below) if and only if the set $\{a_n : n \in \mathbb{Z}, n \geq n_0\}$ is bounded (above, below).

Definition 2.4. A sequence $\{a_n\}_{n=1}^{\infty}$ is called

increasing if $a_n \leq a_{n+1}$ for all indices n,
strictly increasing if $a_n < a_{n+1}$ for all indices n,
decreasing if $a_n \geq a_{n+1}$ for all indices n,
strictly decreasing if $a_n > a_{n+1}$ for all indices n,
monotonic if it is either increasing or decreasing,
strictly monotonic if it is either strictly increasing or strictly decreasing.

Theorem 2.7. *Every convergent sequence is bounded.*

Proof. Here we use the fact, which is easy to prove (by induction), that a set of the form $\{a_n : n \in \mathbb{Z}, n_0 \leq n \leq n_1\}$ is bounded (and indeed has minimum and maximum). We also use that the union of two bounded sets is bounded. Since (a_n) converges, there are L and k such that $|a_n - L| < 1$ for all $n \geq k$. Hence $\{a_n : n < k\}$ and $\{a_n : n \geq k\} \subset (L-1, L+1)$ are bounded. So the sequence is bounded by the above note. ☐

Although each unbounded sequence diverges, there are some divergent sequences which have a limiting behaviour which we want to exploit. Recall that we use the notation '$n \to \infty$' in the definition of the limit. This extends to '$a_n \to \infty$', and we have therefore the

Definition 2.5. 1. We say that a_n tends to infinity as n tends to infinity and write $a_n \to \infty$ as $n \to \infty$ if for every $A \in \mathbb{R}$ there is $K \in \mathbb{R}$ such that $a_n > A$ for all $n \geq K$.
2. We say that a_n tends to minus infinity as n tends to infinity and write $a_n \to -\infty$ as $n \to \infty$ if for every $A \in \mathbb{R}$ there is $K \in \mathbb{R}$ such that $a_n < A$ for all $n \geq K$.

Notes. 1. For the definition of $a_n \to \infty(-\infty)$ as $n \to \infty$, we may restrict A to $A > A_0$ $(A < A_0)$ for suitable numbers A_0.
2. For $a_n \to \infty$ as $n \to \infty$ we also write $\lim\limits_{n\to\infty} a_n = \infty$ and say that (a_n) diverges to ∞.
3. For $a_n \to -\infty$ as $n \to \infty$ we also write $\lim\limits_{n\to\infty} a_n = -\infty$ and say that (a_n) diverges to $-\infty$.

Example 2.2.1. 1. Let $a_n = (-1)^n$. Then (a_n) is bounded but not convergent.
2. Let $a_n = n$. Then (a_n) is unbounded (by the Archimedean principle). Hence (a_n) diverges.

Example 2.2.2. Prove that $n^2 - n^3 + 10 \to -\infty$ as $n \to \infty$.

Solution. Let $A < 0$ and put $K = -A + 11$. Then $K > 11$, and for $n \geq K$ we have

$$\begin{aligned} n^2 - n^3 + 10 &= -n^2(n-1) + 10 \leq -10n^2 + 10 \leq -10K^2 + 10 \\ &= -10(K-1)(K+1) \leq -100(K+1) \leq -K = A - 11 \leq A. \end{aligned}$$

Theorem 2.8. 1. *Let (a_n) be a sequence with $a_n > 0$ for all indices n. Then*

$$\lim_{n\to\infty} a_n = \infty \Leftrightarrow \lim_{n\to\infty} \frac{1}{a_n} = 0.$$

2. *Let (a_n) be a sequence with $a_n < 0$ for all indices n. Then*

$$\lim_{n\to\infty} a_n = -\infty \Leftrightarrow \lim_{n\to\infty} \frac{1}{a_n} = 0.$$

Proof. 1. Assume that $\lim_{n\to\infty} a_n = \infty$. To show that $\lim_{n\to\infty} \frac{1}{a_n} = 0$ let $\varepsilon > 0$. Put $A = \frac{1}{\varepsilon}$. Then there is K such that $0 < A < a_n$ for all $n \geq K$. Then

$$0 < \frac{1}{a_n} < \frac{1}{A} = \varepsilon$$

for these n, which shows that $\lim_{n\to\infty} \frac{1}{a_n} = 0$.

Conversely, assume that $\lim_{n\to\infty} \frac{1}{a_n} = 0$. Let $A \in \mathbb{R}$ and put $A_0 = \max\{A, 1\}$. Then $A_0 > 0$ and put $\varepsilon = \frac{1}{A_0} > 0$. Hence there is $K \in \mathbb{R}$ such that $\frac{1}{a_n} = \left|\frac{1}{a_n}\right| < \varepsilon$ for all $n \geq K$. This give $a_n > \frac{1}{\varepsilon} = A_0 \geq A$ for all $n \geq K$.

The proof of part 2 is similar. □

Apart from this rule, there are more rules for infinite limits. Some of these rules are listed below. [Note to lecturers: do not write this down, just refer the students to the notes.]

Theorem 2.9. *Consider $k \in \mathbb{R}$ and sequences with the following properties as $n \to \infty$:*
$a_n \to \infty$, $\quad b_n \to \infty$, $\quad c_n \to c \in \mathbb{R}$, $\quad d_n \to -\infty$.
Then, as $n \to \infty$,

(a) $ka_n \to \begin{cases} \infty & \text{if } k>0, \\ -\infty & \text{if } k<0, \\ 0 & \text{if } k=0. \end{cases}$

(b) $a_n + b_n \to \infty$,

(c) $a_n + c_n \to \infty$,

(d) $-d_n \to \infty$,

(e) $a_n c_n \to \begin{cases} \infty & \text{if } c>0, \\ -\infty & \text{if } c<0. \end{cases}$

(f) $a_n b_n \to \infty$.

Proof. See tutorials. □

It is convenient to extend the notions of supremum and infimum to unbounded set.

Definition 2.6. Let $A \subset \mathbb{R}$ be nonempty. If A is not bounded above, we write $\sup A = \infty$, and if A is not bounded below, we write $\inf A = -\infty$.

We know that not every sequence converges, and it may be hard to decide if a sequence converges. However, the situation is different with monotonic sequences:

Theorem 2.10. 1. Let (a_n) be an increasing sequence. If (a_n) is bounded, then (a_n) converges, and $\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}$. If (a_n) is not bounded, then (a_n) diverges to ∞.
2. Let (a_n) be a decreasing sequence. If (a_n) is bounded, then (a_n) converges, and $\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}$. If (a_n) is not bounded, then (a_n) diverges to $-\infty$.

Proof. 1. Assume (a_n) is bounded. Put $L = \sup\{a_n : n \in \mathbb{N}\}$ and let $\varepsilon > 0$.
We must show that there is $K \in \mathbb{N}$ such that $n \geq K$ gives $L - \varepsilon < a_n < L + \varepsilon$.
Indeed by the definition of the supremum we have $a_n \leq L < L + \varepsilon$ for all n and by Theorem 1.9 there is K such that $L - \varepsilon < a_K$. Then we have for all $n \geq K$ that

$$L - \varepsilon < a_K \leq a_n.$$

Hence $L - \varepsilon < a_n < L + \varepsilon$ for all $n \geq K$, which proves $a_n \to L$ as $n \to \infty$.
Now assume that (a_n) is not bounded. Since $a_0 \leq a_n$ for all n, (a_n) is bounded below. Hence (a_n) is not bounded above. Let $A \in \mathbb{R}$. Then there is an index K such that $a_K > A$, and thus $a_n \geq a_K > A$ for all $n \geq K$. □

Example 2.2.3. Let $a_n = \left(1+\frac{1}{n}\right)^n$ and $b_n = \left(1+\frac{1}{n}\right)^{n+1}$. Then $a_n < b_n$ for all $n \in \mathbb{N}$, (a_n) is an increasing sequence, (b_n) is a decreasing sequence, both sequences converge to the same number, and the limit is denoted by e, Euler's number.

Proof. The inequality $a_n < b_n$ is obvious from

$$b_n = a_n\left(1+\frac{1}{n}\right).$$

Using Bernoulli's inequality, see Example 1.3.2, we calculate

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\left(\frac{n+2}{n+1}\right)^{n+1}}{\left(\frac{n+1}{n}\right)^n} = \left(\frac{n(n+2)}{(n+1)^2}\right)^{n+1}\frac{n+1}{n}\\ &= \left(1-\frac{1}{(n+1)^2}\right)^{n+1}\frac{n+1}{n}\\ &\geq \left(1-\frac{n+1}{(n+1)^2}\right)\frac{n+1}{n} = \frac{n}{n+1}\,\frac{n+1}{n} = 1\end{aligned}$$

and

$$\begin{aligned}\frac{b_n}{b_{n+1}} &= \frac{\left(\frac{n+1}{n}\right)^{n+1}}{\left(\frac{n+2}{n+1}\right)^{n+2}} = \left(\frac{(n+1)^2}{n(n+2)}\right)^{n+2}\frac{n}{n+1}\\ &= \left(1+\frac{1}{n(n+2)}\right)^{n+2}\frac{n}{n+1}\\ &\geq \left(1+\frac{n+2}{n(n+2)}\right)\frac{n}{n+1} = \frac{n+1}{n}\,\frac{n}{n+1} = 1.\end{aligned}$$

Hence the sequence (a_n) is increasing and bounded (by b_1). By Theorem 2.10, (a_n) has a limit α. Similarly, (b_n) has a limit β. Finally

$$\beta = \lim_{n\to\infty} b_n = \lim_{n\to\infty}\left(\left(1+\frac{1}{n}\right)a_n\right) = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)\lim_{n\to\infty} a_n = \alpha. \qquad \square$$

Definition 2.7. 1. With every sequence (a_n) which is bounded above, we can associate the sequence of numbers

$$\sup\{a_k : k \geq n\}.$$

Since $\{a_k : k \geq n\} \supset \{a_k : k \geq n+1\}$, this sequence is decreasing, and we denote its limit by

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k : k \geq n\}.$$

If (a_n) is not bounded above, we put $\limsup_{n\to\infty} a_n = \infty$.

2. Similarly, if the sequence (a_n) is bounded below, the sequence of numbers

$$\inf\{a_k : k \geq n\}$$

is increasing, and we denote its limit by

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k : k \geq n\}.$$

If (a_n) is not bounded below, we put $\liminf_{n\to\infty} a_n = -\infty$.

Note. 1. Since $\inf S \le \sup S$ for every nonempty set S, it follows that $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$, where we write $-\infty < x$ and $x < \infty$ for each $x \in \mathbb{R}$.
2. The sequence (a_n) is bounded if and only if both $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$ are real numbers (also called finite).

We have now the following characterization of the limit of a sequence and its existence:

Theorem 2.11. 1. *The sequence (a_n) converges if and only if $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$ are finite and equal, and then*

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n.$$

2. $\lim_{n\to\infty} a_n = \infty \Leftrightarrow \liminf_{n\to\infty} a_n = \infty$ *and* $\limsup_{n\to\infty} a_n = \infty$.
3. $\lim_{n\to\infty} a_n = -\infty \Leftrightarrow \liminf_{n\to\infty} a_n = -\infty$ *and* $\limsup_{n\to\infty} a_n = -\infty$.

Proof. 1. For a sequence (a_n), denote $b_n = \inf\{a_k : k \ge n\}$ and $c_n = \sup\{a_k : k \ge n\}$. Note that $b_n \le a_n \le c_n$.
1. Assume that (a_n) converges to L. Let $\varepsilon > 0$. Then there is $K \in \mathbb{N}$ such that for all $n \ge K$, $L - \frac{\varepsilon}{3} < a_n < L + \frac{\varepsilon}{3}$.
Hence

$$L - \frac{\varepsilon}{3} \le b_n \le a_n \le c_n \le L + \frac{\varepsilon}{3}.$$

Since (b_n) is increasing and (c_n) is decreasing, we have

$$L - \frac{\varepsilon}{3} \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le L + \frac{\varepsilon}{3}.$$

Hence

$$0 \le \limsup_{n\to\infty} a_n - \liminf_{n\to\infty} a_n \le \frac{2\varepsilon}{3} < \varepsilon.$$

From Lemma 2.1 we obtain that $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$.
Conversely, if $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$, then it follows from $b_n \le a_n \le c_n$ and the Sandwich Theorem that (a_n) converges and that

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n.$$

2., 3. Exercise. □

Another important concept is that of a Cauchy sequence:

Definition 2.8 (Cauchy sequence)**.** A seqence (a_n) is called a Cauchy sequence if for all $\varepsilon > 0$ there is $K \in \mathbb{R}$ such that for all $n, m \in \mathbb{N}$ with $n, m \ge K$, $|a_n - a_m| < \varepsilon$.

Theorem 2.12. *A sequence (a_n) converges if and only if it is a Cauchy sequence.*

Proof. Let (a_n) be a convergent sequence with limit L. Let $\varepsilon > 0$ and let K such that $|a_n - L| < \frac{\varepsilon}{2}$ for $n \ge K$. Then it follows for $n, m \ge K$ that

$$|a_n - a_m| = |(a_n - L) - (a_m - L)| \le |a_n - L| + |a_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Conversely, assume that (a_n) is a Cauchy sequence. Let $\varepsilon > 0$ and choose K such that $|a_n - a_m| < \frac{\varepsilon}{3}$ for all $m, n \ge K$. Then

$$\begin{aligned}
& a_m - \frac{\varepsilon}{3} < a_n < a_m + \frac{\varepsilon}{3} \\
\Rightarrow\ & a_m - \frac{\varepsilon}{3} \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le a_m + \frac{\varepsilon}{3} \\
\Rightarrow\ & 0 \le \limsup_{n\to\infty} a_n - \liminf_{n\to\infty} a_n \le \left(a_m + \frac{\varepsilon}{3}\right) - \left(a_m - \frac{\varepsilon}{3}\right) = \frac{2\varepsilon}{3} < \varepsilon.
\end{aligned}$$

By Lemma 2.1, $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$, and an application of Theorem 2.11, part 1, completes the proof. □

Tutorial 2.2.1. 1. Prove Theorem 2.9.

2. Use suitable rules or first principles to find

(a) $\lim_{n\to\infty}(n^2+2n-10)$ (b) $\lim_{n\to\infty}(n-\frac{1}{n})$ (c) $\lim_{n\to\infty}\dfrac{n^3-3n^2}{n+1}$

3. Prove that if $\lim_{n\to\infty}|a_n|=\infty$, then (a_n) diverges.

4. Prove that if $p\in\mathbb{N}$, $p>0$, then $n^p\to\infty$ as $n\to\infty$.

5. Define a sequence as follows:

$$a_0=0,\ a_1=\frac{1}{2},\ a_{n+1}=\frac{1}{3}(1+a_n+a_{n-1}^3) \text{ for } n\geq 2.$$

(a) Use induction to show that $0\leq a_n<\frac{2}{3}$ for all $n\in\mathbb{N}$.
(b) Use induction to show that $a_n\leq a_{n+1}$ for all $n\in\mathbb{N}$.
(c) Explain why we may conclude that (a_n) converges.
(d) Using the fact that $\lim_{n\to\infty}a_n=\lim_{n\to\infty}a_{n-1}=\lim_{n\to\infty}a_{n+1}$, find $\lim_{n\to\infty}a_n$.

6. Let $\lim_{n\to\infty}a_n=\infty$, $\lim_{n\to\infty}b_n=\infty$, $\lim_{n\to\infty}c_n=0$. Show, by giving examples, that no general conclusion can be made about the behaviour of the following sequences:
(a) a_n-b_n, (b) a_nc_n, (c) $\dfrac{a_n}{b_n}$, (d) $\dfrac{a_n}{c_n}$.

7. Let (a_n) and (b_n) be sequences such that $a_n\leq b_n$ for all $n\in\mathbb{N}$. Show that $\liminf_{n\to\infty}a_n\leq\liminf_{n\to\infty}b_n$ and $\limsup_{n\to\infty}a_n\leq\limsup_{n\to\infty}b_n$.

8. (a) Show that $\exp(x)=\lim_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n$ exists for all $x\in\mathbb{R}$ and that $\exp(1)=e$.
Hint. Adapt the proof of Example 2.2.3.
(b) Use Bernoulli's inequality to prove that

$$\lim_{n\to\infty}\left(\frac{1+\frac{x+y}{n}}{1+\frac{x+y}{n}+\frac{xy}{n^2}}\right)^n=1$$

for all $x,y\in\mathbb{R}$.
(c) Use (b) to show that $\exp(x+y)=\exp(x)\exp(y)$ for all $x,y\in\mathbb{R}$.
(d) Show that $\exp(x)\geq 1+x$ for all $x>0$.
(e) Show that $\exp(x)>0$ for all $x\in\mathbb{R}$.
(f) Show that exp is strictly increasing.
(g) Show that $\exp(n)=e^n$ for all $n\in\mathbb{Z}$.

Chapter 3

Limits and Continuity of Functions

The single most important concept in all of analysis is that of a limit. Every single notion of analysis is encapsulated in one sense or another to that of a limit. In the previous chapter, you learnt about limits of sequences. Here this notion is extended to limits of functions, which leads to the notion of continuity. You have learnt an intuitive version in Calculus I. Here you will learn a precise definition and you will learn how to prove the results you have learnt in Calculus I.

In this course all functions will have domains and ranges which are subsets of $\mathbb{R}$ unless otherwise stated. Such functions are called real functions.

3.1 Limits of Functions

Definition 3.1. Let $a \in \mathbb{R}$. An interval of the form (c, d) with $c < a < d$ is called a **neighbourhood of** a, and the set $(c, d) \setminus \{a\}$ is called a **deleted neighbourhood of** a.

In the definition of the limit of sequences you have seen formal definitions for 'n tends to ∞' and 'a_n tends to L', and the latter one has an obvious generalization to 'x tends to a' and '$f(x)$ tends to L', which you will encounter in the next definition.

Definition 3.2 (Limit of a function)**.** Let f be a real function, $a, L \in \mathbb{R}$ and assume that the domain of f contains a deleted neighbourhood of a, that is, $f(x)$ is defined for all x in a deleted neighbourhood of f.

Then '$f(x) \to L$ as $x \to a$' is defined to mean:

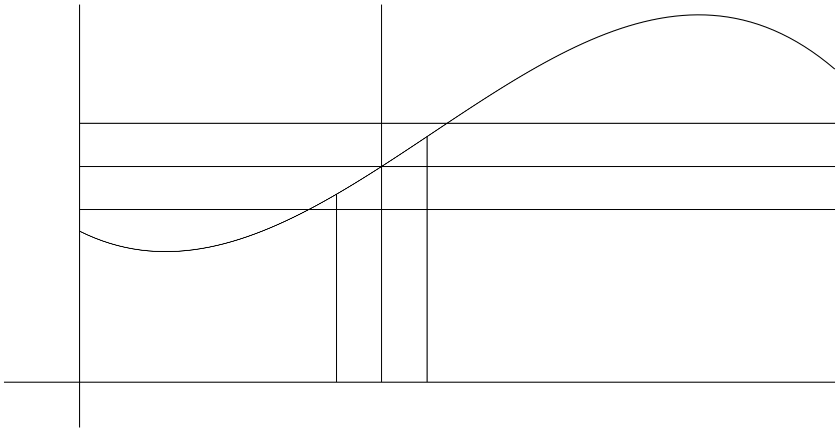
$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x \in (a - \delta, a + \delta) \setminus \{a\}, \; f(x) \in (L - \varepsilon, L + \varepsilon)$,
i. e., $\forall \varepsilon > 0 \, \exists \delta > 0 \, (0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon)$.

If $f(x) \to L$ as $x \to a$, then we write $\lim\limits_{x \to a} f(x) = L$.

Note. 1. L in the definition above is also called the limit of the function at a. For the definition of the limit of a function at a, the number $f(a)$ is never used, and indeed, the function need not be defined at a.

2. Observe that the number δ will depend on ε, and also on a, if we vary a.

Exercise. Use the notation in the above definition to label the following graph:



Example 3.1.1. Prove, from first principles, that $x^2 \to 4$ as $x \to 2$.

Solution. Let $\varepsilon > 0$. We must show that there is $\delta > 0$ such that $0 < |x-2| < \delta$ implies $|x^2-4| < \varepsilon$. Since we want to use $|x-2| < \delta$ to get $|x^2-4| < \varepsilon$, we try to factor out $|x-2|$ from $|x^2-4|$. Thus

$$|x^2-4| = |x-2|\,|x+2| \leq |x-2|\,(|x-2|+4).$$

It is often convenient to make an initial restriction on δ, like $\delta \leq 1$. Then we can continue the above estimate to obtain from $|x-a| < \delta$ that

$$|x^2-4| \leq |x-2|\,(1+4) = 5|x-2|.$$

If we now put

$$\delta = \min\left\{1, \frac{\varepsilon}{5}\right\},$$

then we can conclude (note that we already used $\delta \leq 1$) for $0 < |x-2| < \delta$ that

$$|x^2-4| \leq 5|x-2| < 5\frac{\varepsilon}{5} = \varepsilon.$$

To justify the notation $\lim\limits_{x\to a} f(x) = L$ we have to show

Theorem 3.1. *If $f(x) \to L$ as $x \to a$, then L is unique.*

Proof. Let $f(x) \to L$ and $f(x) \to M$ as $x \to a$. We must show that $L = M$.
So let $\varepsilon > 0$. Then there are $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$0 < |x-a| < \delta_1 \Rightarrow |f(x)-L| < \frac{\varepsilon}{2} \quad \text{and} \quad 0 < |x-a| < \delta_2 \Rightarrow |f(x)-M| < \frac{\varepsilon}{2}.$$

For $\delta = \min\{\delta_1, \delta_2\}$ and $0 < |x-a| < \delta$ (we indeed only need one such x here) it follows that

$$|L-M| = |(L-f(x)) - (f(x)-M)| \leq |f(x)-L| + |f(x)-M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $L = M$ by Lemma 2.1. □

Definition 3.3 (One sided limits)**.** 1. Let f be a real function, $a, L \in \mathbb{R}$ and assume that the domain of f contains an interval (a,d) with $d > a$, that is, $f(x)$ is defined for all x in (a,d).
Then '$f(x) \to L$ as $x \to a^+$' is defined to mean:
$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ (a < x < a+\delta \Rightarrow |f(x)-L| < \varepsilon)$.
If $f(x) \to L$ as $x \to a^+$, then we write $\lim\limits_{x\to a^+} f(x) = L$.
2. Let f be a real function, $a, L \in \mathbb{R}$ and assume that the domain of f contains an interval (c,a) with $c < a$, that is, $f(x)$ is defined for all x in (c,a).
Then '$f(x) \to L$ as $x \to a^-$' is defined to mean:
$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ (a-\delta < x < a \Rightarrow |f(x)-L| < \varepsilon)$.
If $f(x) \to L$ as $x \to a^-$, then we write $\lim\limits_{x\to a^-} f(x) = L$.

Theorem 3.2. *If $f(x)$ is defined in a deleted neighbourhood of a, then*

$$\lim_{x\to a} f(x) = L \Leftrightarrow \lim_{x\to a^-} f(x) = L = \lim_{x\to a^+} f(x).$$

Proof. $\Rightarrow$: Assume that $\lim_{x\to a} f(x) = L$ and let $\varepsilon > 0$. Then

$$\exists\, \delta > 0\ (x \in (a-\delta, a+\delta),\ x \neq a \Rightarrow |f(x) - L| < \varepsilon)$$

gives

$$\exists\, \delta > 0\ (x \in (a-\delta, a) \Rightarrow |f(x) - L| < \varepsilon) \quad \text{and} \quad \exists\, \delta > 0\ (x \in (a, a+\delta) \Rightarrow |f(x) - L| < \varepsilon).$$

Hence $\lim_{x\to a^-} f(x) = L$ and $\lim_{x\to a^+} f(x) = L$.
$\Leftarrow$: Assume that $\lim_{x\to a^-} f(x) = L = \lim_{x\to a^+} f(x)$ and let $\varepsilon > 0$. Then there are $\delta_- > 0$ and $\delta_+ > 0$ such that

$$x \in (a-\delta_-, a) \Rightarrow |f(x) - L| < \varepsilon \quad \text{and} \quad x \in (a, a+\delta_+) \Rightarrow |f(x) - L| < \varepsilon.$$

Let $\delta = \min\{\delta_-, \delta_+\}$. Then

$$\begin{aligned} x \in (a-\delta, a+\delta),\ x \neq a &\Rightarrow x \in (a-\delta_-, a) \text{ or } x \in (a, a+\delta_+) \\ &\Rightarrow |f(x) - L| < \varepsilon. \end{aligned}$$

This proves that $\lim_{x\to a} f(x) = L$. □

Example 3.1.2. Let $f(x) = \dfrac{x}{|x|}$ for $x \in \mathbb{R} \setminus \{0\}$. Then $f(x) = 1$ if $x > 0$ and $f(x) = -1$ if $x < 0$. It is easy to prove from the definition that $\lim_{x\to 0^+} f(x) = 1$ and $\lim_{x\to 0^-} f(x) = -1$. Now it follows from Theorem 3.2 that the function f does not have a limit as x tends to a.

Tutorial 3.1.1. 1. Prove from the definitions that
(a) $2 + x + x^2 \to 8$ as $x \to 2$, (b) $\dfrac{1}{x-2} \to -\dfrac{2}{3}$ as $x \to \dfrac{1}{2}$, (c) $\lim_{x\to 1^-} \sqrt{1-x} = 0$.
2. By negating the definition of limit of a function show that the statement $f(x) \not\to L$ as $x \to a$ is equivalent to the following:
$\exists\, \varepsilon > 0\ \forall\, \delta > 0\ \exists\, x$ with $0 < |x - a| < \delta$ such that $|f(x) - L| \geq \varepsilon$.
3. For $x \in \mathbb{R} \setminus \{0\}$ let $f(x) = \sin \dfrac{1}{x}$. Prove that f does not tend to any limit as $x \to 0$.
4. Let $f(x) = x - \lfloor x \rfloor$. For each integer n, find $\lim_{x\to n^-} f(x)$ and $\lim_{x\to n^+} f(x)$ if they exist.

3.2 Limits at Infinity and Infinite Limits

In $x \to a$ the inequalities $0 < |x - a| < \delta$ occur. For $x \to \infty$, this has to be replaced by $x > K$. Hence we have the following

Definition 3.4. 1. Let f be a real function defined on a set containing an interval of the form (c, ∞). Then '$f(x) \to L$ as $x \to \infty$' is defined to mean:
$\forall\, \varepsilon > 0\ \exists\, K(> 0)\ (x > K \Rightarrow |f(x) - L| < \varepsilon)$.
If $f(x) \to L$ as $x \to \infty$, then we write $\lim_{x\to\infty} f(x) = L$.
2. Let f be a real function defined on a set containing an interval of the form $(-\infty, c)$.
Then '$f(x) \to L$ as $x \to -\infty$' is defined to mean:
$\forall\, \varepsilon > 0\ \exists\, K(< 0)\ (x < K \Rightarrow |f(x) - L| < \varepsilon)$.
If $f(x) \to L$ as $x \to -\infty$, then we write $\lim_{x\to-\infty} f(x) = L$.

Example 3.2.1. Let $f(x) = \dfrac{1}{x}$. Find $\lim\limits_{x\to\infty} f(x)$ and $\lim\limits_{x\to-\infty} f(x)$.

Solution. Guess that both limits are 0. Let $\varepsilon > 0$ and put $K = \dfrac{1}{\varepsilon}$. Then $x > K$ gives

$$\left|\frac{1}{x}\right| = \frac{1}{x} < \frac{1}{K} = \varepsilon.$$

So $\lim\limits_{x\to\infty} f(x) = 0$.
Now let $\varepsilon > 0$ and put $K = -\dfrac{1}{\varepsilon}$. Then $x < K$ gives

$$\left|\frac{1}{x}\right| = -\frac{1}{x} < \frac{1}{-K} = \varepsilon.$$

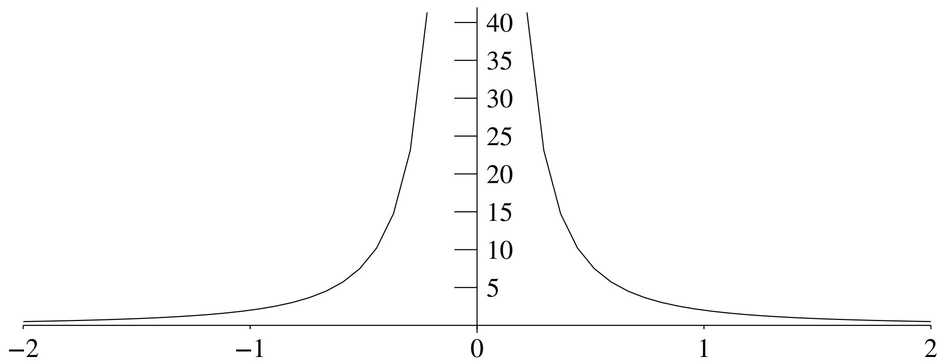
So $\lim\limits_{x\to-\infty} f(x) = 0$.

For infinite limits we have similar definitions:

Definition 3.5. Let f be a real function whose domain includes a deleted neighbourhood of the number a.
1. '$f(x) \to \infty$ as $x \to a$' is defined to mean:
$\forall\, K(> 0)\ \exists\, \delta > 0\ (0 < |x - a| < \delta \Rightarrow f(x) > K)$.
If $f(x) \to \infty$ as $x \to a$, then we write $\lim\limits_{x\to a} f(x) = \infty$.
2. '$f(x) \to -\infty$ as $x \to a$' is defined to mean:
$\forall\, K(< 0)\ \exists\, \delta > 0\ (0 < |x - a| < \delta \Rightarrow f(x) < K)$.
If $f(x) \to -\infty$ as $x \to a$, then we write $\lim\limits_{x\to a} f(x) = -\infty$.

Example 3.2.2. Prove that $\dfrac{2}{x^2} \to \infty$ as $x \to 0$.

Solution. The following sketch illustrates the result.



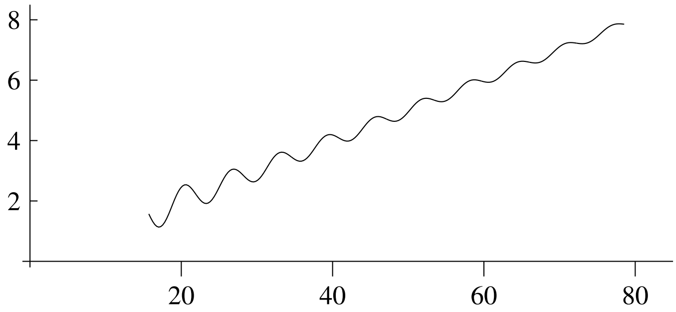
We may choose K to be sufficiently large. So choose $K > 1$ and let $\delta = \frac{1}{K}$. Then, for $0 < |x| < \delta$, and because of $\delta < 1$ we have

$$\frac{2}{x^2} = \frac{1}{|x|}\frac{2}{|x|} \geq \frac{2}{|x|} > \frac{2}{\delta} = 2K > K.$$

Definition 3.6. Let f be a real function defined on a set containing an interval of the form (c, ∞).
Then '$f(x) \to \infty$ as $x \to \infty$' is defined to mean:
$\forall\, A(> 0)\ \exists\, K(> 0)\ (x > K \Rightarrow f(x) > A)$.
If $f(x) \to \infty$ as $x \to \infty$, then we write $\lim\limits_{x\to\infty} f(x) = \infty$.

Similar definition hold with all other combinations of limits involving ∞ and $-\infty$, including one-sided limits.

Example 3.2.3. Prove that $f(x) = \dfrac{x}{10} + 10\,\dfrac{\sin x}{x} \to \infty$ as $x \to \infty$. A section of the function's graph is plotted below.



Additional notes. 1. You have learnt different notations for limits depending on whether real numbers, ∞ or $-\infty$ are considered. However, using the notion of (deleted) neighbourhood, one can give just one definition which applies to each case. To define (deleted) neighbourhoods of ∞ and $-\infty$, we first observe that ∞ and $-\infty$ are just symbols for our purposes, and thus never belong to sets of numbers which we consider. Therefore, no distinction is made here between neighbourhoods and deleted neighbourhoods of ∞ and $-\infty$, respectively.
A (deleted) neighbourhood of ∞ is any set of the form (c, ∞) with $c \in \mathbb{R}$.
A (deleted) neighbourhood of $-\infty$ is any set of the form $(-\infty, d)$ with $d \in \mathbb{R}$.
Now let $a, L \in \mathbb{R} \cup \{\infty, -\infty\}$ and f be a function defined in a deleted neighbourhood of a. Then $f(x) \to L$ as $x \to a$ if and only if for each neighbourhood U of L there is a deleted neighbourhood V of a such that $x \in V$ implies $f(x) \in U$.
If we define one-sided deleted neighbourhoods of $a \in \mathbb{R}$ to be the sets (c, a) and (a, d), respectively, then the above formulation also covers one-sided limits.
2. What we call neighbourhood here should actually be called basic neighbourhood. In general topology, a neighborhood is any subset (of a given set) which contains a basic neighbourhood. But since we do not need general neighbourhoods, we conveniently dropped the word 'basic'.

Example 3.2.4. Let $f(x) = \dfrac{1}{x}$. Find $\lim\limits_{x\to 0^-} f(x)$, $\lim\limits_{x\to 0^+} f(x)$, $\lim\limits_{x\to 0} f(x)$.

Solution. Let $K > 0$ and put $\delta = \dfrac{1}{K}$. Then, for $0 < x < \delta$, $\dfrac{1}{x} > \dfrac{1}{\delta} = K$. Thus $\lim\limits_{x\to 0^+} f(x) = \infty$.
Now let $K < 0$ and put $\delta = -\dfrac{1}{K} > 0$. Then, for $-\delta < x < 0$, $\dfrac{1}{x} < \dfrac{1}{-\delta} = K$. Thus $\lim\limits_{x\to 0^-} f(x) = -\infty$.
Since $\lim\limits_{x\to 0^-} f(x) \neq \lim\limits_{x\to 0^+} f(x)$, $\lim\limits_{x\to 0} f(x)$ does not exist.

Tutorial 3.2.1. 1. Prove from the definitions that
(a) $\lim\limits_{x\to-\infty} \dfrac{1-3x}{2x-1} = -\dfrac{3}{2}$, (b) $\dfrac{1}{x-1} \to \infty$ as $x \to 1^+$, (c) $\dfrac{1}{x-1} \to -\infty$ as $x \to 1^-$.

3.3 Limit Laws

Rather than calculating limits from the definition, in general one will use limit laws. In this section we state and prove some of these laws.

Theorem 3.3 (Limit Laws)**.** *Let $a, c \in \mathbb{R}$ and suppose that the real functions f and g are defined in a deleted neighbourhood of a and that $\lim_{x\to a} f(x) = L \in \mathbb{R}$ and $\lim_{x\to a} g(x) = M \in \mathbb{R}$ both exist. Then*

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

(b) $\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = L + M$.

(c) $\lim_{x\to a}[f(x) - g(x)] = \lim_{x\to a} f(x) - \lim_{x\to a} g(x) = L - M$.

(d) $\lim_{x\to a}[cf(x)] = c\lim_{x\to a} f(x) = cL$.

(e) $\lim_{x\to a}[f(x)g(x)] = \left[\lim_{x\to a} f(x)\right]\left[\lim_{x\to a} g(x)\right] = LM$.

(f) *If* $M \neq 0$, $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \dfrac{L}{M}$.

(g) *If* $L \neq 0$ *and* $M = 0$, $\lim_{x\to a} \dfrac{f(x)}{g(x)}$ *does not exist.*

(h) *If* $n \in \mathbb{N}$, $\lim_{x\to a}[f(x)^n] = \left[\lim_{x\to a} f(x)\right]^n = L^n$.

(i) $\lim_{x\to a} x = a$.

(j) *If* $n \in \mathbb{N}$, $\lim_{x\to a} x^n = a^n$.

(k) *If* $n \in \mathbb{N}$, $\lim_{x\to a} \sqrt[n]{x} = \sqrt[n]{a}$. *If n is even, we assume that $a > 0$.*

(l) *If* $n \in \mathbb{N}$, $\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x\to a} f(x)} = \sqrt[n]{L}$.

(m) *If* $\lim_{x\to a} |f(x)| = 0$, *then* $\lim_{x\to a} f(x) = 0$.

Proof. The proofs are similar to those in Theorem 2.3 and we will only prove (b), (e), (f).
(b) Let $\varepsilon > 0$. Then there are numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that
(i) $|f(x) - L| < \frac{\varepsilon}{2}$ if $0 < |x - a| < \delta_1$,
(ii) $|g(x) - M| < \frac{\varepsilon}{2}$ if $0 < |x - a| < \delta_2$.
Put $\delta = \min\{\delta_1, \delta_2\}$. For $0 < |x - a| < \delta$ we have $0 < |x - a| < \delta_1$ and $0 < |x - a| < \delta_2$ and therefore

$$|(f(x) + g(x)) - (L + M)| = |(f(x) - L) + (g(x) - M)| \le |f(x) - L| + |g(x) - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $f(x) + g(x)$ converges to $L + M$ as $x \to a$.
(e) We consider 2 cases: one special case, to which then the general case is reduced.
Case I: $L = M = 0$. Let $\varepsilon > 0$ and put $\varepsilon' = \min\{1, \varepsilon\}$. Then $0 < \varepsilon' \le \varepsilon$, and there are numbers δ_1 and δ_2 such that
(i) $|f(x)| < \varepsilon'$ if $0 < |x - a| < \delta_1$,
(ii) $|g(x)| < \varepsilon'$ if $0 < |x - a| < \delta_2$.
Put $\delta = \min\{\delta_1, \delta_2\}$. For $0 < |x - a| < \delta$ we have

$$|f(x)g(x)| = |f(x)|\,|g(x)| < {\varepsilon'}^2 \le \varepsilon' \le \varepsilon.$$

Case II: L and M are arbitrary. Then

$$f(x)g(x) = (f(x) - L)(g(x) - M) + L(g(x) - M) + f(x)M.$$

By (a), (b), (c), $(f(x) - M) \to 0$ and $(g(x) - M) \to 0$ as $x \to a$, and by (b), (c), (d), and Case I, it follows that

$$\lim_{x\to a} f(x)g(x) = 0 + L \cdot 0 + LM = LM.$$

(f) First consider $f = 1$. Since $M \neq 0$ and $g(x) \to M$ as $x \to a$, there is $\delta_0 > 0$ such that $|g(x) - M| < \frac{|M|}{2}$ for $0 < |x - a| < \delta_0$. Then, for $0 < |x - a| < \delta_0$,

$$\begin{aligned} |g(x)| &\geq |M| - |g(x) - M| \geq |M| - \frac{|M|}{2} = \frac{|M|}{2} \\ \Rightarrow \left|\frac{1}{g(x)} - \frac{1}{M}\right| &= \frac{|M - g(x)|}{|g(x)M|} \leq \frac{2|g(x) - M|}{|M|^2}. \end{aligned}$$

Now let $\varepsilon > 0$ and $\delta_1 > 0$ such that $0 < |x - a| < \delta_1$ implies $|g(x) - M| \leq \frac{|M|^2}{2}\varepsilon$. Put $\delta = \min\{\delta_0, \delta_1\}$. It follows for $0 < |x - a| < \delta$ that

$$\left|\frac{1}{g(x)} - \frac{1}{M}\right| \leq \frac{2|g(x) - M|}{|M|^2} < 2\frac{|M|^2}{2}\varepsilon\,\frac{1}{|M|^2} = \varepsilon.$$

The general case now follows with (d):

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} f(x) \cdot \lim_{x\to a} \frac{1}{g(x)} = L \cdot \frac{1}{M}.$$

□

Recall that a polynomial function is of the form

$$f(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_2x^2 + b_1x + b_0$$

with $b_i \in \mathbb{R}$ for $i = 1, 2, \ldots, n$ and n any non-negative integer. A rational function is of the form $f(x) = \frac{p(x)}{q(x)}$ with $p(x)$ and $q(x)$ polynomials. We then have the following as a consequence of Theorem 2.3, (b),(d),(i),(j).

Corollary 3.4. *If f is a polynomial or a rational function and a is in the domain of f, then $\lim_{x\to a} f(x) = f(a)$.*

Corollary 3.5. *All the limit rules in Theorem 3.3 remain true if $x \to a$ is replaced by any of the following: $x \to a^+$, $x \to a^-$, $x \to \infty$, $x \to -\infty$.*

Proof. For $x \to a^+$ and $x \to a^-$ one just has to replace $0 < |x - a| < \delta$ with $0 < x - a < \delta$ and $-\delta < -(x - a) < 0$, respectively, in the proof of each of the statements. For $x \to \infty$ and $x \to -\infty$, the proofs are very similar to those for sequences. □

Similar rules hold if the functions have infinite limits. We state some of the results for $x \to a$, observing that there are obvious extensions as in Corollary 3.5.

Theorem 3.6. *Assume that $\lim_{x\to a} f(x) = \infty$, $\lim_{x\to a} g(x) = \infty$ and $\lim_{x\to a} h(x) = c \in \mathbb{R}$. Then*
(a) $f(x) + g(x) \to \infty$ *as* $x \to a$,
(b) $f(x) + h(x) \to \infty$ *as* $x \to a$,
(c) $f(x)g(x) \to \infty$ *as* $x \to a$,
(d) $f(x)h(x) \to \begin{cases} \infty & \text{if } c>0 \\ -\infty & \text{if } c<0 \end{cases}$ *as* $x \to a$.

Proof. We prove (c) and leave the other parts as exercises.
Let $A > 1$ (we can take any $A \in \mathbb{R}$, but the restriction $A > 1$ is convenient here). Then there are $\delta_1 > 0$ and $\delta_2 > 0$ such that
(i) $f(x) > A$ if $0 < |x - a| < \delta_1$,
(ii) $g(x) > A$ if $0 < |x - a| < \delta_2$.

With $\delta = \min\{\delta_1, \delta_2\}$ it follows for $0 < |x-a| < \delta$ that

$$f(x)g(x) > A \cdot A > A.$$

$\square$

Theorem 3.7 (Sandwich Theorem)**.** *Let $a \in \mathbb{R} \cup \{\infty, -\infty\}$ and assume that f, g and h are real functions defined in a deleted neighbourhood of a. If $f(x) \le g(x) \le h(x)$ for x in a deleted neighbourhood of a and $\lim_{x\to a} f(x) = L = \lim_{x\to a} h(x)$, then $\lim_{x\to a} g(x) = L$.*

Proof. Note that $L \in \mathbb{R} \cup \{\infty, -\infty\}$. We will prove this theorem in the case $a \in \mathbb{R}$ and $L \in \mathbb{R}$. The other cases are left as an exercise.
Let $\varepsilon > 0$. Then there are δ_1 and δ_2 such that
(i) $|f(x) - L| < \varepsilon$ if $0 < |x-a| < \delta_1$,
(ii) $|h(x) - L| < \varepsilon$ if $0 < |x-a| < \delta_2$.
Put $\delta = \min\{\delta_1, \delta_2\}$. Then, for $0 < |x-a| < \delta$,

$$L - \varepsilon < f(x) < L + \varepsilon, \quad L - \varepsilon < h(x) < L + \varepsilon$$

gives

$$L - \varepsilon < f(x) \le g(x) \le h(x) < L + \varepsilon.$$

Hence $|g(x) - L| < \varepsilon$ if $0 < |x-a| < \varepsilon$. $\square$

Theorem 3.8. *Let f be defined on an interval (a, b), where $a = -\infty$ and $b = \infty$ are allowed. With the convenient notation $a^+ = -\infty$ if $a = -\infty$ and $b^- = \infty$ if $b = \infty$, we obtain*
(a) *if f is increasing, then*
$\lim_{x\to b^-} f(x) = \sup\{f(x) : x \in (a,b)\}$ *and* $\lim_{x\to a^+} f(x) = \inf\{f(x) : x \in (a,b)\}$*;*
(b) *if f is decreasing, then*
$\lim_{x\to b^-} f(x) = \inf\{f(x) : x \in (a,b)\}$ *and* $\lim_{x\to a^+} f(x) = \sup\{f(x) : x \in (a,b)\}$.

Proof. Since all four cases have similar proofs, we only prove (a) in the case $b \in \mathbb{R}$.
Let $L = \sup\{f(x) : x \in (a,b)\}$.
Case I: $L \in \mathbb{R}$.
Let $\varepsilon > 0$. By Theorem 1.9 there is $c \in (a,b)$ such that $L - \varepsilon < f(c)$. Put $\delta = b - c > 0$. Now let $b - \delta < x < b$, i. e., $x \in (c,b)$. Then $c < x$ gives $f(c) \le f(x)$ since f is increasing and $f(x) \le L$ for all $x \in (c,b) \subset (a,b)$ by definition of the supremum, so that

$$L - \varepsilon < f(c) \le f(x) \le L < L + \varepsilon$$

for these x. By definition, this mean $f(x) \to L$ as $x \to b^-$.
Case II: $L = \infty$. In this case, $\{f(x) : x \in (a,b)\}$ is not bounded above. Therefore, for each $A \in \mathbb{R}$ there is $c \in (a,b)$ such that $f(c) > A$. Since f is increasing, it follows for all $x \in (c,b)$ that $A < f(c) \le f(x)$. Therefore $f(x) \to \infty$ as $x \to b^-$. $\square$

Tutorial 3.3.1. 1. Let n be a positive integer. Prove that
(a) $\lim_{x\to\infty} x^n = \infty$,
(b) $\lim_{x\to-\infty} x^n = \begin{cases} \infty & \text{if n is even,} \\ -\infty & \text{if n is odd,} \end{cases}$
(c) $\lim_{x\to 0^+} x^{-n} = \infty$,
(d) $\lim_{x\to 0^-} x^{-n} = \begin{cases} \infty & \text{if n is even,} \\ -\infty & \text{if n is odd.} \end{cases}$
2. (a) Let f, g be defined in a deleted neighbourhood of a and assume that $f(x) < g(x)$ for all x in a deleted neighbourhood of a. Show that if $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$ exist, then $L \le M$.
(b) Give examples for $L < M$ and for $L = M$ in (a).
(c) Formulate and prove the result corresponding to (a) for one-sided limits.

3. Using rules for limits, determine the behaviour of $f(x)$ as x tends to the given limit:

(a) $f(x) = \dfrac{4x}{3-x}$ as $x \to 3^-$,

(b) $f(x) = \dfrac{(x-4)(x-1)}{x-2}$ as $x \to 2^+$,

(c) $f(x) = \dfrac{2x+1}{x^2-x}$ as $x \to 0^+$.

3.4 Continuity of Functions

Definition 3.7 (Continuity of a function at a point)**.** Let f be a real function, $a \in \mathbb{R}$ and assume that the domain of f contains a neighbourhood of a, that is, $f(x)$ is defined for all x in a neighbourhood of a.
We say that f is **continuous at** a if
$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x \in (a-\delta, a+\delta), \; f(x) \in (f(a)-\varepsilon, f(a)+\varepsilon)$,
i. e., $\forall \varepsilon > 0 \, \exists \delta > 0 \, (|x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon)$.

Note that the sketch on page 25 also illustrates the definition of continuity.

We realize that the definition of the existence of a limit of a function at a point and continuity at that point are very similar, but that there are subtle (and important) differences:
For limits, f does not need to be defined at a, and even if $f(a)$ exists, this value is not used at all when finding the limit of the function f at a.

We conclude

$$\begin{aligned} f \text{ is continuous at } a &\Leftrightarrow \forall \varepsilon > 0 \, \exists \delta > 0 \, (|x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon) \quad \text{(by definition)} \\ &\Leftrightarrow \forall \varepsilon > 0 \, \exists \delta > 0 \, (0 < |x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon) \quad (\because x = a \Rightarrow f(x) = f(a)) \\ &\Leftrightarrow \lim_{x \to a} f(x) = f(a). \end{aligned}$$

Hence we have shown

Theorem 3.9. *f is continuous at a if and only if the following three conditions are satisfied:*

1. $f(a)$ is defined, i. e., a is in the domain of f,

2. $\lim_{x \to a} f(x)$ exists, i.e., $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x)$, and

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

Example 3.4.1. 1. Let

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

Then $\lim_{x \to 2} f(x) = 4$ exists, but this limit is different from $f(2) = 2$. Hence f is not continuous at 2.

2. Let $f(x) = \dfrac{\sin x}{x}$ for $x \neq 0$ while f is not defined at $x = 0$. Then $\lim_{x \to 0} f(x) = 1$ exists, but f is not defined at 0. Hence f is not continuous at 0.
3. Let

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Then $\lim_{x \to 0} f(x) = 1$ exists, and $f(0) = 1$. Hence f is continuous at 0.

Theorem 3.10. *If f and g are continuous at a and if $c \in \mathbb{R}$, then*

1. *the sum $f + g$,*
2. *the difference $f - g$,*
3. *the product fg,*
4. *the quotient $\dfrac{f}{g}$ if $g(a) \neq 0$ and*
5. *the scalar multiple cf*

are functions that are also continuous at a.

Proof. The statements follow immediate from the limit laws, Theorem 3.3, and Theorem 3.9. For example, for 3. we have

$$\lim_{x\to a} f(x) = f(a) \text{ and } \lim_{x\to a} g(x) = g(a),$$

and then Theorem 3.3 gives

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

Then Theorem 3.9 says that fg is continuous at a. □

Recall that the composite $g\circ f$ of two functions f and g is defined by $(g\circ f)(x) = g(f(x))$.

Theorem 3.11. *If f is continuous at a and g is continuous at $f(a)$, then $g\circ f$ is continuous at a.*

Proof. Let $\varepsilon > 0$. Since g is continuous at $f(a)$, there is $\eta > 0$ such that

$$|y - f(a)| < \eta \Rightarrow |g(y) - g(f(a))| < \varepsilon. \tag{1}$$

Since f is continuous at a, there is $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \eta. \tag{2}$$

Putting $y = f(x)$ in (1) it follows from (1) and (2) that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \eta \Rightarrow |g(f(x)) - g(f(a))| < \varepsilon$$

that is,

$$|x - a| < \delta \Rightarrow |(g\circ f)(x) - (g\circ f)(a)| < \varepsilon.$$

Hence $g\circ f$ is continuous at a. □

Definition 3.8. 1. A function f is **continuous from the right at** a if $\lim_{x\to a^+} f(x) = f(a)$.
2. A function f is **continuous from the left at** a if $\lim_{x\to a^-} f(x) = f(a)$.

Example 3.4.2. 1. Let

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

Determine the right and left continuity of f at $x = 0$.

Solution. $f(0) = 0$ whilst

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} \frac{|x| + x}{2x} = \lim_{x\to 0^-} \frac{-x + x}{2x} = \lim_{x\to 0^-} 0 = 0$$

and

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} \frac{|x| + x}{2x} = \lim_{x\to 0^+} \frac{x + x}{2x} = \lim_{x\to 0^-} 1 = 1.$$

Since $f(0) = 0 = \lim_{x\to 0^-} f(x)$, f is continuous from the left at $x = 0$. Since $f(0) = 0 \neq 1 = \lim_{x\to 0^+} f(x)$, f is not continuous from the right at $x = 0$.

Note. 1. It is easy to show that f is continuous at a if and only if f is continuous from the right and continuous from the left at a.
2. If $a \in \mathrm{dom}(f)$ and if there is $\varepsilon > 0$ such that $\mathrm{dom}(f) \cap (a-\varepsilon, a+\varepsilon) = (a-\varepsilon, a]$, then we say that f is continuous at a if $\lim_{x\to a^-} = f(a)$.
3. If $a \in \mathrm{dom}(f)$ and if there is $\varepsilon > 0$ such that $\mathrm{dom}(f) \cap (a-\varepsilon, a+\varepsilon) = [a, a+\varepsilon)$, then we say that f is continuous at a if $\lim_{x\to a^+} = f(a)$.
4. The convention in 2 and 3 is consistent with what you will learn in General Topology about continuity. Just note that the condition $|f(x) - f(a)| < \varepsilon$ has to be checked for all $x \in \mathrm{dom}(f)$ which satisfy $|x - a| < \delta$.

Lemma 3.12. *If $f(x) \to b$ as $x \to a$ (a^+, a^-) and g is continuous at b, then $g(f(x)) \to g(b)$ as $x \to a$ (a^+, a^-), which can be written, e. g., as*

$$\lim_{x\to a} g(f(x)) = g\left(\lim_{x\to a} f(x)\right).$$

Proof. The function

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \mathrm{dom}(f),\ x \neq a, \\ b & \text{if } x = a, \end{cases}$$

is continuous (from the right, from the left) at a. Hence the result follows from Theorem 3.11. □

Definition 3.9. A function is continuous on a set $X \subset \mathbb{R}$ if f is continuous at each $x \in X$. Here continuity is understood in the sense of the above note with $X = \mathrm{dom}(f)$.

Example 3.4.3. Show that $f(x) = \sqrt{x^2 - 4}$ is continuous.

Solution. The domain of f is $\{x \in \mathbb{R} : |x| \geq 2\} = (-\infty, -2] \cup [2, \infty)$. By Theorem 3.10, the function $x \mapsto x^2 - 4$ is continuous on $\mathbb{R}$, and by Theorem 3.3 (k), the square root is continuous at each positive number. So also the composite function f is continuous at a. Also, the proof of Theorem 3.3 (k) can be easily adapted to show that the square root is continuous from the right at 0. Then it easily follows that f is continuous (from the right) at 2 and continuous (from the left) at -2.

Theorem 3.13. *The following functions are continuous on their domains.*

1. *Polynomials* $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0$, $a_i \in \mathbb{R}$, $n \in \mathbb{N}$.
2. *Rational functions* $\dfrac{p(x)}{q(x)}$, *p and $q \neq 0$ polynomials.*
3. *Sums, differences, products and quotients of continuous functions.*
4. *Root functions.*
5. *The trigonometric functions* $\sin x$, $\cos x$, $\tan x$, $\operatorname{cosec} x$, $\sec x$ *and* $\cot x$.
6. *The exponential function* $\exp(x)$.
7. *The absolute value function* $|x|$.

Proof. 1, 2 and 3 easily follow from previous theorems on limits and continuity, as does 7. However, 7 can be easily proved dircetly: For each $\varepsilon > 0$ let $\delta = \varepsilon$. Then, for $|x - a| < \delta$ we have

$$\big|\,|x| - |a|\,\big| \leq |x - a| < \delta = \varepsilon.$$

The continuity of sin and cos follows from the sum of angles formulae and from the limits proved in Calculus I (the proofs used the Sandwich Theorem, which now has been proved). The continuity of the other trigonometric functions then follows from part 3.
Finally, the continuity of exp is a tutorial problem. □

Theorem 3.14. *Let $a \in \mathbb{R}$ and let f be a real function which is defined in a neighbourhood of a. Then f is continuous at a if and only if for each sequence (x_n) in* $\mathrm{dom}(f)$ *with $\lim_{n\to\infty} x_n = a$ the sequence $f(x_n)$ satisfies $\lim_{n\to\infty} f(x_n) = f(a)$.*

Proof. $\Rightarrow$: Let (x_n) be a sequence in $\mathrm{dom}(f)$ with $\lim_{n\to\infty} x_n = a$. We must show that $\lim_{n\to\infty} f(x_n) = f(a)$. Hence let $\varepsilon > 0$. Since f is continuous at a, there is $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon. \tag{1}$$

Since $\lim_{n\to\infty} x_n = a$, there is $K \in \mathbb{R}$ such that for $n > K$, $|x_n - a| < \delta$. But then, by (1), $|f(x_n) - f(a)| < \varepsilon$ for $n > K$.

$\Leftarrow$: (indirect proof) Assume that f is not continuous at a. Then

$$\exists\, \varepsilon > 0\ \forall\, \delta > 0\ \exists\, x \in \mathrm{dom}(f),\ |x - a| < \delta \text{ and } |f(x) - f(a)| \geq \varepsilon.$$

In particular, for $\delta = \frac{1}{n}$, $n = 1, 2, \ldots$ we find $x_n \in \mathrm{dom}(f)$ such that $|x_n - a| < \dfrac{1}{n}$ and $|f(x_n) - f(a)| \geq \varepsilon$. But then $\lim_{n\to\infty} x_n = a$, whereas $(f(x_n))$ does not converge to $f(a)$. ☐

Tutorial 3.4.1. 1. Consider the function

$$f(x) = \begin{cases} \dfrac{\lfloor x \rfloor}{x} & \text{if } x \neq 0, \\ -1 & \text{if } x = 0. \end{cases}$$

Investigate continuity from the left and the right at $x = 0$, $x = \pi$ and $x = 1$.

2. Let $f(x) = x \sin \frac{1}{x}$ for $x \neq 0$ and $g(x) = \begin{cases} 0 & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$
Show that $f(x) \to 0$ as $x \to 0$ and that $g(x) \to 0$ as $x \to 0$, but that $g(f(x))$ does not have a limit as $x \to 0$. Explain this behaviour.

3. Find the values of a and b which make the function

$$f(x) = \begin{cases} x - 1 & \text{if } x \leq -2, \\ ax^2 + c & \text{if } -2 < x < 1, \\ x + 1 & \text{if } x \geq 1, \end{cases}$$

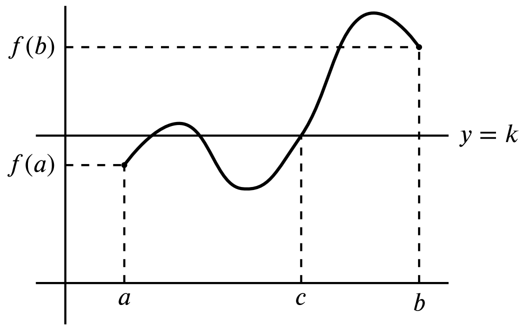
continuous at $x = -2$ and $x = 1$.

4. Prove that if $\lim_{x\to 0^-} f(x)$ exists, then $\lim_{x\to 0^+} f(-x) = \lim_{x\to 0^-} f(x)$.

5. Prove that exp is continuous. You may use the following steps.
(a) The inequality $\exp(x) \geq 1 + x$ is true for all $x \in \mathbb{R}$.
(b) $\lim_{x\to 0^-} \exp(x) = 1$.
(c) $\lim_{x\to 0^+} \exp(x) = 1$. **Hint.** Use tutorial problem 4.

3.5 The Intermediate Value Theorem

The following important theorem on continuous functions tells us that the graph of a continuous function cannot jump from one side of a horizontal line $y = k$ to the other without intersecting the line at least once.



Theorem 3.15 (Intermediate Value Theorem (IVT))**.** *Suppose that f is continuous on the closed interval $[a, b]$ with $f(a) \neq f(b)$. Then for any number k between $f(a)$ and $f(b)$ there exists a number c in the open interval (a, b) such that $f(c) = k$.*

Proof. Let
$$g(x) = f(x) - k, \quad (x \in [a, b]).$$
Then g is continuous, and $g(a)$ and $g(b)$ have opposite signs: $g(a)g(b) < 0$.
Let $[a_0, b_0] = [a, b]$ and use bisection to define intervals $[a_n, b_n]$ as follows: If $[a_n, b_n]$ with $g(a_n)g(b_n) < 0$ has been found, let d be the midpoint of the interval $[a_n, b_n]$. If $g(d) = 0$, the result follows with $c = d$. If $g(d)$ has the same sign as $g(b_n)$, then $g(a_n)$ and $g(d)$ have opposite signs, and putting $a_{n+1} = a_n$, $b_{n+1} = d$, we have $g(a_{n+1})g(b_{n+1}) < 0$. Otherwise, if $g(d)$ has the opposite sign to $g(b_n)$, we put $a_{n+1} = d$, $b_{n+1} = b_n$ and get again $g(a_{n+1})g(b_{n+1}) < 0$.
If this procedure does not stop, we obtain an increasing sequence (a_n) and a decreasing sequence (b_n), both of which converge by Theorem 2.10. We observe that
$$b_n = a_n + \frac{1}{2}(b_{n-1} - a_{n-1}) = a_n + 2^{-n}(b - a).$$
Then
$$c := \lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n + \lim_{n\to\infty} 2^{-n}(b - a) = \lim_{n\to\infty} a_n.$$
Since $a \leq c \leq b$ and g is continuous at c, it follows in view of Theorem 3.14 that
$$g(c)^2 = \lim_{n\to\infty} g(a_n) \lim_{n\to\infty} g(b_n) = \lim_{n\to\infty} g(a_n)g(b_n) \leq 0.$$
Therefore $g(c) = 0$, which gives $f(c) = k$.
Since $f(a) \neq k$ and $f(b) \neq k$, it follows that $c \neq a$ and $c \neq b$, so that $a < c < b$. $\square$

Note. You have seen the definition of interval in first year and you will recall that that the definition required several cases, depending on whether the endpoints belong to the interval or not and whether the interval is bounded (above, below), namely (a, b), $[a, b)$, $(a, b]$, $[a, b]$, $(-\infty, b)$, $(-\infty, b]$, (a, ∞), $[a, \infty)$, $(-\infty, \infty)$ where $a, b \in \mathbb{R}$ and $a < b$. However, intervals can be characterized by one common property. For this we need the following notion: A subset S of $\mathbb{R}$ is called a **singleton** if the set S has exactly one element.

Definition 3.10. 1. A set $S \subset \mathbb{R}$ is called an interval if
(i) $S \neq \emptyset$,
(ii) S is not a singleton,
(iii) if $x, y \in S$, $x < y$, then each $z \in \mathbb{R}$ with $x < z < y$ satisfies $z \in S$.
2. An interval of the form $[a, b]$ with $a < b$ is called a closed bounded interval.

Note. A subset S of $\mathbb{R}$ is an interval if and only if it contains at least two elements and if all real numbers between any two elements in S also belong to S.

Definition 3.11. For a function $f : X \to Y$ and $A \subset X$, the set
$$f(A) = \{y \in Y : \exists\, x \in A \cap \mathrm{dom}(f),\ f(x) = y\} = \{f(x) : x \in A \cap \mathrm{dom}(f)\}$$
is called the **image of A under f**.

Corollary 3.16. *Let I be an interval and let f be a continuous real function on I. Then $f(I)$ is either an interval or a singleton.*

Proof. Since $I \neq \emptyset$, there is $x \in I$ and so $f(x) \in f(I)$. Hence $f(I) \neq \emptyset$. Hence we have to show that if $f(I)$ is not a singleton, then it is an interval, that is, we must show that for any $x, y \in I$ with $f(x) < f(y)$ and any $k \in \mathbb{R}$ with $f(x) < k < f(y)$ there is $c \in I$ such that $f(c) = k$.
Indeed, from $f(x) \neq f(y)$, we have $x \neq y$. If $x < y$, then f is continuous on $[x, y]$, and by the intermediate value theorem, there is $c \in (x, y)$ with $f(c) = k$. A similar argument holds for $x > y$. □

Example 3.5.1. Let $f(x) = x^2$. Then $f((-1, 2)) = [0, 4)$. Notice that $I = (-1, 2)$ is an open interval, while $f(I)$ is not.

Theorem 3.17. *Let f be a real function which is continuous on $[a, b]$, where $a < b$. Then f is bounded on $[a, b]$, i. e., $f([a, b])$ is bounded.*

Proof. Assume that $f([a, b])$ is unbounded. Let d be the midpoint of the interval $[a, b]$. Then at least one of the sets $f([a, d])$, $f([d, b])$ would be unbounded, because otherwise $f([a, b]) = f([a, d]) \cup f([d, b])$ would be bounded. By induction, we find subintervals $[a_n, b_n]$ of $[a, b]$ such that (a_n) is increasing, (b_n) is decreasing, $f([a_n, b_n])$ is unbounded, and

$$b_n = a_n + \frac{1}{2}(b_{n-1} - a_{n-1}) = 2^{-n}(b - a),$$

see the proof of Theorem 3.15, and we infer that both sequences converge with

$$c := \lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n \in [a, b].$$

Since f is continuous at c, there is $\delta > 0$ such that

$$f((c - \delta, c + \delta) \cap [a, b]) \subset (f(c) - 1, f(c) + 1).$$

Since $c = \lim_{n\to\infty} b_n$, there is $K \in \mathbb{N}$ such that $|b_n - c| < \delta$ for $n > K$. Similarly, there is $M \in \mathbb{N}$ such that $|a_n - c| < \delta$ for $n > M$. Let $n > \max\{K, M\}$. Then

$$c - \delta < a_n \leq c \leq b_n < c + \delta,$$

and

$$f([a_n, b_n]) \subset f((c - \delta, c + \delta) \cap [a, b]) \subset (f(c) - 1, f(c) + 1)$$

would give the contradiction that $f([a_n, b_n])$ would have to be bounded as well as unbounded. □

Theorem 3.18. *A continuous function on a closed bounded interval achieves its supremum and infimum.*

Proof. Let $[a, b]$ be a closed bounded interval and f be a continuous function on $[a, b]$. We must show that there are $x_1, x_2 \in [a, b]$ such that $f(x_1) = \inf f([a, b])$ and $f(x_2) = \sup f([a, b])$.
We are going to show the latter; the proof of the first statement is similar.
By Theorem 3.17, $S = f([a, b])$ is bounded. Let $M = \sup S$. By proof of contradiction, assume that $f(x) \neq M$ for all $x \in [a, b]$. Define

$$g(x) = \frac{1}{M - f(x)}, \quad x \in [a, b].$$

By Theorem 3.13, g is continuous on $[a, b]$, and by Theorem 3.17, $g([a, b])$ is bounded. So there is $K \in \mathbb{R}$ such that $0 < g(x) \leq K$ for all $x \in [a, b]$. Hence for all $x \in [a, b]$:

$$\frac{1}{K} \leq \frac{1}{g(x)} = M - f(x) \Rightarrow f(x) \leq M - \frac{1}{K},$$

so that the number $M - \frac{1}{K} < M$ would be an upper bound of $f([a, b])$. This contradicts the fact that $M = \sup f([a, b])$. □

Corollary 3.19. *If f is continuous on $[a, b]$, $a < b$, then either $f([a, b])$ is a singleton or $f([a, b]) = [c, d]$ with $c < d$.*

Proof. From Corollary 3.16 and Theorems 3.17 and 3.18 it follows that $f([a, b])$ is either a singleton or a bounded interval which contains both its infimum and supremum. But such an interval is of the form $[c, d]$ with $c < d$. □

Theorem 3.20. *Let I be an interval and $f : I \to \mathbb{R}$ be a stricly monotonic continuous function. Then $f(I)$ is an interval, and the inverse function $f^{-1} : f(I) \to \mathbb{R}$ is continuous.*

Proof. By Corollary 3.16, $f(I)$ is an interval. Assume that f is strictly increasing. Then also f^{-1} is strictly increasing. Let $b \in f(I)$, i. e., $b = f(a)$ for some $a \in I$. If a is not the left endpoint of I, then b is not the left endpoint of $f(I)$, and for $y \in f(I)$ with $y < b = f(a)$ we have $f^{-1}(y) < f^{-1}(f(a)) = a$. Therefore, f^{-1} is bounded above and increasing on $f(I) \cap (-\infty, b)$, and thus

$$\alpha = \lim_{y \to b^-} f^{-1}(y)$$

exists and $\alpha \le f^{-1}(b) = a$, see Theorem 3.8. Let $x_0 \in I \cap (-\infty, a)$. Then $f(x_0) < f(a) = b$ and hence

$$x_0 = f^{-1}(f(x_0)) \le \alpha \le a.$$

Since x_0 and a belong to the interval I, also $\alpha \in I$. Since f is continuous, it follows from Lemma 3.12 that

$$f(\alpha) = f\left(\lim_{y \to b^-} f^{-1}(y)\right) = \lim_{y \to b^-} f(f^{-1}(y)) = \lim_{y \to b^-} y = b$$

which gives

$$\lim_{y \to b^-} f^{-1}(y) = \alpha = f^{-1}(b).$$

Therefore f^{-1} is continuous from the left. Similarly, one can show that f^{-1} is continuous from the right. Therefore f^{-1} is continuous. The case f strictly decreasing is similar. □

Tutorial 3.5.1. 1. Prove that a subset S of $\mathbb{R}$ is an interval if and only if S has the form (a, b), $[a, b)$, $(a, b]$, $[a, b]$, $(-\infty, b)$, $(-\infty, b]$, (a, ∞), $[a, \infty)$ or $(-\infty, \infty)$ where $a, b \in \mathbb{R}$ and $a < b$.
Hint. Consider $\inf S$ and $\sup S$.

2. Let f be a real function and let $\emptyset \ne A \subset B \subset \mathrm{dom}(f)$ such that for each $x \in A$ there is $\varepsilon > 0$ such that $(x - \varepsilon, x] \subset A$ or $[x, x + \varepsilon) \subset A$ or $(x - \varepsilon, x + \varepsilon) \subset A$, and the same property for B. Show that if f is continuous on B, then f is also continuous on A.

3. **A fixed point theorem**. Let $a < b$ and let f be a continuous function on $[a, b]$ such that $f([a, b]) \subset [a, b]$. Show that there is $x \in [a, b]$ such that $f(x) = x$.

4. Let I be an interval and f be a continuous function on I such that $f(I)$ is unbounded. What can you say about $f(I)$? Find examples which illustrate your answer.

5. Let $f : [0, 1] \to \mathbb{R}$ be a continuous function which only assumes rational values. Show that f is constant.

6. Find a continuous function $f : [-1, 1] \to \mathbb{R}$ which is one-to-one when restricted to rational numbers in $[-1, 1]$ but which is not one-to-one on the whole interval $[-1, 1]$.

Chapter 4

Differentiation

4.1 Revision – Self Study

[Note to lecturers: None of these results have to be presented formally. But, time permitting, some of these definitions and results may be discussed in class; foremost, of course, the definition of the derivative itself.]

In this section we recall definitions and results which have been stated and proved in first year calculus. These results must be known but will not be examined directly.

In this section let $A \subset \mathbb{R}$ such that for each $a \in A$ there is $\varepsilon > 0$ such that $(a-\varepsilon, a] \subset A$, $[a, a+\varepsilon) \subset A$ or $(a-\varepsilon, a+\varepsilon) \subset A$.

Definition 4.1. Let $f : A \to \mathbb{R}$ and $a \in A$.
1. f is called differentiable at a if the limit

$$f'(x) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}$$

exists. The number $f'(x)$ is called the derivative of f.
2. f is called differentiable on A if f is differentiable at each $a \in A$.

Note. We also use the notations $\dfrac{df}{dx}$ or $\dfrac{d}{dx}f$ for f'.

Theorem 4.1. *Let $f : A \to \mathbb{R}$ and $a \in A$. If f is differentiable at a, then f is continuous at a.*

Proof. From

$$f(x) = f(a) + \frac{f(x)-f(a)}{x-a}(x-a)$$

it follows that

$$\lim_{x\to a} f(x) = f(a) + \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \lim_{x\to a}(x-a) = f(a) + [f'(a)](0) = f(a).$$ □

Theorem 4.2 (Rules for the derivative)**.** *Let $f, g : A \to \mathbb{R}$, $a \in A$, f and g differentiable at a, and $c \in \mathbb{R}$. Then*
1. *Linearity of the derivative: $f+g$ and cf are differentiable at a, and*

$$(f+g)'(a) = f'(a) + g'(a) \quad \text{and} \quad (cf)'(a) = cf'(a).$$

2. *Product Rule: fg is differentiable at a, and*

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

3. *Quotient Rule: If $\dfrac{f}{g}$ is defined at a, i. e., $g(a) \neq 0$, then $\dfrac{f}{g}$ is differentiable at a, and*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.$$

Proof. 1. From first principles and rules of limits, we have

$$\begin{aligned}(f+g)'(a) &= \lim_{x\to a}\frac{(f+g)(x)-(f+g)(a)}{x-a}\\ &= \lim_{x\to a}\frac{[f(x)+g(x)]-[f(a)+g(a)]}{x-a}\\ &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}+\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\\ &= f'(a)+g'(a)\end{aligned}$$

and

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

2. Observing Theorem 4.1 we have

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

3. We first consider the case $f=1$:

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

To prove the statement for general f, we use the product rule:

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \left(f\frac{1}{g}\right)'(a)=f'(a)\left(\frac{1}{g}\right)(a)+f(a)\left(\frac{1}{g}\right)'(a)=\frac{f'(a)}{g(a)}+f(a)\left(-\frac{g'(a)}{[g(a)]^2}\right)\\ &= \frac{f'(a)g(a)-f(a)g'(a)}{[g(a)]^2}.\end{aligned}$$

□

Example 4.1.1. 1. If $c\in\mathbb{R}$, then $\frac{dc}{dx}=0$.
2. For $n\in\mathbb{N}^*$ we have $\dfrac{d}{dx}x^n=nx^{n-1}$.

Proof. 1. follows immediately from the definition.
2. We present an alternate proof to that given in first year: proof by induction. Again, for $n=1$ the statement is straightforward:

$$\frac{d}{dx}x=\lim_{y\to x}\frac{y-x}{y-x}=1=(1)x^{1-1}.$$

Now assume the statement is true for $n \in \mathbb{N}^*$. Then, making use of the Product Rule,

$$\frac{d}{dx}x^{n+1} = \frac{d}{dx}((x^n)(x)) = \left(\frac{d}{dx}x^n\right)(x) + x^n\frac{d}{dx}x = nx^{n-1}(x) + x^n(1) = (n+1)x^n = (n+1)x^{(n+1)-1}. \qquad \square$$

As in first year calculus, we now obtain the derivative $\frac{d}{dx}x^n = nx^{n-1}$ for negative integers n with the aid of the quotient rule. Also, the derivatives of the trigonometric functions sin, cos, tan have been derived in first year calculus and should be known.

The proof of the differentiability of e^x in first year calculus was incomplete; it will be given in the next section. Also, the differentiability of the inverse trigonometric functions, $\ln x$ and x^r for $r \in \mathbb{R} \setminus \mathbb{N}$ will be postponed to the next section.

Theorem 4.3 (Fermat's Theorem)**.** *Let I be an interval, $f : I \to \mathbb{R}$, and let c be in the interior of I. If f has a local maximum or local minimum at c and f is differentiable at c, then $f'(c) = 0$.*

Proof. Since c is an interior point of I, there is $\varepsilon_0 > 0$ such that $(c - \varepsilon_0, c + \varepsilon_0) \subset I$. Assume that f has a local maximum at c. Then there is $\varepsilon \in (0, \varepsilon_0)$ such that $f(x) \le f(c)$ for all $x \in (c - \varepsilon, c + \varepsilon)$. Therefore

$$\frac{f(x) - f(c)}{x - c} \begin{cases} \ge 0 & \text{for } x \in (c - \varepsilon, c), \\ \le 0 & \text{for } x \in (c, c + \varepsilon). \end{cases}$$

Hence, since f is differentiable at c,

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

and

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

Therefore $0 \le f'(c) \le 0$, which proves $f'(c) = 0$.
The case of a local minimum at c can be dealt with in a similar way, or we may use the fact that then $-f$ has a local maximum at c and therefore $-f'(c) = 0$ by the first part of the proof. $\square$

Theorem 4.4 (Rolle's Theorem)**.** *Let $a < b$ be real numbers and $f : [a, b] \to \mathbb{R}$ be a function having the following properties:*
1. *f is continuous on the closed interval $[a, b]$,*
2. *f is differentiable on the open interval (a, b),*
3. *$f(a) = f(b)$.*

Then there is $c \in (a, b)$ such that $f'(c) = 0$.

Proof. If f is constant, then $f' = 0$, and the statement is true for any $c \in (a, b)$. If f is not constant, then there is $x_0 \in [a, b]$ such that $f(x_0) \ne f(a)$. We may assume $f(x_0) > f(a)$. Otherwise, $f(x_0) < f(a)$, and we can consider $-f$. Since f is continuous on $[a, b]$ by property 1, f has a maximum on $[a, b]$, see Corollary 3.19. That is, there is some $c \in [a, b]$ such that $f(c) \ge f(x)$ for all $x \in [a, b]$. In particular, $f(c) \ge f(x_0) > f(a)$. Since $f(a) = f(b)$, we have $c \ne a$ and $c \ne b$, and therefore $c \in (a, b)$. Hence $f'(c) = 0$ by Fermat's Theorem. $\square$

Theorem 4.5 (First Mean Value Theorem)**.** *Let $a < b$ be real numbers and $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b).*
Then there is $c \in (a, b)$ such that

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

Proof. The function

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

is continuous on $[a, b]$, differentiable on (a, b), and $g(a) = f(a) = g(b)$. Hence g satisfies the assumptions of Rolle's theorem. Therefore, there is $c \in (a, b)$ such that $g'(c) = 0$. But

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

and substituting c completes the proof. □

Tutorial 4.1.1. Let $a < b$ be real numbers, $f, g : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b).
1. Show that g is injective on $[a, b]$ if $g'(x) \neq 0$ for all $x \in (a, b)$.
2. Prove the Second Mean Value Theorem: If $g'(x) \neq 0$ for all $c \in (a, b)$, then there is $c \in (a, b)$ such that

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

Hint. Consider the function

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - (g(a)).$$

3. Show that the statement of the Second Mean Value Theorem remains correct if one replaces the condition that $g'(x) \neq 0$ for all $x \in (a, b)$ with the weaker condition that $g(a) \neq g(b)$ and that there is no $x \in (a, b)$ with $g'(x) = f'(x) = 0$.
4. Prove the following one-sided version of l'Hôpital's Rule: If $f(a) = g(a) = 0$, $g'(x) \neq 0$ for x near a and $\lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$ exists, then

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

***Tutorial 4.1.2.** Let $-\infty \leq a < b \leq \infty$ and let $f, g : (a, b) \to \mathbb{R}$ be differentiable on (a, b) such that $g'(x) \neq 0$ for all $x \in (a, b)$. Prove the following one-sided versions of l'Hôpital's Rule.
1. If $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = 0$ and $\lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$ exists as a proper or improper limit, then

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

2. If $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = \infty$ and $\lim\limits_{x\to a^+} \frac{f'(x)}{g'(x)}$ exists as a proper or improper limit, then

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

4.2 The Chain Rule and Inverse Functions

Proposition 4.6. *Let I be an interval, $f : I \to \mathbb{R}$ and $a \in I$. The following are equivalent:*
(i) *f is differentiable at a.*
(ii) *There is $\gamma \in \mathbb{R}$ such that the function*

$$f^*(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} & \text{if } x \in I \setminus \{a\}, \\ \gamma & \text{if } x = a, \end{cases}$$

is continuous at a.
(iii) *There is a function $\tilde{f} : I \to \mathbb{R}$ such that $\tilde{f}$ is continuous at a and*

$$f(x) = f(a) + \tilde{f}(x)(x - a), \quad x \in I.$$

(iv) *There are $\gamma \in \mathbb{R}$ and a function $\varepsilon : I \to \mathbb{R}$ such that $\varepsilon(a) = 0$, ε is continuous at a, and*

$$f(x) = f(a) + \gamma(x - a) + \varepsilon(x)(x - a), \quad x \in I.$$

If (ii) *or* (iv) *hold, then the number γ is unique and $\gamma = f'(a)$. If* (iii) *holds, then $\tilde{f}(a) = f'(a)$.*

Proof. (i)⇒(ii): If f is differentiable at a, then

$$\lim_{x\to a} f^*(x) = f'(x),$$

so that f^* is continuous at a for $\gamma = f'(a)$.
(ii)⇒(iii): We simply put $\tilde{f} = f^*$. In particular, $\tilde{f}(a) = \gamma$.
(iii)⇒(iv): We have

$$f(x) = f(a) + \tilde{f}(a)(x-a) + (\tilde{f}(x) - \tilde{f}(a))(x-a).$$

Then (iv) holds with $\varepsilon(x) = \tilde{f}(x) - \tilde{f}(a)$ since $\varepsilon(a) = 0$ and the continuity of $\tilde{f}$ at a implies the continuity of ε at a.
(iv)⇒(i): For $x \in I \setminus \{a\}$ we have

$$\frac{f(x) - f(a)}{x - a} = \gamma + \varepsilon(x) \to \gamma \text{ as } x \to a.$$

Hence f is differentiable at a, and $\gamma = f'(a)$. □

Theorem 4.7 (Chain Rule)**.** *Let I and J be intervals, $g : J \to \mathbb{R}$ and $f : I \to \mathbb{R}$ with $f(I) \subset J$, and let $a \in I$. Assume that f is differentiable at a and that g is differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and*

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

Proof. Using the property (iii) of Proposition 4.6 we obtain a function $\tilde{f}$ on I which is continuous at a with

$$f(x) = f(a) + \tilde{f}(x)(x-a)$$

and a function $\tilde{g}$ on J which is continuous at $f(a)$ with

$$g(y) = g(f(a)) + \tilde{g}(y)(y - f(a)).$$

Then we can write

$$\begin{aligned}(g \circ f)(x) = g(f(x)) &= g(f(a)) + \tilde{g}(f(x))(f(x) - f(a)) \\ &= (g \circ f)(a) + \tilde{g}(f(x))\tilde{f}(x)(x-a).\end{aligned}$$

Since f and $\tilde{f}$ are continuous at a and since g is continuous at $f(a)$, $(\tilde{g} \circ f)\tilde{f}$ is continuous at a, see Theorem 3.11, and by Proposition 4.6, $(g \circ f)'(a) = \tilde{g}(f(a))\tilde{f}(a) = g'(f(a))f'(a)$. □

Theorem 4.8. *Let I be an interval, let $f : I \to \mathbb{R}$ be continuous and strictly increasing or decreasing, and $b \in f(I)$. Assume that f is differentiable at $a = f^{-1}(b)$ with $f'(a) \neq 0$. Then f^{-1} is differentiable at $b = f(a)$ and*

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))} \quad \textit{or, equivalently,} \quad (f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Proof. By Proposition 4.6 there is a function $\tilde{f}$ on I which is continuous at a such that

$$f(x) - f(a) = \tilde{f}(x)(x-a), \tag{$*$}$$

and $\tilde{f}(a) = f'(a)$. By Theorem 3.20, f^{-1} is continuous, and therefore $\tilde{f} \circ f^{-1}$ is continuous at $f(a)$. Since $f'(a) \neq 0$, $\tilde{f}(a) \neq 0$, and since f is injective, $(*)$ gives that $\tilde{f}(x) \neq 0$ for $x \in I \setminus \{a\}$. Therefore the function $\dfrac{1}{\tilde{f} \circ f^{-1}}$ is well-defined on $f(I)$ and continuous at $f(a)$. Setting $x = f^{-1}(y)$ for $y \in f(I)$ we get

$$y - f(a) = \tilde{f}(f^{-1}(y))(f^{-1}(y) - a),$$

or, equivalently,

$$f^{-1}(y) - f^{-1}(f(a)) = \frac{1}{\tilde{f}(f^{-1}(y))}(y - f(a)).$$

By Proposition 4.6, f^{-1} is differentiable at $f(a)$, and

$$(f^{-1})'(f(a)) = \frac{1}{\tilde{f}(f^{-1}(f(a)))} = \frac{1}{f'(a)}.$$ □

This result now allows us to find the derivatives of arcsin and arctan.

Example 4.2.1. Show that $\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ is continuous on $[-1,1]$ and differentiable on $(-1,1)$, and find $\dfrac{d}{dx}\arcsin$.

Solution. $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$ is strictly increasing and continuous. By Theorem 3.20, arcsin is continuous on $[-1,1]$. Since

$$\frac{d}{dx}\sin x = \cos x \neq 0 \quad \text{for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

that is, if $\sin x \in (-1,1)$, it follows by Theorem 4.8 that arcsin is differentiable on $(-1,1)$ and

$$\frac{d}{dx}\arcsin x = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}$$

since $\cos t > 0$ for $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

Finally, in this section we shall show that $e^x = \exp(x)$ is differentiable and find its derivative. This will then allow us to find the derivative of $\ln x$ as an inverse function.

Theorem 4.9 (Derivative of e^x)**.**
1. $e^x \geq 1 + x$ *for* $x \in \mathbb{R}$ *and* $e^x \leq \dfrac{1}{1-x}$ *for* $x < 1$.
2. e^x *is differentiable and* $\dfrac{d}{dx}e^x = e^x$.

Proof. 1. Let $x \in \mathbb{R}$ and $n \in \mathbb{N}$ with $n > |x|$. Then $\frac{x}{n} > -1$, and with the aid of Bernoulli's inequality we calculate

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

Thus we have shown that $e^x \geq 1 + x$ for all $x \in \mathbb{R}$. Replacing x with $-x$ gives $e^{-x} \geq 1 - x$. Taking inverses for $x < 1$, i.e., $1 - x > 0$, leads to $e^x = \dfrac{1}{e^{-x}} \leq \dfrac{1}{1-x}$ for $x < 1$.
2. The above estimates give for $h < 1$ that

$$h \leq e^h - 1 \leq \frac{1}{1-h} - 1 = \frac{h}{1-h} \Rightarrow \begin{cases} 1 \leq \dfrac{e^h-1}{h} \leq \dfrac{1}{1-h} & \text{if } h>0, \\ \dfrac{1}{1-h} \leq \dfrac{e^h-1}{h} \leq 1 & \text{if } h<0. \end{cases}$$

Application of the Sandwich Theorem leads to

$$\lim_{h\to0}\frac{e^h-1}{h} = 1.$$

Therefore

$$\frac{d}{dx}e^x = \lim_{h\to0}\frac{e^{x+h}-e^x}{h} = \lim_{h\to0}\frac{e^xe^h-e^x}{h} = e^x\lim_{h\to0}\frac{e^h-1}{h} = e^x. \qquad \square$$

Tutorial 4.2.1. 1. Find the derivatives of
(a) $\arctan x$, (b) $\arccos x$, (c) $\ln x$ and $\ln|x|$,
(d) $\ln|f(x)|$, where $f : I \to \mathbb{R} \setminus \{0\}$ is differentiable on I.

2. Let $x > 0$ and let $r = \dfrac{p}{q}$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}^*$. Define $x^r = (x^p)^{\frac{1}{q}}$.
Show that $x^r = \exp(r\ln x)$.

3. In view of tutorial problem 2 above, we define $x^r = \exp(r\ln x)$ for all $x > 0$ and $r \in \mathbb{R}$.
(a) Show that $\lim\limits_{x\to-\infty}\exp(x) = 0$.
(b) Show that $\lim\limits_{x\to\infty}\dfrac{1}{x^r} = 0$ for all $r > 0$.

Chapter 5

Series

5.1 Definitions and Basic Examples

Given a sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers, the symbol

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

is called a **series** (of real numbers).

Definition 5.1. Let $\sum_{n=1}^{\infty} a_n$ be a series.

1. The number $s_n = \sum_{i=1}^{n} a_i$ is called the ***n*-th partial sum of the series**.

2. The series $\sum_{n=1}^{\infty} a_n$ is said to **converge** if the sequence (s_n) converges.
In this case, the number $s = \lim_{n\to\infty} s_n$ is called the **sum of the series** and we write

$$\sum_{n=1}^{\infty} a_n = s$$

A series which does not converge is said to **diverge** or be **divergent**.

Note. 1. Observe that $\sum_{n=1}^{\infty} a_n$ denotes a series (convergent or divergent) as well as its sum if it converges.
2. A series does not have to start at $n = 1$. The change of notation is obvious for other starting indices.

Example 5.1.1 (See Calculus I)**.** [Note to lecturers: Students have seen this in Calculus I, so one does not need to write down much.] Consider the **geometric series**

$$\sum_{n=0}^{\infty} ar^n, \quad a \neq 0,\ r \in \mathbb{R}.$$

Recall the partial sums for $r \neq 1$:

$$\begin{aligned} s_n &= a + ar + ar^2 + \cdots + ar^n, \\ rs_n &= \phantom{a + {}} ar + ar^2 + \cdots + ar^n + ar^{n+1}. \end{aligned}$$

Subtracting these equations gives

$$s_n - rs_n = a - ar^{n+1},$$

and so

$$s_n = \frac{a(1-r^{n+1})}{1-r}.$$

By Theorem 2.5, (s_n) and thus the geometric series converges if $|r| < 1$, with

$$\sum_{n=0}^{\infty} ar^n = \lim_{n\to\infty} s_n = \frac{a}{1-r},$$

and diverges if $|r| > 1$ or $r = -1$. Finally, for $r = 1$, $s_n = a(n+1)$, and thus $s_n \to \infty$ as $n \to \infty$.

Thus we have shown

Theorem 5.1. *The geometric series*

$$\sum_{n=0}^{\infty} ar^n, \quad a \neq 0,$$

is convergent if $|r| < 1$, *with sum*

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}.$$

If $|r| \geq 1$, *the geometric series diverges.*

Example 5.1.2. Is the series $\sum_{n=1}^{\infty} 2^{2n}3^{1-n}$ convergent or divergent?

Solution. The n-th term of the series is obviously of the form ar^n:

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

So $r = \frac{4}{3} > 1$, and the series diverges.

Example 5.1.3. Write the number $5.4\overline{417} = 5.4417417417\ldots$ as a ratio of integers.

Solution. Write

$$5.4\overline{417} = 5.4 + \frac{417}{10^4} + \frac{417}{10^7} + \frac{417}{10^{10}} + \ldots.$$

Starting with the second term we have a geometric series with $a = \frac{417}{10^{-4}}$ and $r = 10^{-3}$. Hence the series converges, and

$$\begin{aligned} 5.4\overline{417} &= 5.4 + \frac{\frac{417}{10^{-4}}}{1-10^{-3}} = 5.4 + \frac{417}{10^4 - 10} \\ &= \frac{54}{10} + \frac{417}{9990} = \frac{539877}{9990} = \frac{179959}{3330}. \end{aligned}$$

The following theorem has been proved in Calculus I.

Theorem 5.2. *If the series* $\sum_{n=1}^{\infty} a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$.

Proof. Let

$$s_n = a_1 + a_2 + \ldots a_n.$$

Then

$$a_n = s_n - s_{n-1}.$$

Since $\sum_{n=1}^{\infty} a_n$ converges,

$$\lim_{n\to\infty} s_n = s$$

exists. Since also $n-1 \to \infty$ as $n \to \infty$,

$$\lim_{n\to\infty} s_{n-1} = s.$$

Hence

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s - s = 0.$$

The contrapositive statement to Theorem 5.2 is very useful:

Theorem 5.3 (Test for Divergence). *If $\lim_{n\to\infty} a_n$ does not exist or if $\lim_{n\to\infty} a_n \neq 0$, then the series $\sum_{n=1}^{\infty} a_n$ diverges.*

Note. From $\lim_{n\to\infty} a_n = 0$ nothing can be concluded about the convergence of the series $\sum_{n=1}^{\infty} a_n$.

From Theorem 2.3 we immediately infer

Theorem 5.4 (Sum Laws). *Let $c \in \mathbb{R}$ and suppose that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ both converge. Then also $\sum_{n=1}^{\infty}(ca_n)$, $\sum_{n=1}^{\infty}(a_n + b_n)$ and $\sum_{n=1}^{\infty}(a_n - b_n)$ converge, and*

1. $\sum_{n=1}^{\infty}(ca_n) = c\sum_{n=1}^{\infty} a_n$,
2. $\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n$,
3. $\sum_{n=1}^{\infty}(a_n - b_n) = \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} b_n$,

Example 5.1.4. Find the sum of the series $\sum_{n=1}^{\infty}\left(\frac{5}{n(n+1)} + \frac{1}{5^n}\right)$.

Solution. Recall from Calculus I that

$$\sum_{n=1}^{m} \frac{1}{n(n+1)} = \sum_{n=1}^{m}\left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{m+1}.$$

Hence

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

And $\sum_{n=1}^{\infty} \frac{1}{5^n}$ is a geometric series with $a = \frac{1}{5}$ and $r = \frac{1}{5}$. Hence

$$\sum_{n=1}^{\infty} \frac{1}{5^n} = \frac{1}{5}\frac{1}{1-\frac{1}{5}} = \frac{1}{4}.$$

Therefore

$$\sum_{n=1}^{\infty}\left(\frac{5}{n(n+1)} + \frac{1}{5^n}\right) = 5\sum_{n=1}^{\infty}\frac{1}{n(n+1)} + \sum_{n=1}^{\infty}\frac{1}{5^n} = 5 + \frac{1}{4} = \frac{21}{4}.$$

Note. For convergence it does not matter at which index the series starts. But it matters for the sum.

Theorem 5.5. *The series* $\sum_{n=1}^{\infty} a_n$ *converges if and only if for each* $\varepsilon > 0$ *there is* $K \in \mathbb{N}$ *such that for all* $m \geq k \geq K$, $\left|\sum_{n=k}^{m} a_n\right| < \varepsilon$.

Proof. Let $s_k = \sum_{n=1}^{k} a_n$. Then

$$\begin{aligned}\sum_{n=1}^{\infty} a_n \text{ converges } &\Leftrightarrow (s_n)_{n=1}^{\infty} \text{ converges} \qquad \text{(by definition)}\\ &\Leftrightarrow \forall \varepsilon > 0\ \exists K \in \mathbb{N}\ \forall m \geq k \geq K,\ |s_m - s_{k-1}| < \varepsilon \qquad \text{(by Theorem 2.12)}\\ &\Leftrightarrow \forall \varepsilon > 0\ \exists K \in \mathbb{N}\ \forall m \geq k \geq K,\ \left|\sum_{n=k}^{m} a_n\right| < \varepsilon.\end{aligned}$$

□

Tutorial 5.1.1. 1. Test each of the following series for convergence or divergence:
(a) $\sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^n$, (b) $\sum_{n=1}^{\infty} n \sin\left(\frac{1}{n}\right)$, (c) $\sum_{n=1}^{\infty} \frac{n^2 - 1}{n - 50n^2}$.
2. Which of the following is valid? Justify your conclusions.
(a) If $a_n \to 0$ as $n \to \infty$, then $\sum_{n=1}^{\infty} a_n$ is convergent.
(b) If $a_n \not\to 0$ as $n \to \infty$, then $\sum_{n=1}^{\infty} a_n$ is divergent.
(c) If $\sum_{n=1}^{\infty} a_n$ is divergent, then $a_n \not\to 0$ as $n \to \infty$.

5.2 Convergence Tests for Series

Lemma 5.6. *Let* $a_n \geq 0$ *for all* $n \in \mathbb{N}^*$ *and let* $s_k = \sum_{n=1}^{k} a_n$. *Then* $\sum_{n=1}^{\infty} a_n$ *converges if and only if* (s_n) *is bounded.*

Proof. Since $s_{n+1} = s_n + a_{n+1} \geq s_n$, it follows that (s_n) is an increasing sequence. By Theorem 2.10, this sequence and hence the series converges if and only if the sequence (s_n) is bounded. □

Theorem 5.7 (Comparison Test). *Let* $\sum_{n=1}^{\infty} a_n$ *and* $\sum_{n=1}^{\infty} b_n$ *be series with nonnegative terms and assume that* $a_n \leq b_n$ *for all* $n \in \mathbb{N}^*$.
(i) *If* $\sum_{n=1}^{\infty} b_n$ *converges, then also* $\sum_{n=1}^{\infty} a_n$ *converges.*
(ii) *If* $\sum_{n=1}^{\infty} a_n$ *diverges, then also* $\sum_{n=1}^{\infty} b_n$ *diverges.*

Proof. Let $s_k = \sum_{n=1}^{k} a_n$ and $t_k = \sum_{n=1}^{k} b_n$. Then both (s_k) and (t_k) are increasing sequences with $s_k \leq t_k$. Hence, if $\sum_{n=1}^{\infty} b_n$ converges, then (t_k) is bounded, say $t_k \leq M$ for all $k \in \mathbb{N}$, and $s_k \leq t_k \leq M$ for all $k \in \mathbb{N}$. Hence (s_k) is a bounded sequence and thus converges by Lemma 5.6.
(ii) is the contrapositive of (i). □

Example 5.2.1. Test the series $\sum_{n=1}^{\infty} \frac{\sin^2 n + 10}{n + 2^n}$ for convergence.

Solution. Putting

$$a_n = \frac{\sin^2 n + 10}{n + 2^n}$$

it follows that $0 < a_n < 11 \left(\frac{1}{2}\right)^n =: b_n$. By Theorem 5.1, the series with general term b_n converges. Hence $\sum_{n=1}^{\infty} \frac{\sin^2 n + 10}{n + 2^n}$ converges by the comparison test.

Definition 5.2. 1. A series $\sum_{n=1}^{\infty} a_n$ is called **absolutely convergent** if the series of its absolute values $\sum_{n=1}^{\infty} |a_n|$ converges.

2. A series $\sum_{n=1}^{\infty} a_n$ is called **conditionally convergent** if it is convergent but not absolutely convergent.

Theorem 5.8. *Every absolutely convergent series is convergent.*

Proof. Let $\sum_{n=1}^{\infty} a_n$ be absolutely convergent. Let $\varepsilon > 0$. Then, by Theorem 5.5, there is $K \in \mathbb{N}$ such that for all $m \geq k \geq K$, $\sum_{n=k}^{m} |a_n| < \varepsilon$. Because of

$$\left| \sum_{n=k}^{m} a_n \right| \leq \sum_{n=k}^{m} |a_n|$$

it follows that $\left| \sum_{n=k}^{m} a_n \right| < \varepsilon$ for these k, m, and therefore $\sum_{n=1}^{\infty} a_n$ converges by Theorem 5.5.

Alternative proof. By assumption and Theorem 5.4, both $\sum_{n=1}^{\infty} |a_n|$ and $\sum_{n=1}^{\infty} 2|a_n|$ converge. From

$$0 \leq a_n + |a_n| \leq 2|a_n|$$

and the comparison test, it follows that also $\sum_{n=1}^{\infty}(a_n + |a_n|)$ converges. Again by Theorem 5.4 it follows that

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

converges. □

Recall from Calculus I that there are convergent series which are not absolutely convergent.

Definition 5.3. An **alternating series** is a series of the form

$$\sum_{n=1}^{\infty}(-1)^n b_n \qquad \text{or} \qquad \sum_{n=1}^{\infty}(-1)^{n-1} b_n \qquad \text{with } b_n \geq 0.$$

Theorem 5.9 (Alternating series test). *If the alternating series* $\sum_{n=1}^{\infty}(-1)^n b_n$ *or* $\sum_{n=1}^{\infty}(-1)^{n-1} b_n$ *satisfies*

(i) $b_n \geq b_{n+1}$ *for all n,*
(ii) $\lim_{n\to\infty} b_n = 0$,
then the series converges.

Proof. For $k \in \mathbb{N}^*$ and $m \in \mathbb{N}$ we have

$$(-1)^k \sum_{n=k}^{k+2m} (-1)^n b_n = (b_k - b_{k+1}) + (b_{k+2} - b_{k+3}) + \cdots + (b_{k+2m-2} - b_{k+2m-1}) + b_{k+2m}$$
$$= b_k - (b_{k+1} - b_{k+2}) - \cdots - (b_{k+2m-1} - b_{k+2m}).$$

Hence

$$0 \le b_{k+2m} \le (-1)^k \sum_{n=k}^{k+2m} (-1)^n b_n \le b_k.$$

Similarly,

$$0 \le (-1)^k \sum_{n=k}^{k+2m+1} (-1)^n b_n \le b_k - b_{k+2m+1} \le b_k.$$

Now let $\varepsilon > 0$. Since $b_k \to 0$ as $k \to \infty$, there is $K \in \mathbb{N}$ such that $b_K < \varepsilon$. Hence for all $l \ge k \ge K$:

$$\left| \sum_{n=k}^{l} (-1)^n b_n \right| \le b_k \le b_K < \varepsilon.$$

Hence the alternating series converges. □

Note. $\lim_{n\to\infty} b_n = 0$ is necessary by Theorem 5.3 since $\lim_{n\to\infty} (-1)^n b_n = 0 \Leftrightarrow \lim_{n\to\infty} b_n = 0 \Leftrightarrow \lim_{n\to\infty} (-1)^{n-1} b_n = 0$.

Example 5.2.2. The alternating series $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{3n}{4n-1}$ does not converge since

$$\lim_{n\to\infty} \frac{3n}{4n-1} = \frac{3}{4} \neq 0$$

and thus (ii) is not satisfied, which is necessary for convergence. See the note following the statement of the Alternating Series Test.

Worked Example 5.2.3. Find whether the series $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n^2}{n^3+1}$ is convergent.

Solution. Clearly, we have an alternating series with

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{n^2}{n^3+1} = 0.$$

To show that $b_n \ge b_{n+1}$ we have at least the following 2 choices:
1. Show the inequality directly by cross-multiplication and simplification.
2. Show that $f(x) = \dfrac{x^2}{x^3+1}$ has a negative derivative for sufficiently large x.

Alternatively, write

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

Then

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n^2}{n^3+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} + \sum_{n=1}^{\infty} \frac{(-1)^n}{n(n^3+1)}$$

is the sum of two alternating series which converge by the alternating series test; hence the series converges.

Theorem 5.10 (Ratio Test). (i) *If* $\limsup_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$, *then the series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*

(ii) *If* $\liminf_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = l > 1$, *then the series* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof. Note that the ratio test assumes $a_n \neq 0$ for all $n \in \mathbb{N}$.
(i) Let $\varepsilon > 0$ such that $L + \varepsilon < 1$.
Then there is $K \in \mathbb{N}$ such that $\left|\frac{a_{n+1}}{a_n}\right| < L + \varepsilon$ for all $n \geq K$. Hence for $m > K$:

$$|a_m| = |a_K| \left|\frac{a_{K+1}}{a_K}\right| \cdot \dots \cdot \left|\frac{a_m}{a_{m-1}}\right| < |a_K|(L+\varepsilon)^{m-K}. \tag{*}$$

Since $\sum_{m=K}^{\infty} |a_K|(L+\varepsilon)^{m-K}$ is a convergent geometric series, it follows from (*) and the Comparison Test, Theorem 5.7, that $\sum_{m=K}^{\infty} a_m$ converges absolutely. Hence also $\sum_{n=1}^{\infty} a_n$ converges absolutely.
(ii) Let $l' \in (1, l)$.
Then there is $K \in \mathbb{N}$ such that $\left|\frac{a_{n+1}}{a_n}\right| > l'$ for all $n \geq K$. Hence for $m > K$:

$$|a_m| = |a_K| \left|\frac{a_{K+1}}{a_K}\right| \cdot \dots \cdot \left|\frac{a_m}{a_{m-1}}\right| > |a_K|,$$

so that $a_n \not\to 0$ as $n \to \infty$. Hence $\sum_{n=1}^{\infty} a_n$ diverges by the Test for Divergence, Theorem 5.3. □

Theorem 5.11 (Root Test). (i) *If* $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = L < 1$, *then the series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*
(ii) *If* $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = L > 1$, *then the series* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof. (i) Let $\varepsilon > 0$ such that $L + \varepsilon < 1$.
Then there is $K \in \mathbb{N}$ such that $\sqrt[n]{|a_n|} < L+\varepsilon$ for all $n \geq K$. Hence $|a_n| < (L+\varepsilon)^n$ for $n \geq K$: Since $\sum_{n=K}^{\infty}(L+\varepsilon)^n$ is a convergent geometric series, it follows from the comparison test, Theorem 5.7, that $\sum_{n=K}^{\infty} a_n$ converges absolutely.
Hence also $\sum_{n=1}^{\infty} a_n$ converges absolutely.
(ii) Let $L' \in (1, L)$.
Then for each $K \in \mathbb{N}$ there is $m \geq K$ such that $\sqrt[m]{|a_m|} > L'$. Hence $|a_m| > (L')^m > 1$ for this m, and we conclude $a_n \not\to 0$ as $n \to \infty$.
Indeed, if $a_n \to 0$ as $n \to \infty$, for $\varepsilon = 1$ there would be $K \in \mathbb{N}$ such that $|a_n| < 1$ for all $n > K$.
Hence $\sum_{n=1}^{\infty} a_n$ diverges by the Test for Divergence, Theorem 5.3. □

Tutorial 5.2.1. 1. Test $\sum_{n=1}^{\infty} \left\{ \left(\frac{1}{2}\right)^n + 5\left(\frac{3}{4}\right)^n \right\} n \sin\left(\frac{1}{n}\right)$ for convergence.
2. Prove that the sequence $(a_n)_{n=1}^{\infty}$ converges if and only if
(i) $(a_{2n})_{n=1}^{\infty}$ converges,
(ii) $(a_{2n-1})_{n=1}^{\infty}$ converges,
(iii) $(a_n - a_{n-1}) \to 0$ as $n \to \infty$.
3. Use Tutorial 2 to prove the Alternating Series Test.
Hint. Show that $(s_{2n})_{n=1}^{\infty}$ and $(s_{2n-1})_{n=1}^{\infty}$ are monotonic sequences.

4. Use the alternating series test, ratio test or root test to test for convergence:

(a) $\displaystyle\sum_{n=1}^{\infty} \left(\frac{(-1)^n n}{2n+1}\right)^{2n}$, (b) $\displaystyle\sum_{n=1}^{\infty} \frac{n!2^n}{(2n)!}$,

(c) $\displaystyle\sum_{n=1}^{\infty} (-1)^n \left(e - \left(1+\frac{1}{n}\right)^n\right)$, (d) $\displaystyle\sum_{n=1}^{\infty} \frac{n^n}{(n!)^2}$,

(e) $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n}$.

5.3 Power Series

Definition 5.4. Let a and c_n, $n \in \mathbb{N}$, be real numbers. A series

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

is called a **power series in** $(x - a)$ or a **power series centred at** a or a **power series about** a.

Note that $(x - a)^0 = 1$. For $x = a$, all terms from the second onwards are 0, so the series converges to c_0 for $x = a$. Each power series defines a function whose domain is those $x \in \mathbb{R}$ for which the series converges.

Notation. With each power series $\displaystyle\sum_{n=0}^{\infty} c_n(x-a)^n$ we associate a number R or ∞, called the **radius of convergence** of the series, which is defined as follows:

(i) $R = 0$ if $\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} = \infty$,

(ii) $R = \dfrac{1}{\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$ if $0 < \displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} \in \mathbb{R}$,

(iii) $R = \infty$ if $\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} = 0$.

Theorem 5.12. *There are three alternatives for the domain of a power series* $\displaystyle\sum_{n=0}^{\infty} c_n(x-a)^n$*:*
(i) *If $R = 0$, then the series converges only for $x = a$.*
(ii) *If $R = \infty$, then the series converges absolutely for all $x \in \mathbb{R}$.*
(iii) *If $0 < R \in \mathbb{R}$, then the series converges absolutely if $|x - a| < R$ and diverges if $|x - a| > R$.*

Note that (iii) says that the series converges for x in the interval $(a - R, a + R)$ and diverges outside $[a - R, a + R]$. For $x = a - R$ and $x = a + R$ anything can happen, see Calculus I. In any case, the domain of the series is an interval, called the **interval of convergence**.

Proof. (i): We note that if $\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} = \infty$, then for $x \neq a$ also

$$\limsup_{n\to\infty} \sqrt[n]{|c_n(x-a)^n|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|}\,|x-a| = \infty.$$

In view of the root test, this shows that the power series diverges if $x \neq a$.
(ii), (iii): If $\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} \in \mathbb{R}$, then

$$\limsup_{n\to\infty} \sqrt[n]{|c_n(x-a)^n|} = |x-a| \limsup_{n\to\infty} \sqrt[n]{|c_n|}.$$

Hence the power series converges for all $x \in \mathbb{R}$ if $\displaystyle\limsup_{n\to\infty} \sqrt[n]{|c_n|} = 0$, proving (ii).

If finally $0 < \limsup_{n\to\infty} \sqrt[n]{|c_n|} \in \mathbb{R}$, then, by the Root Test, the series converges if
$|x-a| \limsup_{n\to\infty} \sqrt[n]{|c_n|} < 1$, i. e., $|x-a| < \dfrac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$,
and the series diverges if
$|x-a| > \dfrac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$. □

Tutorial 5.3.1. 1. (a) Let $c > 1$ and put $c_n = \sqrt[n]{c} - 1$.
(i) Show that $c_n \geq 0$.
(ii) Show that $\limsup_{n\to\infty} c_n \leq 0$. **Hint.** Use Bernoulli's inequality.
(iii) Conclude that $\lim_{n\to\infty} \sqrt[n]{c} = 1$.
(b) Use (a) to show that $\lim_{n\to\infty} \sqrt[n]{c} = 1$ for all $c > 0$.

2. Consider $\sum_{n=1}^{\infty} a_n$ with $a_n \neq 0$ for all $n \in \mathbb{N}$. Show that

$$\liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \leq \liminf_{n\to\infty} \sqrt[n]{|a_n|} \leq \limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|.$$

What can you say if $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists or is ∞?

3. Consider the power series $\sum_{n=1}^{\infty} a_n(x-a)^n$ with $a_n \neq 0$ for all $n \in \mathbb{N}$. Using tutorial problem 2 above or otherwise, prove that if $R = \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|$ exists or is ∞, then R is the radius of convergence of the power series.

4. Prove that $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

5. Find the radius and interval of convergence for each of the following power series:

(a) $\sum_{n=1}^{\infty} \frac{(2x)^n}{n}$, (b) $\sum_{n=0}^{\infty} \frac{x^n}{n^n}$, (c) $\sum_{n=0}^{\infty} n^n x^n$,

(d) $\sum_{n=1}^{\infty} \frac{(x-1)^n}{3^n\sqrt{n}}$, (e) $\sum_{n=1}^{\infty} \frac{(-2x)^n}{n^3}$, (f) $\sum_{n=0}^{\infty} (-1)^n x^n$,

(g) $\sum_{n=0}^{\infty} \left(\frac{n}{n+1}\right)^{n^2} x^n$, (h) $\sum_{n=1}^{\infty} \frac{(x+3)^n}{n^3}$, (i) $\sum_{n=0}^{\infty} \frac{(nx)^n}{(2n)!}$.

Chapter 6

Riemann Integration

6.1 Suprema and Infima (self study)

Recall that if a nonempty set $A \subset \mathbb{R}$ is bounded above, then the supremum of A is the **least upper bound** of A, and if A is bounded below, then the infimum of A is the **greatest lower bound** of A. Hence for all $x \in A$ we have $\inf A \le x \le \sup A$ and for each $\varepsilon > 0$ there exist $p, q \in A$ such that $p < \varepsilon + \inf A$ and $q > -\varepsilon + \sup A$, see Theorems 1.9 and 1.12.

Lemma 6.1. *Let A and B be non-empty bounded sets of real numbers, let $c \in \mathbb{R}$, and define*

$$cA := \{cx \mid x \in A\}, \quad c + A := \{c + x \mid x \in A\} \quad \textit{and} \quad A + B := \{x + y \mid x \in A, y \in B\}.$$

Then cA, $c + A$ and $A + B$ are bounded, and
(a) $\sup(cA) = c \sup A, \quad \inf(cA) = c \inf A, \quad$ *if* $c \ge 0$,
(b) $\sup(cA) = c \inf A, \quad \inf(cA) = c \sup A, \quad$ *if* $c \le 0$,
(c) $\sup(c + A) = c + \sup A, \quad \inf(c + A) = c + \inf A$,
(d) $\sup(A + B) = \sup A + \sup B, \quad \inf(A + B) = \inf A + \inf B$.

Proof. (a) Let $M = \sup A$. Then $x \le M$ for all $x \in A$ and therefore $cx \le cM$ for all $cx \in cA$. Hence cA is bounded above with bound cM. This shows that $L = \sup(cA)$ satisfies $L \le cM$.
If $c = 0$ then $L = \sup\{0\} = 0 = 0 \cdot M = cM$, and if $c > 0$, the the above inequality gives

$$cM = c \sup(c^{-1}(cA)) \le \sup(cA) = L.$$

Hence $L \le cM$ and $cM \le L$, which proves (a) for supremum.
The proof for the infimum is similar, or we can use Proposition 1.10, with $A = -(-A)$.
(b) Since $cA = -((-c)A)$ and $-c \ge 0$, this follows from (a) and Proposition 1.10.
(d) This is Tutorial 1.2.1, part 7.
(c) is a special case of (d) if one observes $c + A = \{c\} + A$ and $\inf\{c\} = \sup\{c\} = c$. ☐

Lemma 6.2. *If $\emptyset \ne A \subset B \subset \mathbb{R}$, then*

$$\sup A \le \sup B, \quad \inf A \ge \inf B.$$

Proof. Assume B is bounded. Let $a \in A$. Then $a \in B$, and hence $\inf B \le a \le \sup B$. It follows that A is bounded, that $\inf B$ is a lower bound of A and that $\sup A$ is an upper bound of A. Hence $\inf A \ge \inf B$ and $\sup A \le \sup B$. The cases that $\inf B = -\infty$ or $\sup B = \infty$ are similar. ☐

Lemma 6.3. *Let A and B be nonempty subsets of $\mathbb{R}$ such that $x \le y$ for all $x \in A$ and $y \in B$. Then A is bounded above, B is bounded below, and*

$$\sup A \le \inf B.$$

Proof. Let $b \in B$. Then $a \le b$ for all $a \in A$. So b is an upper bound of A. Hence A is bounded above and $\sup A \le b$ since $\sup A$ is the least upper bound. Then $\sup A \le b$ for all $b \in B$, and therefore $\sup A$ is a lower bound for B. Thus $\sup A \le \inf B$ follows. ☐

6.2 Riemann Sums

The aim of this part of your course is to formalize the Riemann integral and consequently prove the major theorems concerning its properties.

Definition 6.1. A **partition**, P, of the interval $[a, b]$ is a finite set of points $a = x_0 < x_1 < \cdots < x_N = b$ which for brevity will be denoted

$$P = \{a = x_0 < x_1 < \cdots < x_N = b\}.$$

The set of all partitions of the interval $[a, b]$ will be denoted by $\mathcal{P}(a, b)$.

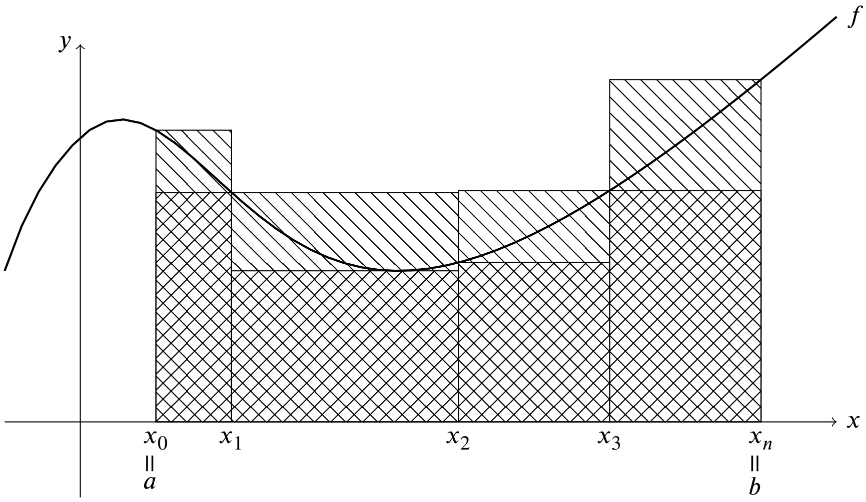
Next we consider the upper and lower Riemann sums with respect to a partition.

Definition 6.2. Let f be a bounded function on $[a, b]$, (i.e. $f : [a, b] \to \mathbb{R}$ and there exists a constant K such that $|f(x)| \leq K$ for all $x \in [a, b]$) and let $P = \{a = x_0 < \cdots < x_N = b\}$ be a partition of the closed finite interval $[a, b]$.
(a) We define **the upper sum** of f over $[a, b]$ with respect to the partition P by

$$U(f, P) := \sum_{j=1}^{N} (x_j - x_{j-1}) \sup\{f(x) \mid x \in [x_{j-1}, x_j]\}.$$

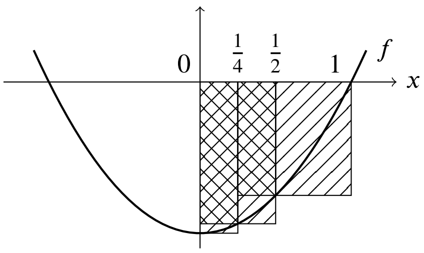
(b) We define **the lower sum** of f over $[a, b]$ with respect to the partition P by

$$L(f, P) := \sum_{j=1}^{N} (x_j - x_{j-1}) \inf\{f(x) \mid x \in [x_{j-1}, x_j]\}.$$



Observe that if $f(x) \geq 0$ on $[a, b]$, then both $L(f, P)$ and $U(f, P)$ are approximations of the area under the curve $y = f(x)$ for $x \in [a, b]$. Here $L(f, P)$ under approximates the area, while $U(f, P)$ over approximates the area.

Worked Example 6.2.1. Find $U(f, P)$ and $L(f, P)$ for $f(x) = x^2 - 1$ and $P = \{0 < 1/4 < 1/2 < 1\}$.
Solution.



Upper Sum:

$$\sup_{x\in[0,\frac{1}{4}]} f(x) = \frac{1}{16} - 1 = \frac{-15}{16}$$

$$\sup_{x\in[\frac{1}{4},\frac{1}{2}]} f(x) = \frac{1}{4} - 1 = \frac{-3}{4}$$

$$\sup_{x\in[\frac{1}{2},1]} f(x) = 1 - 1 = 0$$

$$\Rightarrow U(f,P) = \frac{-15}{16}\left(\frac{1}{4}-0\right) + \frac{-3}{4}\left(\frac{1}{2}-\frac{1}{4}\right) + 0\left(1-\frac{1}{2}\right) = -\frac{27}{64}$$

Lower Sum:

$$\inf_{x\in[0,\frac{1}{4}]} f(x) = 0 - 1 = -1, \qquad \inf_{x\in[\frac{1}{4},\frac{1}{2}]} f(x) = \frac{1}{16} - 1 = \frac{-15}{16}, \qquad \inf_{x\in[\frac{1}{2},1]} f(x) = \frac{1}{4} - 1 = \frac{-3}{4}$$

$$\Rightarrow L(f,P) = -1\left(\frac{1}{4}-0\right) + \frac{-15}{16}\left(\frac{1}{2}-\frac{1}{4}\right) + \frac{-3}{4}\left(1-\frac{1}{2}\right) = -\frac{55}{64}$$

Remark. From 1st year: $\int_0^1 (x^2-1)\,dx = \left[\frac{x^3}{3} - x\right]_0^1 = -\frac{2}{3}$ so we observe that

$$-\frac{55}{64} = L(x^2-1,P) \le \int_0^1 (x^2-1)\,dx \le U(f,P) = -\frac{27}{64}.$$

The following lemma gives some useful bounds.

Lemma 6.4. *Let f be a bounded function on the closed interval $[a,b]$ and let P be a partition of $[a,b]$. Then*

$$(b-a)\inf\{f(x) \mid x\in[a,b]\} \le L(f,P) \le U(f,P) \le (b-a)\sup\{f(x) \mid x\in[a,b]\}.$$

Proof. From the definition of supremum and infimum:

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

Also, since $[x_{j-1},x_j]\subset[a,b]$, it follows from Lemma 6.2 that

$$\sup_{x\in[x_{j-1},x_j]} f(x) \le \sup_{x\in[a,b]} f(x),$$
$$\inf_{x\in[a,b]} f(x) \le \inf_{x\in[x_{j-1},x_j]} f(x).$$

Thus

$$\inf_{x\in[a,b]} f(x) \le \inf_{x\in[x_{j-1},x_j]} f(x) \le \sup_{x\in[x_{j-1},x_j]} f(x) \le \sup_{x\in[a,b]} f(x).$$

Multiplying by $(x_j - x_{j-1})$ (>0) and summing up over j gives

$$\begin{array}{ccccccc}
\sum_{j=1}^{n}(x_j-x_{j-1})\inf_{x\in[a,b]} f(x) & \le & \sum_{j=1}^{n}(x_j-x_{j-1})\inf_{x\in[x_{j-1},x_j]} f(x) & \le & \sum_{j=1}^{n}(x_j-x_{j-1})\sup_{x\in[x_{j-1},x_j]} f(x) & \le & \sum_{j=1}^{n}(x_j-x_{j-1})\sup_{x\in[a,b]} f(x) \\
\| & & \| & & \| & & \| \\
(b-a)\inf_{x\in[a,b]} f(x) & & L(f,P) & & U(f,P) & & (b-a)\sup_{x\in[a,b]} f(x)
\end{array}$$

□

Worked Example 6.2.2. Let $f(x) = x^2$ and $P = \{1 < 4/3 < 5/3 < 2\}$. With $a = 1$ and $b = 2$ we have

$$\inf_{[1,2]} f = 1, \quad \sup_{[1,2]} f = 4, \quad b - a = 1.$$

Thus we conclude

$$1 \le L(f, P) \le U(f, P) \le 4.$$

Note to lecturers: This calculation may be skipped. In fact, since f is increasing on $[1, 2]$, infimum and supremum on each subdivision interval are taken at the left and right endpoint, respectively. Therefore

$$\begin{aligned} U(f,P) &= \frac{1}{3}\cdot\frac{16}{9} + \frac{1}{3}\cdot\frac{25}{9} + \frac{1}{3}\cdot 4 = \frac{77}{27} \\ L(f,P) &= \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot\frac{16}{9} + \frac{1}{3}\cdot\frac{25}{9} = \frac{50}{27}, \end{aligned}$$

i. e.

$$\begin{array}{ccccccc} 1 & \le & \frac{50}{27} & \le & \frac{77}{27} & \le & 4 \\ \| & & \| & & \| & & \| \\ (b-a)\inf\limits_{[1,2]} f & & L(f,P) & & U(f,P) & & (b-a)\sup\limits_{[1,2]} f \end{array}$$

Tutorial 6.2.1. 1. Let $f(x) = x - \frac{1}{2}$ and $P = \{0 < 1/2 < 1 < 3/2 < 2\}$. Find $U(f, P)$ and $L(f, P)$ and compare them with $\int_0^2 f(x)\,dx$.
2. Let $f(x) = x$ and $P_N = \{-1 < -1 + \frac{1}{N} < -1 + \frac{2}{N} < \cdots < -1 + \frac{3N}{N} = 2\}$, $N \in \mathbb{N}^*$.
(i) Find $\sup\left\{ f(x) \mid x \in \left[-1 + \frac{j-1}{N}, -1 + \frac{j}{N}\right]\right\}$ and $\inf\left\{ f(x) \mid x \in \left[-1 + \frac{j-1}{N}, -1 + \frac{j}{N}\right]\right\}$.

(ii) Hence compute $U(f, P_N)$ and $L(f, P_N)$.

(iii) Find $\lim\limits_{N\to\infty} U(f, P_N)$ and $\lim\limits_{N\to\infty} L(f, P_N)$.

(iv) Compare these limits with $\int_{-1}^2 x\,dx$.

3. Verify Lemma 6.4 for the function $\operatorname{sgn}(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$ on the interval $[-2, 3]$ with partition $P = \{-2 < -1 < 0 < 1 < 2 < 3\}$.

6.3 Refinements

Definition 6.3. Let $P = \{a = x_0 < x_1 < \cdots < x_m = b\}$ and $Q = \{a = y_0 < y_1 < \cdots < y_n = b\}$ be partitions of $[a, b]$. Then Q is said to be a **refinement** of P if $P \subset Q$ i.e. $\{x_0, x_1, \ldots, x_m\} \subset \{y_0, y_1, \ldots, y_n\}$.

Q is called a refinement because Q chops $[a, b]$ into "smaller" pieces than P did. Or you can think of P as first cutting $[a, b]$ into pieces and then Q additionally chops some of those pieces into smaller pieces.

Worked Example 6.3.1. 1. $P = \{0 < 1/2 < 1\}$ and $Q = \{0 < 1/3 < 2/3 < 1\}$ are both partitions of $[0, 1]$ but Q is not a refinement of P as $\frac{1}{2}$ belongs to the partition P but not to Q.
2. $P = \{0 < 1/2 < 1\}$ and $Q = \{0 < 1/4 < 2/4 < 3/4 < 1\}$ are both partitions of $[0, 1]$ and Q is a refinement of P.
3. Let $P_n = \{0 < \frac{1}{2^n} < \frac{2}{2^n} < \cdots < \frac{2^n}{2^n}\}$. Then P_n is a partition of $[0, 1]$ for each $n \in \mathbb{N}$ and $P_n \subset P_{n+1}$ since $\frac{j}{2^n} = \frac{2j}{2^{n+1}}$.

Lemma 6.5. *Let f be a bounded function on $[a,b]$ and let P and R be partitions of $[a,b]$ with R a refinement of P. Then*

$$L(f,P) \leq L(f,R) \leq U(f,R) \leq U(f,P).$$

Proof. Since R can be formed by adding a finite number of partition points to P, proceeding inductively it suffices to consider the case where R is formed by introducing one extra point to P:
Suppose $P = \{x_0 < x_1 < \cdots < x_N\}$ and $R = \{x_0 < x_1 < \cdots < x_{k-1} < y < x_k < \cdots < x_N\}$. We have

$$\begin{aligned} U(f,P) &= \sum_{j=1}^{n}(x_j - x_{j-1}) \sup_{x\in[x_{j-1},x_j]} f(x) \\ &= \sum_{\substack{j=1 \\ j\neq k}}^{n}(x_j - x_{j-1}) \sup_{x\in[x_{j-1},x_j]} f(x) + (x_k - x_{k-1}) \sup_{x\in[x_{k-1},x_k]} f(x) \end{aligned}$$

But $Q = \{x_{k-1} < y < x_k\}$ is a partition of $[x_{k-1}, x_k]$, and Lemma 6.4 gives

$$(y - x_{k-1}) \sup_{x\in[x_{k-1},y]} f(x) + (x_k - y) \sup_{x\in[y,x_k]} f(x) \leq (x_k - x_{k-1}) \sup_{x\in[x_{k-1},x_k]} f(x).$$

Hence

$$\begin{aligned} U(f,P) &\geq \sum_{\substack{j=1 \\ j\neq k}}^{n}(x_j - x_{j-1}) \sup_{x\in[x_{j-1},x_j]} f(x) + (y - x_{k-1}) \sup_{x\in[x_{k-1},y]} f(x) + (x_k - y) \sup_{x\in[y,x_k]} f(x) \\ &= U(f,R). \end{aligned}$$

The inequality $L(f,P) \leq L(f,R)$ follows in a similar way, and $L(f,R) \leq U(f,R)$ has been shown in Lemma 6.4. □

Definition 6.4. If P and Q are partitions of $[a,b]$, then $P \cup Q$ (with elements listed in increasing order and without repetitions) is called the **common refinement** of P and Q.

Worked Example 6.3.2. Let $P = \{-1 < 0 < 1\}$ and $Q = \{-1 < -1/3 < 1/3 < 1\}$. Then the common refinement of P and Q is $P \cup Q = \{-1 < -1/3 < 0 < 1/3 < 1\}$.

Lemma 6.6. *If P and Q are partitions of $[a,b]$ and f is a bounded function on $[a,b]$, then*

$$L(f,Q) \leq U(f,P).$$

Proof. Since

$$P \subset P \cup Q \quad \text{and} \quad Q \subset P \cup Q,$$

Lemma 6.5 gives

$$L(f,Q) \leq L(f,P\cup Q), \quad U(f,P\cup Q) \leq U(f,P) \quad \text{and} \quad L(f,P\cup Q) \leq U(f,P\cup Q).$$

Hence

$$L(f,Q) \leq L(f,P\cup Q) \leq U(f,P\cup Q) \leq U(f,P).$$ □

Tutorial 6.3.1. 1. Let $P_n = \{0 < \frac{1}{n} < \frac{2}{n} < \cdots < \frac{n}{n}\}$. Then P_n is a partition of $[0,1]$ for each $n \in \mathbb{N}$. Show that P_{n+1} is not a refinement of P_n for $n \geq 2$.
2. Let $f(x) = x^3 + x^2$, $P = \{-1 < 0 < 1\}$, $Q = \{-1 < -1/2 < 0 < 1\}$. Compute $L(f,P)$, $L(f,Q)$, $U(f,P)$, $U(f,Q)$ and compare their values.

6.4 The Riemann Integral

Definition 6.5. Let f be a bounded function on the interval $[a, b]$.
(a) We define the **lower integral** of f over $[a, b]$ as

$$\underline{\int_a^b} f(t)\,dt := \sup\{L(f, P) \mid P \text{ a partition of } [a, b]\}.$$

(b) We define the **upper integral** of f over $[a, b]$ as

$$\overline{\int_a^b} f(t)\,dt := \inf\{U(f, P) \mid P \text{ a partition of } [a, b]\}.$$

(c) We say that f is Riemann integrable on $[a, b]$ if

$$\underline{\int_a^b} f(t)\,dt = \overline{\int_a^b} f(t)\,dt,$$

in which case the common value is denoted by $\int_a^b f(t)\,dt$ and is called the **Riemann integral of f over** $[a, b]$.

Note. By the properties of suprema and infima, see Lemmas 6.6 and 6.3,

$$\underline{\int_a^b} f(t)\,dt \le \overline{\int_a^b} f(t)\,dt.$$

Hence f is Riemann integrable if and only if

$$\overline{\int_a^b} f(t)\,dt \le \underline{\int_a^b} f(t)\,dt.$$

Theorem 6.7 (Integrability Condition)**.** *A bounded function f on the interval $[a, b]$ is Riemann integrable if and only if for each $\varepsilon > 0$ there exists a partition P of $[a, b]$ such that*

$$U(f, P) - L(f, P) < \varepsilon.$$

Proof. $\Rightarrow$: Let f be Riemann integrable on $[a, b]$ and let $\varepsilon > 0$. Then there is a partition P of $[a, b]$ such that

$$\overline{\int_a^b} f(t)\,dt \le U(f, P) < \overline{\int_a^b} f(t)\,dt + \frac{\varepsilon}{2}$$

and there is a partition Q of $[a, b]$ such that

$$\underline{\int_a^b} f(t)\,dt - \frac{\varepsilon}{2} < L(f, Q) \le \underline{\int_a^b} f(t)\,dt.$$

Let $R = P \cup Q$. Then, in view of Lemma 6.5,

$$\int_a^b f(t)\,dt - \frac{\varepsilon}{2} = \underline{\int_a^b} f(t)\,dt - \frac{\varepsilon}{2} < L(f, Q) \le L(f, R)$$

$$\le U(f, R) \le U(f, P) < \overline{\int_a^b} f(t)\,dt + \frac{\varepsilon}{2} = \int_a^b f(t)\,dt + \frac{\varepsilon}{2}.$$

In particular,

$$U(f, R) < \int_a^b f(t)\,dt + \frac{\varepsilon}{2},$$

$$-L(f, R) < -\int_a^b f(t)\,dt + \frac{\varepsilon}{2}.$$

Summing up gives

$$U(f, R) - L(f, R) < \varepsilon.$$

$\Leftarrow$: Suppose that there are $\varepsilon > 0$ and a partition P_ε of $[a, b]$ such that

$$U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon.$$

Then from

$$\underline{\int_a^b} f(t)\,dt \ge L(f, P_\varepsilon),$$

$$\overline{\int_a^b} f(t)\,dt \le U(f, P_\varepsilon)$$

we get

$$\overline{\int_a^b} f(t)\,dt - \underline{\int_a^b} f(t)\,dt \le U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon.$$

Hence

$$\forall\, \varepsilon > 0 \quad 0 \le \overline{\int_a^b} f(t)\,dt - \underline{\int_a^b} f(t)\,dt \le \varepsilon,$$

giving

$$\overline{\int_a^b} f(t)\,dt = \underline{\int_a^b} f(t)\,dt.$$

□

Theorem 6.8. *If f is an increasing function on the interval $[a, b]$, then f is Riemann integrable on $[a, b]$.*

Proof. By assumption, $f(a) \le f(x) \le f(b)$ for all $x \in [a, b]$. Hence f is bounded.
Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that

$$(b - a)\frac{f(b) - f(a)}{N} < \varepsilon.$$

Consider the partition

$$P = \left\{ a < a + \frac{b-a}{N} < a + \frac{2(b-a)}{N} < \cdots < a + \frac{(N-1)(b-a)}{N} < a + \frac{N(b-a)}{N} = b \right\}.$$

Observing that over each subinterval of $[a, b]$ the increasing function f takes its minimum at the left endpoint and its maximum at the right endpoint of the interval, it follows that

$$U(f,P) = \sum_{j=1}^{N} \frac{b-a}{N} \sup \left\{ f(x) \,\middle|\, x \in \left[a + \frac{(j-1)(b-a)}{N}, a + \frac{j(b-a)}{N} \right] \right\} = \sum_{j=1}^{N} \frac{b-a}{N} f\left(a + \frac{j(b-a)}{N} \right),$$
$$L(f,P) = \sum_{j=1}^{N} \frac{b-a}{N} \inf \left\{ f(x) \,\middle|\, x \in \left[a + \frac{(j-1)(b-a)}{N}, a + \frac{j(b-a)}{N} \right] \right\} = \sum_{j=1}^{N} \frac{b-a}{N} f\left(a + \frac{(j-1)(b-a)}{N} \right).$$

Hence

$$\begin{aligned} U(f,P) - L(f,p) &= \sum_{j=1}^{N} \frac{b-a}{N} f\left(a + \frac{j(b-a)}{N} \right) - \sum_{j=0}^{N-1} \frac{b-a}{N} f\left(a + \frac{j(b-a)}{N} \right) \\ &= \frac{b-a}{N}(f(b) - f(a)) < \varepsilon. \end{aligned}$$

Hence f is integrable by Theorem 6.7. $\square$

Worked Example 6.4.1. 1. Let $f(x) = x$. Then f is Riemann integrable on $[0, 1]$ and $\int_0^1 f(t)\,dt = \frac{1}{2}$. To see this let $P_n = \{0 < \frac{1}{n} < \frac{2}{n} < \cdots < \frac{n}{n}\}$. Then

$$L(f,P_n) = \sum_{j=1}^{n} \frac{1}{n} \frac{j-1}{n} = \frac{1}{n^2} \frac{(n-1)n}{2} = \frac{1}{2}\left(1 - \frac{1}{n}\right)$$

and

$$U(f,P_n) = \sum_{j=1}^{n} \frac{1}{n} \frac{j}{n} = \frac{1}{n^2} \frac{n(n+1)}{2} = \frac{1}{2}\left(1 + \frac{1}{n}\right).$$

Thus $U(f, P_n) - L(f, P_n) = \frac{1}{n}$, which can be made arbitrarily small but taking n sufficiently large. Thus f is Riemann integrable. Also the integral is sandwiched between $L(f, P_n)$ and $U(f, P_n)$, both of which converge to $\frac{1}{2}$, giving $\int_0^1 f(t)\,dt = \frac{1}{2}$.

2. Let $f(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ 1 & \text{if } x \in \mathbb{Q} \end{cases}$. Then f is not Riemann integrable on $[0, 1]$, since for each partition P of $[0, 1]$, $U(f, P) = 1$, while $L(f, P) = 0$.

3. Let $f(x) = \operatorname{sgn}(x)$ on $[-1, 1]$. Let $P_n := \{-1 < -\frac{1}{n} < \frac{1}{n} < 1\}$ then $U(\operatorname{sgn}, P_n) = \frac{2}{n}$ and $L(\operatorname{sgn}, P_n) = -\frac{2}{n}$. Thus $U(\operatorname{sgn}, P_n) - L(\operatorname{sgn}, P_n) = \frac{4}{n} \to 0$ as $n \to \infty$. For each $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $\frac{4}{n} < \varepsilon$. Hence sgn is Riemann integrable on $[-1, 1]$. Finally, $\lim_{n\to\infty} U(\operatorname{sgn}, P_n) = \lim_{n\to\infty} L(\operatorname{sgn}, P_n) = 0$ gives $\int_{-1}^1 \operatorname{sgn}(x)\,dx = 0$.

Tutorial 6.4.1. Let $f : [a, b] \to \mathbb{R}$ be bounded. For $n \in \mathbb{N}^*$ let $P_n \in \mathcal{P}(a, b)$ such that $\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n)$. Show that f is Riemann integrable on $[a, b]$ and that $\int_a^b f(x)\,dx = \lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n)$.

Tutorial 6.4.2. Show that $f(x) = x^2$ is Riemann integrable on $[0, 1]$ and that $\int_0^1 f(t)\,dt = \frac{1}{3}$.

6.5 Properties of the Integral

Theorem 6.9 (Homogeneity)**.** *If f is a bounded Riemann integrable function on the interval $[a, b]$ and if $c \in \mathbb{R}$, then cf is Riemann integrable and*

$$\int_a^b cf(t)\,dt = c\int_a^b f(t)\,dt.$$

Proof. Note that whenever f is bounded, then also cf is bounded.
Case I: $c \geq 0$.
Let $P = \{a = x_0 < x_1 < \cdots < x_N = b\}$ be a partition of $[a, b]$. Then, by Lemma 6.1,

$$\begin{aligned}
U(cf, P) &= \sum_{j=1}^{N}(x_j - x_{j-1}) \sup\{cf(x) \mid x \in [x_{j-1}, x_j]\} \\
&= \sum_{j=1}^{N}(x_j - x_{j-1})c \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\
&= c\sum_{j=1}^{N}(x_j - x_{j-1}) \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\
&= cU(f, P).
\end{aligned}$$

Thus

$$\begin{aligned}
\overline{\int_a^b} cf(t)\,dt &= \inf\{U(cf, P) \mid P \in \mathcal{P}(a, b)\} \\
&= \inf\{cU(f, P) \mid P \in \mathcal{P}(a, b)\} \\
&= c\inf\{U(f, P) \mid P \in \mathcal{P}(a, b)\} \\
&= c\overline{\int_a^b} f(t)\,dt.
\end{aligned}$$

Interchanging sup and inf we get $L(cf, P) = cL(f, P)$ and then

$$\underline{\int_a^b} cf(t)\,dt = c\underline{\int_a^b} f(t)\,dt.$$

But $\underline{\int_a^b} f(t)\,dt = \overline{\int_a^b} f(t)\,dt$ now gives

$$\underline{\int_a^b} cf(t)\,dt = c\underline{\int_a^b} f(t)\,dt = c\int_a^b f(t)\,dt = c\overline{\int_a^b} f(t)\,dt = \overline{\int_a^b} cf(t)\,dt.$$

Hence cf is Riemann integrable and $\displaystyle\int_a^b cf(t)\,dt = c\int_a^b f(t)\,dt$.

Case II: $c = -1$. Let $P = \{a = x_0 < x_1 < \cdots < x_N = b\}$ be a partition of $[a, b]$. Then, using the results of Section 6.1,

$$\begin{aligned} U(-f, P) &= \sum_{j=1}^{N}(x_j - x_{j-1}) \sup\{-f(x) \mid x \in [x_{j-1}, x_j]\} \\ &= -\sum_{j=1}^{N}(x_j - x_{j-1}) \inf\{f(x) \mid x \in [x_{j-1}, x_j]\} \\ &= -L(f, P). \end{aligned}$$

Thus

$$\begin{aligned} \overline{\int_a^b} (-f(t))\, dt &= \inf\{U(-f, P) \mid P \in \mathcal{P}(a, b)\} \\ &= \inf\{-L(f, P) \mid P \in \mathcal{P}(a, b)\} \\ &= -\sup\{L(f, P) \mid P \in \mathcal{P}(a, b)\} \\ &= -\underline{\int_a^b} f(t)\, dt. \end{aligned}$$

Since so far we have only used that f and $-f$ are bounded, we can replace f by $-f$ and obtain

$$\overline{\int_a^b} f(t)\, dt = \overline{\int_a^b} (-(-f(t)))\, dt = -\underline{\int_a^b} (-f(t))\, dt,$$

and therefore

$$\underline{\int_a^b} (-f(t))\, dt = -\overline{\int_a^b} f(t)\, dt.$$

Thus

$$\int_a^b (-f(t))\, dt = -\overline{\int_a^b} f(t)\, dt = -\int_a^b f(t)\, dt = -\underline{\int_a^b} f(t)\, dt = \overline{\int_a^b} (-f(t))\, dt.$$

This shows that $-f$ is Riemann integrable and that

$$\int_a^b (-f(t))\, dt = -\int_a^b f(t)\, dt.$$

Case III: $c < 0$.
By Case I, $(-c)f$ is Riemann integrable, and

$$\int_a^b (-c)f(t)\, dt = (-c)\int_a^b f(t)\, dt.$$

Then, by Case II, $-(-c)f$ is Riemann integrable, and

$$\int_a^b (-(-c)f(t))\, dt = -\int_a^b (-c)f(t)\, dt.$$

Thus cf is Riemann integrable, and

$$\int_a^b cf(t)\,dt = -\int_a^b (-c)f(t)\,dt = -(-c)\int_a^b f(t)\,dt = c\int_a^b f(t)\,dt. \qquad \square$$

Lemma 6.10. *If f and g are bounded functions on $[a,b]$ and P is a partition of $[a,b]$, then*

$$U(f+g,P) \le U(f,P) + U(g,P),$$
$$L(f+g,P) \ge L(f,P) + L(g,P).$$

Proof. Let $P = \{x_0 < x_1 < \cdots < x_N\}$ be a partion of $[a,b]$. For each $j \in \{1,\dots,N\}$ and $t \in [x_{j-1},x_j]$ we have

$$f(t) + g(t) \le \sup\{f(x) \mid x \in [x_{j-1},x_j]\} + \sup\{g(x) \mid x \in [x_{j-1},x_j]\},$$

whence

$$\sup\{f(x)+g(x) \mid x \in [x_{j-1},x_j]\} \le \sup\{f(x) \mid x \in [x_{j-1},x_j]\} + \sup\{g(x) \mid x \in [x_{j-1},x_j]\}.$$

Multiplying with $(x_j - x_{j-1})$ and summing up over j gives

$$U(f+g,P) \le U(f,P) + U(g,P).$$

$L(f+g,P) \ge L(f,P)+L(g,P)$ can be shown similarly, or we may use the estimate for upper sums and $L(f,P) = -U(-f,P)$. $\square$

Theorem 6.11 (Additivity)**.** *If f and g are bounded Riemann integrable functions on the interval $[a,b]$, then $f+g$ is Riemann integrable and*

$$\int_a^b (f(t)+g(t))\,dt = \int_a^b f(t)\,dt + \int_a^b g(t)\,dt.$$

Proof. Let $\varepsilon > 0$ and choose $P_1, P_2 \in \mathcal{P}(a,b)$ such that

$$\overline{\int_a^b} f(t)\,dt \ge U(f,P_1) - \frac{\varepsilon}{2},$$
$$\overline{\int_a^b} g(t)\,dt \ge U(g,P_2) - \frac{\varepsilon}{2}.$$

Let $P = P_1 \cup P_2$. By Lemmas 6.5 and 6.10,

$$U(f+g,P) \le U(f,P) + U(g,P) \le U(f,P_1) + U(g,P_2).$$

Thus

$$\begin{aligned}\overline{\int_a^b} (f(t)+g(t))\,dt &\le U(f+g,P) \le U(f,P_1) + U(g,P_2)\\ &\le \overline{\int_a^b} f(t)\,dt + \overline{\int_a^b} g(t)\,dt + \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows that

$$\overline{\int_a^b} (f(t)+g(t))\,dt \le \overline{\int_a^b} f(t)\,dt + \overline{\int_a^b} g(t)\,dt.$$

Similarly,

$$\underline{\int_a^b} (f(t) + g(t))\, dt \geq \underline{\int_a^b} f(t)\, dt + \underline{\int_a^b} g(t)\, dt.$$

But since f and g are integrable over $[a, b]$

$$\overline{\int_a^b} f(t)\, dt + \overline{\int_a^b} g(t)\, dt = \int_a^b f(t)\, dt + \int_a^b g(t)\, dt = \underline{\int_a^b} f(t)\, dt + \underline{\int_a^b} g(t)\, dt,$$

and therefore

$$\overline{\int_a^b} (f(t) + g(t))\, dt \leq \int_a^b f(t)\, dt + \int_a^b g(t)\, dt \leq \underline{\int_a^b} (f(t) + g(t))\, dt,$$

which shows that $f + g$ is integrable, and

$$\int_a^b (f(t)\, dt + g(t))\, dt = \int_a^b f(t)\, dt + \int_a^b g(t)\, dt. \qquad \square$$

Theorem 6.12. *If f is a bounded Riemann integrable function on the interval $[a, b]$ and on the interval $[b, c]$, then f is Riemann integrable on $[a, c]$ and*

$$\int_a^c f(x)\, dx = \int_a^b f(x)\, dx + \int_b^c f(x)\, dx.$$

Proof. Let P be a partition of $[a, b]$ and Q be a partition of $[b, c]$. Writing

$$\begin{aligned} P &= \{a = x_0 < x_1 < \cdots < x_m = b\}, \\ Q &= \{b = x_m < x_{m+1} < \cdots < x_n = c\}, \end{aligned}$$

we have that $R = \{a = x_0 < x_1 < \cdots < x_n = c\}$ is a partition of $[a, c]$ and

$$\begin{aligned} U(f, R) &= \sum_{j=1}^{n} (x_j - x_{j-1}) \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\ &= \sum_{j=1}^{m} (x_j - x_{j-1}) \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} + \sum_{j=m+1}^{n} (x_j - x_{j-1}) \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\ &= U(f, P) + U(f, Q). \end{aligned}$$

Hence, by Lemmas 6.1 and 6.2,

$$\begin{aligned} \overline{\int_a^c} f(x)\, dx &= \inf\{U(f, S) \mid S \in \mathcal{P}(a, c)\} \\ &\leq \inf\{U(f, R) \mid R \in \mathcal{P}(a, c),\ b \in R\} \\ &= \inf\{U(f, P) + U(f, Q) \mid P \in \mathcal{P}(a, b), Q \in \mathcal{P}(b, c)\} \\ &= \inf\{U(f, P) \mid P \in \mathcal{P}(a, b)\} + \inf\{U(f, Q) \mid Q \in \mathcal{P}(b, c)\} \\ &= \overline{\int_a^b} f(x)\, dx + \overline{\int_b^c} f(x)\, dx. \end{aligned}$$

Similarly

$$\begin{aligned}\underline{\int_a^c} f(x)\,dx &= \sup\{L(f,S) \mid S \in \mathcal{P}(a,c)\}\\ &\geq \sup\{L(f,P)+L(f,Q) \mid P \in \mathcal{P}(a,b), Q \in \mathcal{P}(b,c)\}\\ &= \underline{\int_a^b} f(x)\,dx + \underline{\int_b^c} f(x)\,dx.\end{aligned}$$

Since f is integrable over $[a,b]$ and over $[b,c]$, it follows from above that

$$\overline{\int_a^c} f(x)\,dx \leq \int_a^b f(x)\,dx + \int_b^c f(x)\,dx \leq \underline{\int_a^c} f(x)\,dx.$$

Therefore f is integrable over $[a,c]$ and

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx. \qquad \square$$

Lemma 6.13. *Let f be a continuous function on $[a,b]$ and let $\varepsilon > 0$. Then there is a partition $\mathcal{P} = \{x_0 < x_1 < \cdots < x_N\}$ of $[a,b]$ such that $|f(x) - f(x_{j-1})| \leq \varepsilon$ for all $j \in \{1,\dots,N\}$ and all $x \in [x_{j-1}, x_j]$.*

Proof. Let $\varepsilon > 0$ and put $x_0 = a$. For $j \in \mathbb{N}^*$ we define numbers x_j recursively as follows: Let

$$E_j = \{x \in [x_{j-1}, b] \mid |f(x) - f(x_{j-1})| \geq \varepsilon\}.$$

Note that $E_j = \emptyset$ if $x_{j-1} = b$.
If $E_j = \emptyset$, let $x_j = b$.
If $E_j \neq \emptyset$, let $x_j = \inf E_j$.
Claim 1. If $x_{j-1} < b$, then $x_{j-1} < x_j$.
If $E_j = \emptyset$, then $x_{j-1} < b = x_j$. Now let $E_j \neq \emptyset$. Since f is continuous at x_{j-1}, there is $\delta > 0$ such that $|f(x)-f(x_{j-1})| < \varepsilon$ whenever $x \in [a,b]$ with $|x-x_{j-1}| < \delta$. Hence $x \in E_j$ implies $x \geq x_{j-1}$ and $|x-x_{j-1}| \geq \delta$, and therefore $x \geq x_{j-1} + \delta$. Thus $x_{j-1} + \delta$ is a lower bound of E_j and therefore

$$x_j = \inf E_j \geq x_{j-1} + \delta > x_{j-1}.$$

Claim 2. If $x_{j-1} < b$, then $|f(x) - f(x_{j-1})| < \varepsilon$ for all $x \in [x_{j-1}, x_j)$, $|f(x_j) - f(x_{j-1})| \leq \varepsilon$, and $|f(x_j) - f(x_{j-1})| = \varepsilon$ if $E_j \neq \emptyset$.
Indeed, if $x \in [x_{j-1}, x_j)$, then $x \in [x_{j-1}, b]$ and $x < \inf E_j$, and so $x \notin E_j$. By definition of E_j it follows that $|f(x) - f(x_{j-1})| < \varepsilon$, and, since $x_{j-1} < x_j$ by Claim 1, the continuity of f and the absolute value function give

$$|f(x_j) - f(x_{j-1})| = \lim_{x \to x_j^-} |f(x) - f(x_{j-1})| \leq \varepsilon.$$

If $E_j \neq \emptyset$, then there is a sequence (a_k) in E_j such that

$$\lim_{k\to\infty} a_k = x_j.$$

Observing $|f(a_k) - f(x_{j-1})| \geq \varepsilon$, the same continuity argument as above gives

$$|f(x_j) - f(x_{j-1})| = \lim_{k\to\infty} |f(a_k) - f(x_{j-1})| \geq \varepsilon.$$

Claim 3. There is an index n such that $x_n = b$.

Assume that Claim 3 is false. Then, by Claim 1, (x_j) is an increasing sequence, bounded by b, which therefore has a limit, say $\overline{x}$, see Theorem 6.11, and $\overline{x} \in [a, b]$. Then the continuity of f at $\overline{x}$ gives

$$|f(x_j) - f(x_{j-1})| \le |f(x_j) - f(\overline{x})| + |f(\overline{x}) - f(x_{j-1})| \to 0 \text{ as } j \to \infty.$$

Hence there is j such that $|f(x_j) - f(x_{j-1})| < \varepsilon$, contradicting Claim 2 since $x_j \ne b$ implies $E_j \ne \emptyset$. Putting $N = \min\{j \in \mathbb{N} \mid x_j = b\}$ completes the proof. □

Theorem 6.14. *If f is a continuous function on $[a, b]$, then f is Riemann integrable on $[a, b]$.*

Proof. From Theorem 3.17, we know that a continuous function on a closed interval $[a, b]$ is bounded. Let $\varepsilon > 0$ and let $P = \{x_0 < \cdots < x_N\}$ be a partition of $[a, b]$ according to Lemma 6.13. Then, for $j \in \{1, \dots, N\}$, with Lemma 6.1,

$$\begin{aligned}
\sup_{x\in[x_{j-1},x_j]} f(x) - \inf_{x\in[x_{j-1},x_j]} f(x) &= \sup_{x\in[x_{j-1},x_j]} [f(x) - f(x_{j-1})] - \inf_{x\in[x_{j-1},x_j]} [f(x) - f(x_{j-1})] \\
&= \sup_{x\in[x_{j-1},x_j]} [f(x) - f(x_{j-1})] + \sup_{x\in[x_{j-1},x_j]} [f(x_{j-1}) - f(x)] \\
&\le 2 \sup_{x\in[x_{j-1},x_j]} |f(x) - f(x_{j-1})| \\
&\le 2\varepsilon
\end{aligned}$$

in view of Lemma 6.13, Claim 2. Hence

$$\begin{aligned}
U(f, P) - L(f, P) &= \sum_{j=1}^{N} (x_j - x_{j-1}) \left[\sup_{x\in[x_{j-1},x_j]} f(x) - \inf_{x\in[x_{j-1},x_j]} f(x) \right] \\
&\le \sum_{j=1}^{N} (x_j - x_{j-1}) 2\varepsilon = 2(b - a)\varepsilon.
\end{aligned}$$

By Theorem 6.7, f is Riemann integrable. □

Theorem 6.15 (Fundamental Theorem of Calculus)**.** *Let f be a differentiable function on $[a, b]$ such that f' is continuous on $[a, b]$. Then f' is Riemann integrable on $[a, b]$ and*

$$\int_a^b f'(t)\,dt = f(b) - f(a).$$

Proof. Since f' is continuous, f' is Riemann integrable by the previous theorem. Let $P = \{x_0 < \cdots < x_N\}$ be a partition of $[a, b]$. By the First Mean Value Theorem, Theorem 4.5, for each $j \in \{1, \dots, N\}$ there is $c_j \in (x_{j-1}, x_j)$ such that

$$f(x_j) - f(x_{j-1}) = (x_j - x_{j-1}) f'(c_j).$$

Then

$$\begin{aligned}
L(f', P) &= \sum_{j=1}^{N} (x_j - x_{j-1}) \inf_{x\in[x_{j-1},x_j]} f'(x) \\
&\le \sum_{j=1}^{N} (x_j - x_{j-1}) f'(c_j) \\
&= \sum_{j=1}^{N} f(x_j) - f(x_{j-1}) = f(b) - f(a).
\end{aligned}$$

Hence

$$\underline{\int_a^b} f'(x)\,dx = \sup\{L(f', P) \mid P \in \mathcal{P}(a, b)\} \le f(b) - f(a).$$

Similarly,

$$\overline{\int_a^b} f'(x)\,dx \geq f(b) - f(a).$$

Since f' is Riemann integrable, the upper and lower Riemann integrals are equal, and

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

follows. $\square$

Tutorial 6.5.1. Let f be a bounded Riemann integrable function on $[a, b]$ and let P be a partition of $[a, b]$. Then prove that

$$(b-a)\inf\{f(x) \mid x \in [a,b]\} \leq L(f,P) \leq \int_a^b f(t)\,dt \leq U(f,P) \leq (b-a)\sup\{f(x) \mid x \in [a,b]\}.$$

***Tutorial 6.5.2.** Let $a < b$, let $f : [a, b] \to \mathbb{R}$ and let $P = \{a = x_0 < x_1 < \cdots < x_n = b\} \in \mathcal{P}(a, b)$. Define

$$V(f,P) := \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})|$$

and

$$V_a^b f := \sup\{V(f,P) : P \in \mathcal{P}(a,b)\}.$$

One sets $V_a^a f = 0$. The function f is said to be of bounded variation on $[a, b]$ if $V_a^b f < \infty$.

1. Let $a \leq x < y \leq b$. Show that $V_a^x f + |f(y) - f(x)| \leq V_a^y f$.
2. Assume that f is of bounded variation on $[a, b]$ and for $x \in [a, b]$ define

$$f_1(x) = V_a^x f,$$
$$f_2(x) = V_a^x f - f(x).$$

Show that f_1 and f_2 are increasing on $[a, b]$ and that $f = f_1 - f_2$.

3. Show that f is of bounded variation on $[a, b]$ if and only if f is the difference of two increasing functions on $[a, b]$.
4. Show that if f is of bounded variation on $[a, b]$, then f is Riemann integrable over $[a, b]$.
5. Give an example of a continuous function on $[a, b]$ which is not of bounded variation.

Chapter 7

Introduction to Metric Spaces

7.1 Metric Spaces

Definition 7.1. Let X be a set and $d : X \times X \to \mathbb{R}$. d is called a metric on X if for all $x, y, z \in X$ the following properties are satisfied:

(i) $d(x, y) = 0 \Leftrightarrow x = y$,

(ii) $d(x, y) = d(y, x)$ (symmetry),

(iii) $d(x, z) \le d(x, y) + d(y, z)$ (triangle inequality).

If d is a metric on X, then (X, d) or shortly X is called a metric space.

Remark 7.1.1. If (X, d) is a metric space, then $d(x, y) \ge 0$ for all $x, y \in X$.

Proof. From

$$0 \underset{\text{(i)}}{=} d(x, x) \underset{\text{(iii)}}{\le} d(x, y) + d(y, x) \underset{\text{(ii)}}{=} d(x, y) + d(x, y) = 2d(x, y)$$

it follows that $d(x, y) \ge 0$. □

In the following, $n \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$ unless stated otherwise.

Worked Example 7.1.1. 1. For $x, y \in \mathbb{R}$ let $d(x, y) = |x - y|$. Then $(\mathbb{R}, d)$ is a metric space.
2. For $x, y \in \mathbb{C}$ let $d(x, y) = |x - y|$. Then $(\mathbb{C}, d)$ is a metric space.

3. For $x = (x_j)_{j=1}^n, y = (y_j)_{j=1}^n \in \mathbb{C}^n$ let $d(x, y) = \left(\sum_{j=1}^{n} |x_j - y_j|^2 \right)^{\frac{1}{2}}$. Then $(\mathbb{C}^n, d)$ is a metric space.

Solution. Parts 1 and 2 are clear if one observes the triangle inequality for real and complex numbers.
For the triangle inequality in 3, we use the notation

$$\|x\| = d(x, 0).$$

Then $\|x + y\| \le \|x\| + \|y\|$ can be shown as was done for $\mathbb{R}^n$ in Multivariable Calculus. But rather than repeating that proof we use it to prove the result for $\mathbb{C}^n$. To this end, put $\hat{x} = (|x_j|)_{j=1}^n \in \mathbb{R}^n$. Obviously, $\|\hat{x}\| = \|x\|$ and $\|x + y\| \le \|\hat{x} + \hat{y}\|$ by part 2. Then, from Multivarible Calculus results applicable to $\hat{x}$ and $\hat{y}$, it follows that

$$\|x + y\| \le \|\hat{x} + \hat{y}\| \le \|\hat{x}\| + \|\hat{y}\| = \|x\| + \|y\|.$$

Then, for $x, y, z \in \mathbb{C}^n$,

$$\begin{aligned} d(x, z) = \|x - z\| &= \|(x - y) + (y - z)\| \le \|x - y\| + \|y - z\| \\ &= d(x, y) + d(y, z). \end{aligned}$$

In the following, let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$.

Worked Example 7.1.2. Let Y be a nonempty set and let $B(Y) = B(Y, \mathbb{F}^n)$ be the set of all bounded functions from Y to $\mathbb{F}^n$. Here a function f from Y to $\mathbb{F}^n$ is called bounded if there is a real number M such that $\|f(x)\| = \left(\sum_{j=1}^{n} |f_j(x)|^2\right)^{\frac{1}{2}} \le M$ for all $x \in Y$, where $f_j(x)$ is the j-th component of $f(x)$. For $f, g \in B(Y)$ let

$$d(f,g) = \sup_{x\in Y} \|f(x) - g(x)\|.$$

Then $(B(Y), d)$ is a metric space.

Solution. We have for $f, g \in B(Y)$ that there are $M, N \in \mathbb{R}$ such that $\|f(x)\| \le M$ and $\|g(x)\| \le N$ for all $x \in Y$. Denoting the metric on $\mathbb{C}^n$ from Example 7.1.1 by d_n, it follows for $x \in Y$ that

$$\|f(x) - g(x)\| = d_n(f(x), g(x)) \le d_n(f(x), 0) + d_n(g(x), 0) = \|f(x)\| + \|g(x)\| \le M + N,$$

so that $d(f, g) \in \mathbb{R}$.
Clearly, $d(f, f) = 0$ for all $f \in B(Y)$. If, on the other hand, $f \ne g$, then there is $x_0 \in Y$ such that $f(x_0) \ne g(x_0)$. Hence

$$d(f,g) = \sup_{x\in Y} \|f(x) - g(x)\| \ge \|f(x_0) - g(x_0)\| > 0.$$

From $\|f(x) - g(x)\| = \|g(x) - f(x)\|$ it easily follows that $d(f, g) = d(g, f)$.
Finally, for $f, g, h \in B(Y)$,

$$\begin{aligned}
d(f,h) &= \sup_{x\in Y} \|f(x) - h(x)\| \\
&\le \sup_{x\in Y}(\|f(x) - g(x)\| + \|g(x) - h(x)\|) \\
&\le \sup_{x,y\in Y} (\|f(x) - g(x)\| + \|g(y) - h(y)\|) \\
&= \sup_{x\in Y} \|f(x) - g(x)\| + \sup_{y\in Y} \|g(y) - h(y)\| \\
&= d(f,g) + d(g,h).
\end{aligned}$$

Proposition 7.1. *Let $X = (X, d)$ be a metric space and let $Y \subset X$ be nonempty. For $x, y \in Y$ let $d_Y(x, y) = d(x, y)$. Then $Y = (Y, d_Y) = (Y, d)$ is a metric space.*

Note. When it is clear what the metric is, it is convenient to write briefly X for a metric space, and $Y \subset X$ is called a (metric) subspace of X.

Worked Example 7.1.3. 1. $\mathbb{R}^n$ is a subspace of $\mathbb{C}^n$.
2. If Y is a nonempty set and $A \subset \mathbb{F}^n$, then $B(Y, A)$ is a subspace of $B(Y, \mathbb{F}^n)$.

Definition 7.2. A sequence (x_k) in a metric space (X, d) is called
(i) convergent if there is $x \in X$ such that

$$\lim_{k\to\infty} d(x_k, x) = 0.$$

We then write $x = \lim_{k\to\infty} x_k$ and say that (x_k) converges to x.
(ii) a Cauchy sequence if for all $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for all $k, m \ge N$: $d(x_k, x_m) < \varepsilon$.

Theorem 7.2. *Let (X, d) be a metric space and let (x_k) be a convergent sequence in X. Then*
1. $\lim_{k\to\infty} x_k$ *is unique.*
2. (x_k) *is a Cauchy sequence.*

Proof. 1. Let $x, y \in X$ such that $\lim_{k\to\infty} d(x_k, x) = 0$ and $\lim_{k\to\infty} d(x_k, y) = 0$. Then

$$\begin{aligned}
d(x,y) &\le d(x, x_k) + d(x_k, y) \\
&= d(x_k, x) + d(x_k, y)
\end{aligned}$$

gives

$$0 \le d(x, y) \le \lim_{k\to\infty} d(x_k, x) + \lim_{k\to\infty} d(x_k, y) = 0.$$

Hence $d(x, y) = 0$, giving $x = y$.
2. Let $\varepsilon > 0$. Since there is $x \in X$ such that $\lim_{k\to\infty} d(x_k, x) = 0$, there is $N \in \mathbb{N}$ such that $d(x_k, x) < \frac{\varepsilon}{2}$ for $k \ge N$. Then, for $k, m \ge N$,

$$d(x_k, x_m) \le d(x_k, x) + d(x_m, x) < \varepsilon. \qquad \square$$

Tutorial 7.1.1. 1. Let X be a set and let $f : X \to \mathbb{R}$ be injective.
a) Show that $d(x, y) = |f(x) - f(y)|$ defines a metric on X.
b) Show that a sequence (x_k) in (X, d) converges to some $x \in X$ if and only if the sequence $(f(x_k))$ converges to some $y \in$ range f, and that then $y = f(x)$.
2. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, define $f : \overline{\mathbb{R}} \to \mathbb{R}$ by

$$f(x) = \begin{cases} \arctan x & \text{if } x \in \mathbb{R}, \\ -\dfrac{\pi}{2} & \text{if } x = -\infty \text{ ,} \\ \dfrac{\pi}{2} & \text{if } x = \infty, \end{cases}$$

and let d be as in part 1. Show that a sequence (x_k) in $\mathbb{R}$ converges in $(\overline{\mathbb{R}}, d)$ if and only if either (x_k) converges in $\mathbb{R}$ in the usual sense or $x_k \to -\infty$ or $x_k \to \infty$.
3. Let X be a nonempty set. For $x, y \in X$ define

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \ne y. \end{cases}$$

a) Show that (X, d) is a metric space. (X, d) is called a discrete metric space.
b) Show that a sequence (x_k) converges in (X, d) if and only if there is $N \in \mathbb{N}$ such that $x_k = x_N$ for all $k > N$.
4. Let $x, y \in \mathbb{C}^n$. Prove directly that $|(x, y)| \le \|x\|\|y\|$ and $\|x + y\| \le \|x\| + \|y\|$.

7.2 Complete Metric Spaces

Definition 7.3. A metric space (X, d) is called complete if every Cauchy sequence in (X, d) converges.

Theorem 7.3. *$\mathbb{C}^n$ is complete.*

Proof. Let (z_k) be a Cauchy sequence in $\mathbb{C}^n$. Write $z_k = (x_{k,j} + i y_{k,j})_{j=1}^n$ with $x_{k,j}, y_{k,j} \in \mathbb{R}$. Then

$$|x_{k,j} - x_{l,j}| \le d(z_k, z_l) \to 0 \text{ as } k, l \to \infty.$$

Thus $(x_{k,j})_{j=1}^n$ is a Cauchy sequence in $\mathbb{R}$ for each $j = 1, \dots, n$. By Theorem 2.12, the sequence converges, so that there is $a_j \in \mathbb{R}$ such that

$$\lim_{k\to\infty} x_{k,j} = a_j.$$

Similarly, there is $b_j \in \mathbb{R}$ such that

$$\lim_{k\to\infty} y_{k,j} = b_j.$$

Put

$$z = (a_j + i b_j)_{j=1}^n.$$

Let $\varepsilon > 0$. Then there is $K \in \mathbb{R}$ such that for all $k \ge K$ and all $j = 1, \dots, n$,

$$|x_{k,j} - a_j| < \frac{\varepsilon}{\sqrt{2n}}, \quad |y_{k,j} - b_j| < \frac{\varepsilon}{\sqrt{2n}}.$$

Hence, for $k \geq K$,

$$
\begin{aligned}
d(z_k, z) &= \left(\sum_{j=1}^{n} \left(|x_{k,j} - a_j|^2 + |y_{k,j} - b_j|^2 \right) \right)^{\frac{1}{2}} \\
&< \left(\sum_{j=1}^{n} \left(\frac{\varepsilon^2}{2n} + \frac{\varepsilon^2}{2n} \right) \right)^{\frac{1}{2}} \\
&= \varepsilon.
\end{aligned}
$$

□

Theorem 7.4. *Let Y be a nonempty set. Then $B(Y) = B(Y, \mathbb{F}^n)$ is complete.*

Proof. Let (f_k) be a Cauchy sequence in $B(Y)$. For each $x \in Y$,

$$\|f_k(x) - f_m(x)\| \leq d(f_k, f_m) \to 0 \text{ as } k, m \to 0.$$

Hence $(f_k(x))_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{F}^n$. Since $\mathbb{F}^n$ is complete, there is $f(x) \in \mathbb{F}^n$ such that $f_k(x) \to f(x)$ as $n \to \infty$.
First we will show that f is bounded. Indeed, for each $\varepsilon > 0$ there is $n_\varepsilon \in \mathbb{N}$ such that for all $k, m \geq n_\varepsilon$, $d(f_k, f_m) < \varepsilon$. Also, for all $x \in Y$, there is $m \geq n_1$ such that

$$\|f_m(x) - f(x)\| < 1.$$

Hence

$$
\begin{aligned}
\|f(x)\| &\leq \|f(x) - f_m(x)\| + \|f_m(x) - f_{n_1}(x)\| + \|f_{n_1}(x)\| \\
&< \|f_{n_1}(x)\| + 2.
\end{aligned}
$$

Hence

$$\sup_{x \in Y} \|f(x)\| \leq \sup_{x \in Y} \|f_{n_1}(x)\| + 2 < \infty.$$

Therefore f is bounded.
Finally, for $x \in Y$ choose $m \geq n_{\frac{\varepsilon}{3}}$ such that

$$\|f_m(x) - f(x)\| < \frac{\varepsilon}{4}.$$

Then, for $k \geq n_{\frac{\varepsilon}{4}}$,

$$
\begin{aligned}
\|f_k(x) - f(x)\| &\leq \|f_k(x) - f_m(x)\| + \|f_m(x) - f(x)\| \\
&< \|f_k(x) - f_m(x)\| + \frac{\varepsilon}{4} \\
&\leq d(f_k, f_m) + \frac{\varepsilon}{4} \\
&< \frac{1}{2}\varepsilon.
\end{aligned}
$$

Hence, for $k \geq n_{\frac{\varepsilon}{4}}$,

$$
\begin{aligned}
d(f_k, f) &= \sup_{x \in Y} \|f_k(x) - f(x)\| \\
&\leq \frac{1}{2}\varepsilon \\
&< \varepsilon.
\end{aligned}
$$

This proves that (f_k) converges in $B(Y)$. Since (f_k) was an arbitrary Cauchy sequence in $B(Y)$, it follows that $B(Y)$ is complete. □

Definition 7.4. 1. A subset A of a metric space (X, d) is called **open** if for each $x \in A$ there is $\varepsilon > 0$ such that $B(x, \varepsilon) \subset A$, where $B(x, \varepsilon) = \{y \in X : d(y, x) < \varepsilon\}$ is the ε-ball with centre x.
2. A subset A of a metric space (X, d) is called **closed** if $X \setminus A$ is open.

Theorem 7.5. *A subset A of a metric space (X, d) is closed if and only if for each sequence (x_k) in A which converges to some $x \in X$ it follows that $x \in A$.*

Proof. Let A be closed and let (x_k) be a sequence in A which converges to some $x \in X$. We must show $x \in A$. Assume that $x \notin A$. Since $x \in X \setminus A$ and $X \setminus A$ is open, there is $\varepsilon > 0$ such that $B(x, \varepsilon) \subset X \setminus A$. Since $x_k \in A$ for all $k \in \mathbb{N}$, it follows that $d(x_k, x) \geq \varepsilon$, which contradicts $x_k \to x$.
Conversely, if $A \subset X$ is not closed, then there is $x \in X \setminus A$ such that $B(x, \varepsilon) \not\subset X \setminus A$ for all $\varepsilon > 0$. Therefore, for positive integers k, there is $x_k \in A \cap B(x, \frac{1}{k})$. Thus, for each $\varepsilon > 0$ and $k \geq \frac{1}{\varepsilon}$ we have $d(x_k, x) < \frac{1}{k} \leq \varepsilon$. Thus we have found a sequence (x_k) in A which converges to some $x \in X$ which does not belong to A. □

Theorem 7.6. *Let (X, d) be a complete metric space and let $A \subset X$. Then A is complete if and only if A is closed.*

Proof. Assume that A is complete and let (x_k) be a sequence in A which converges to some $x \in X$. We must show that $x \in A$. But since (x_k) is a convergent sequence in X, it is a Cauchy sequence in X by Theorem 7.2, 2, and therefore also in A. Since A is complete, (x_k) converges to some $a \in A$, and therefore also in X. Since limits are unique by Theorem 7.2, 1, it follows that $x = a \in A$. Hence, by Theorem 7.5, A is closed.
Conversely, assume that A is closed. Let (x_k) be a Cauchy sequence in A. Then (x_k) is a Cauchy sequence in X, and since X is complete, there is $x \in X$ such that $x_k \to x$. Since A is closed, it follows that $x \in A$. Hence each Cauchy sequence in A converges in A. It follows that A is complete. □

Definition 7.5. Let $A \subset \mathbb{R}$ and $f : A \to \mathbb{F}^n$. Then f is **uniformly continuous** if for each $\varepsilon > 0$ there is $\delta > 0$ such that for all $x, y \in A$ $(|x - y| < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon)$.

Remark 7.2.1. It is straightforward to show that a function $f : [a, b] \to \mathbb{C}$ is (uniformly) continuous if and only if its real and imaginary part are (uniformly) continuous and that a function $f : [a, b] \to \mathbb{F}^n$ is (uniformly) continuous if and only if all its component are (uniformly) continuous.

Theorem 7.7. *Let $a, b \in \mathbb{R}$, $a < b$, and let $f : [a, b] \to \mathbb{F}^n$ be continuous. Then f is uniformly continuous.*

Proof. Because of Remark 7.2.1 we only have to consider $f : [a, b] \to \mathbb{R}$. Let $\varepsilon > 0$ and choose $a = x_0 < x_1 < \cdots < x_N = b$ according to Lemma 6.13 with respect to $\frac{\varepsilon}{4}$. Put

$$\delta = \min\{|x_j - x_{j-1}| : j = 1, \ldots, N\}.$$

Now let $x, y \in [a, b]$ such that $|x - y| < \delta$. We want to show $|f(x) - f(y)| < \varepsilon$. This is trivial if $x = y$. Thus let $x \neq y$, and we may assume $x < y$.
There is $j \in \{1, \ldots, N\}$ such that $x \in [x_{j-1}, x_j]$. Hence

$$x_{j-1} \leq x < y < x + \delta \leq x_j + \delta.$$

Case I: $y \leq x_j$. Then $y \in [x_{j-1}, x_j]$, and Lemma 6.13 gives

$$|f(x) - f(y)| \leq |f(x) - f(x_{j-1})| + |f(x_{j-1}) - f(y)| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon.$$

Case II: $y > x_j$. Then $j < N$, and $x_j < y < x_j + \delta \leq x_{j+1}$. Again by Lemma 6.13,

$$|f(x) - f(y)| \leq |f(x) - f(x_{j-1})| + |f(x_{j-1}) - f(x_j)| + |f(x_j) - f(y)| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon.$$ □

For an interval I we denote by $C(I, \mathbb{F}^n)$ the set of continuous functions from I to $\mathbb{F}^n$.

Theorem 7.8. *Let $a, b \in \mathbb{R}$ and $n \in \mathbb{N}\setminus\{0\}$. Then $C([a, b], \mathbb{F}^n)$ is a closed subspaces of $B([a, b], \mathbb{F}^n)$. In particular, $C([a, b], \mathbb{F}^n)$ is complete.*

Proof. By Remark 7.2.1, it suffices to consider $C([a, b], \mathbb{R})$. By Theorem 3.17, each continuous function on $[a, b]$ is bounded, so that $C([a, b], \mathbb{R}) \subset B([a, b], \mathbb{R})$.

To show that $C([a,b],\mathbb{R})$ is closed in $B([a,b],\mathbb{R})$ let (f_k) be a sequence in $C([a,b],\mathbb{R})$ which converges to some f in $B([a,b],\mathbb{R})$. We must show that f is continuous.
Let $\varepsilon > 0$. Then there is $k \in \mathbb{N}$ such that $d(f_k, f) < \frac{\varepsilon}{3}$. Since f_k is uniformly continuous, there is $\delta > 0$ such that

$$|f_k(x) - f_k(y)| < \frac{\varepsilon}{3} \text{ for } x, y \in [a,b], \ |x-y| < \delta.$$

Then, for $x, y \in [a,b]$ and $|x-y| < \delta$,

$$\begin{aligned} |f(x) - f(y)| &\le |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &\le d(f, f_k) + |f_k(x) - f_k(y)| + d(f_k, f) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned}$$

Thus f is uniformly continuous.
By Theorem 7.6, $C([a,b],\mathbb{F}^n)$ is complete. □

Tutorial 7.2.1. 1. Let (X,d) be the metric space from Tutorial 7.1.1, 1. Show that (X,d) is complete if and only if range f is closed in $\mathbb{R}$.
2. Show that the metric space $(\overline{\mathbb{R}}, d)$ from Tutorial 7.1.1, 2, is complete.
3. Show that $\mathbb{R}$ is not complete with respect to the metric given in Tutorial 7.1.1, 2.

7.3 Banach's Fixed Point Theorem

In this section, $X = (X, d_X)$, $Y = (Y, d_Y)$ and $Z = (Z, d_Z)$ are metric spaces.

Definition 7.6. Let $T : X \to Y$. T is called **bounded** if there is a number $M \ge 0$ such that
$d_Y(Tx, Ty) \le M d_X(x,y)$ for all $x, y \in X$.
If T is bounded, then

$$\|T\| = \inf\{M \ge 0 : \forall\, x, y \in X,\ d_Y(Tx, Ty) \le M d_X(x,y)\}$$

is called the **norm** of T.
If T is not bounded, we may write for convenience $\|T\| = \infty$.

Since it is clear that $d_Y(Tx, Ty) \le (\|T\| + \varepsilon) d_X(x,y)$ for all $\varepsilon > 0$, it follows that

$$\forall\, x, y \in X,\ d_Y(Tx, Ty) \le \|T\| d_X(x,y).$$

Proposition 7.9. *Let $T : X \to Y$. Then*

$$\|T\| = \sup_{\substack{x,y\in X \\ x\ne y}} \frac{d_Y(Tx,Ty)}{d_X(x,y)}.$$

Proof. Let

$$A = \left\{ \frac{d_Y(Tx,Ty)}{d_X(x,y)} : x, y \in X,\ x \ne y \right\},\ \alpha = \sup A.$$

If T is bounded, then $\|T\|$ is an upper bound of A and thus $\alpha \le \|T\|$.
On the other hand, $\alpha \ge \dfrac{d_Y(Tx,Ty)}{d_X(x,y)}$ for all $x, y \in X$ with $x \ne y$, and thus $d_Y(Tx,Ty) \le \alpha d_X(x,y)$ for all $x, y \in X$.
Hence $\alpha \ge \|T\|$.
The case that T is not bounded is left as an exercise. □

Notation. 1. If $T : X \to Y$ and $S : Y \to Z$, then we write $ST = S \circ T$.
2. If $T : X \to X$, we define
T^0 by $T^0 x = x$, $x \in X$,
$T^{k+1} = TT^k$ inductively for $k \in \mathbb{N}$.

Proposition 7.10. *If $T : X \to Y$ and $S : Y \to Z$ are bounded, then ST is bounded, and $\|ST\| \le \|S\|\,\|T\|$.*

Proof. For all $x, y \in X$ we have

$$d_Z((ST)x, (ST)y) = d_Z(S(Tx), S(Ty)) \le \|S\| d_Y(Tx, Ty) \le \|S\|\,\|T\| d_X(x, y).$$

Hence, by definition, ST is bounded and $\|ST\| \le \|S\|\,\|T\|$. □

Definition 7.7. Let $T : X \to X$ be bounded.
1. T is called a **contraction** if $\|T\| < 1$.
2. T is called a **generalized contraction** if there is $m \in \mathbb{N}$ such that $\|T^m\| < 1$.

Remark 7.3.1. Let $T : X \to X$ be bounded. Then T is a generalized contraction if and only if

$$\sum_{j=0}^{\infty} \|T^j\| < \infty.$$

Proof. If the sum converges, then $\|T^j\| \to 0$ as $j \to \infty$, and therefore there is $m \in \mathbb{N}$ such that $\|T^m\| < 1$. Conversely, if $\|T^m\| < 1$, then we group the sum into blocks of m terms, writing every natural number as a multiple of m plus its remainder, use

$$\|T^{km+l}\| = \|T^{km}T^l\| \le \|(T^m)^k\|\,\|T^l\| \le \|T^m\|^k \|T^l\|$$

and get

$$\begin{aligned}
\sum_{j=0}^{\infty} \|T^j\| &= \sum_{k=0}^{\infty} \sum_{l=0}^{m-1} \|T^{km+l}\| \\
&\le \sum_{k=0}^{\infty} \sum_{l=0}^{m-1} \|T^m\|^k \|T^l\| \\
&= \left(\sum_{k=0}^{\infty} \|T^m\|^k\right)\left(\sum_{l=0}^{m-1} \|T^l\|\right) \\
&= \frac{1}{1 - \|T^m\|} \sum_{l=0}^{m-1} \|T^l\| < \infty.
\end{aligned}$$

□

Theorem 7.11 (Banach's Fixed Point Theorem)**.** *Let (X, d) be a complete metric space and let $T : X \to X$ be a generalized contraction. Then there is a unique $a \in X$ such that $Ta = a$.*

Proof. Since T is a generalized contraction, there is $m \in \mathbb{N}$ such that $\|T^m\| < 1$.
For uniqueness, let $a, b \in X$ such that $Ta = a$ and $Tb = b$. By induction

$$T^k a = TT^{k-1}a = Ta = a \text{ and } T^k b = b,$$

so that

$$d(a, b) = d(T^k a, T^k b) \le \|T^k\| d(a, b),$$

and $d(a, b) \ne 0$ would give the contradiction $1 \le \|T^m\|$. Hence $d(a, b) = 0$, i. e., $a = b$.
For existence let $x_0 \in X$ and consider the sequence $(T^k x_0)$. Then, for $j, k \in \mathbb{N}$, $k > j$,

$$\begin{aligned}
d(T^k x_0, T^j x_0) &\le d(T^k x_0, T^{k-1} x_0) + d(T^{k-1} x_0, T^{k-2} x_0) + \cdots + d(T^{j+1} x_0, T^j x_0) \\
&\le \|T^{k-1}\|\, d(Tx_0, x_0) + \|T^{k-2}\|\, d(Tx_0, x_0) + \cdots + \|T^j\|\, d(Tx_0, x_0) \\
&= \sum_{l=j}^{k-1} \|T^l\|\, d(Tx_0, x_0)
\end{aligned}$$

Since

$$\sum_{l=0}^{\infty} \|T^l\| < \infty,$$

it follows that $(T^k x_0)$ is a Cauchy sequence. Since X is complete, this Cauchy sequence has a limit, say a. Finally, to show that $Ta = a$, let $\varepsilon > 0$. Then there is K such that for $j \geq K$,

$$d(T^j x_0, a) < \frac{\varepsilon}{1 + \|T\|}.$$

For $k \geq K + 1$ we get

$$\begin{aligned} d(Ta, a) &\leq d(Ta, T^k x_0) + d(T^k x_0, a) \\ &\leq \|T\| d(a, T^{k-1} x_0) + d(T^k x_0, a) \\ &< \|T\| \frac{\varepsilon}{1 + \|T\|} + \frac{\varepsilon}{1 + \|T\|} \\ &= \varepsilon. \end{aligned}$$

Hence $d(Ta, a) = 0$, i. e., $Ta = a$. □

Tutorial 7.3.1. Let (X, d) be a complete metric space and let T be a generalized contraction on X. Let $m \in \mathbb{N}^*$ such that $\|T^m\| < 1$ and let x be the fixed point of T. Prove the following a priori estimate: For any $x_0 \in X$,

$$d(x, x_0) \leq \frac{d(Tx_0, x_0)}{1 - \|T^m\|} \sum_{j=0}^{m-1} \|T^j\|.$$

Hint. Use that $T^m x = x$ and use a "telescoping" sum of distances involving the points $T^m x$, $T^m x_0$, $T^{m-1} x_0$, …, Tx_0.

Tutorial 7.3.2. Newton Raphson method of finding zeros
Let $f : [a, b] \to \mathbb{R}$ be twice continuously differentiable and assume there are $c, d \in [a, b]$, $c < d$, such that $f(c)$ and $f(d)$ have opposite signs. By the Intermediate Value Theorem, f has a zero y in (c, d). Also assume that $f'(x) \neq 0$ for all $x \in [c, d]$. Since f' is continuous, f' is either positive or negative by the Internmediate Value Theorem, so that f is strictly monotonic by the Mean Value Theorem. Hence the zero is unique. To find this zero one can use the Newton Raphson method:
1. Let

$$Tx = x - \frac{f(x)}{f'(x)}, \quad x \in [c, d].$$

Show that for all $x_1, x_2 \in [c, d]$ with $x_1 \neq x_2$ there is ξ between x_1 and x_2 such that

$$Tx_1 - Tx_2 = \frac{f(\xi) f''(\xi)}{(f'(\xi))^2} (x_1 - x_2).$$

2. Show that $\|T\| < 1$ if c and d are sufficiently close to the zero y of f.
3. Show that one can choose c and d in such a way that $T([c, d]) \subset [c, d]$.
Hint. You may show that $T([c, d]) \subset [c, d]$ if $d - y = y - c$, i. e., y is the midpoint of the interval $[c, d]$. However, this is not very practical since you do not know y (otherwise, the whole procedure would be pointless). The easiest way out of this is to choose suitable values (by trial and error), $c_1 < d_1$ in $[c, d]$ such that $f(c_1)$ and $f(c_2)$ have opposite sign, $Tc_1, Tc_2 \in [c_1, c_2]$, and ff'' has no zeros on $[c_1, d_1] \setminus \{y\}$.
*4. Assume that statement 2 holds and define

$$c_1 = \frac{c + \|T\| d}{1 + \|T\|}, \quad d_1 = \frac{d + \|T\| c}{1 + \|T\|}.$$

Show that $c < c_1 < y < d_1 < d$ if $f(c_1)$ and $f(d_1)$ have opposite signs. Then show that $T([c, d]) \subset [c, d]$. Hence show that for any $x \in [c, d]$, $T^k x \to y$ as $k \to \infty$.

Tutorial 7.3.3. Let $c > 0$. Show that $\sqrt{c + \sqrt{c + \sqrt{c + \dots}}}$ converges by using the following proof.
(a) Let $T_c x = \sqrt{c + x}$, $x \in [0, \infty)$. Show that $T_c x > x$ for $x \in [0, 1]$ and that $T_c : \left[\frac{1}{4}, \infty\right) \to \left[\frac{1}{4}, \infty\right)$ is a contraction.
Show that if $x_k \leq 1$ for all $k \in \mathbb{N}$, then (x_k) is (strictly) increasing. (Actually, this never happens.)
(b) Let $x_0 = c$, $x_{k+1} = T_c x_k$ $(k \in \mathbb{N})$.
(c) Show that (x_k) converges as $k \to \infty$.
(d) Find $\lim\limits_{k\to\infty} x_k$.

7.4 Existence of solutions of differential equations

Let $I \subset \mathbb{R}$ be an interval, let $U \subset \mathbb{F}^n$ and let $f : I \times U \to \mathbb{F}^n$ be continuous. Then

$$y' = f(x, y) \tag{7.4.1}$$

is called a **first-order system of differential equations**, and any continuously differentiable function $y : I \to U$ satisfying (7.4.1) is called a solution of (7.4.1).
Let $a \in I$ and $b \in \mathbb{F}^n$. Then a solution of (7.4.1) satisfying

$$y(a) = b \tag{7.4.2}$$

is called a solution of the **initial value problem** (IVP) (7.4.1), (7.4.2).
From the Fundamental Theorem of Calculus, Theorem 6.15, we obtain

Lemma 7.12. *y is a a solution of the IVP (7.4.1), (7.4.2) if and only if $y \in C(I, \mathbb{F}^n)$, $y(I) \subset U$, and*

$$y(x) = b + \int_a^x f(t, y(t))\, dt \quad (x \in I). \tag{7.4.3}$$

Proof. We only have to note that we can apply the Fundamental Theorem of Calculus componentwise and to real and imaginary parts separately. Also, if $y \in C(I, \mathbb{F}^n)$ satisfies (7.4.3), then f is differentiable. $\square$

We are now going to show that under certain conditions, (7.4.3) and thus the IVP (7.4.1), (7.4.2) has a unique solution. We write $C^1(I, \mathbb{F}^n)$ for the set of all continuously differentiable functions from I to $\mathbb{F}^n$.

Theorem 7.13. *We make the following assumptions:*

(i) *Let I be a closed bounded interval, $a \in I$, $A = \max_{x \in I} |x - a|$, $b \in \mathbb{F}^n$, $0 < B \le \infty$,*
$U = \{y \in \mathbb{F}^n : \|y - b\| \le B\}$, $R = I \times U$.

(ii) *Let $f : R \to \mathbb{F}^n$ be continuous and let $M = \sup_{(x,y) \in R} \|f(x, y)\|$.*

(iii) $\exists\, L > 0\ \forall\, (x, y_1), (x, y_2) \in R\ \|f(x, y_1) - f(x, y_2)\| \le L\|y_1 - y_2\|$.
(This is called a Lipschitz condition.)

(iv) $AM \le B$.

Then there is exactly one $y \in C^1(I, \mathbb{F}^n)$ such that

$$
\begin{aligned}
&y(a) = b,\\
&(x, y(x)) \in R \quad (x \in I),\\
&y'(x) = f(x, y(x)) \qquad (x \in I).
\end{aligned}
$$

Proof. Since U is a closed subset of $\mathbb{F}^n$, $C(I, U)$ is a complete metric space. For $y \in C(I, U)$ define

$$
(Ty)(x) = b + \int_a^x f(t, y(t))\, dt \quad (x \in I). \tag{7.4.4}
$$

By Lemma 7.12, the theorem is proved if we show that T maps $C(I, U)$ into $C(I, U)$ and that T has exactly one fixed point.

Thus let $y \in C(I, U)$ and $x \in I$. Clearly, $Ty \in C(I, \mathbb{F}^n)$ and

$$
\begin{aligned}
\|(Ty)(x) - b\| &= \left\| \int_a^x f(t, y(t))\, dt \right\| \\
&\le \left| \int_a^x \|f(t, y(t))\|\ dt \right| \\
&\le \left| \int_a^x M\, dt \right| \\
&= |x - a| M \\
&\le AM \\
&\le B.
\end{aligned}
$$

Thus we have shown that $Ty \in C(I, U)$.

Next we are going to show that T is a generalized contraction by proving the following estimate: For all $n \in \mathbb{N}$, $y_1, y_2 \in C(I, U)$ and $x \in I$,

$$
\left\| (T^k y_1)(x) - (T^k y_2)(x) \right\| \le \frac{L^k |x - a|^k}{k!} d(y_1, y_2). \tag{7.4.5}
$$

The induction base $k = 0$ is trivial since

$$
\|y_1(x) - y_2(x)\| \le \sup_{t \in I} \|y_1(t) - y_2(t)\| = d(y_1, y_2).
$$

For the induction step assume that (7.4.5) is true for k. Then

$$\begin{aligned}
\left\|(T^{k+1}y_1)(x) - (T^{k+1}y_2)(x)\right\| &= \left\|\int_a^x \left[f(t,(T^k y_1)(t)) - f(t,(T^k y_2)(t))\right]\,dt\right\| \\
&\le \left|\int_a^x \left\|f(t,(T^k y_1)(t)) - f(t,(T^k y_2)(t))\right\|\,dt\right| \\
&\le \left|\int_a^x L\left\|(T^k y_1)(t) - (T^k y_2)(t)\right\|\,dt\right| \quad \text{(by assumption (iii))} \\
&\le \left|\int_a^x L\frac{L^k|x-a|^k}{k!}\,dt\right| d(y_1,y_2) \quad \text{(by induction hypothesis)} \\
&= \frac{L^{k+1}|x-a|^{k+1}}{(k+1)!}d(y_1,y_2).
\end{aligned}$$

Taking the maximum over all $x \in I$ in (7.4.5) gives

$$d(T^k y_1, T^k y_2) \le \frac{L^k A^k}{k!} d(y_1, y_2),$$

and therefore

$$\|T^k\| \le \frac{L^k A^k}{k!},$$

which leads to

$$\sum_{k=0}^{\infty} \|T^k\| \le \sum_{k=0}^{\infty} \frac{(LA)^k}{k!} = e^{LA} < \infty.$$

Hence T is a generalized contraction on $C(I, U)$. An application of Banach's Fixed Point Theorem completes the proof. □

Worked Example 7.4.1. Consider the initial value problem $y' = 1 + y^2$, $y(0) = 0$.
a) Find the largest $c > 0$ such that the assumptions of Theorem 7.12 are satisfied for $I = [0, c]$.
b) Find a maximal interval on which a solution exists.

Solution. a) We have $f(x, y) = 1 + y^2$. Hence

$$f(x, y_1) - f(x, y_2) = y_1^2 - y_2^2 = (y_1 + y_2)(y_1 - y_2),$$

so that

$$|f(x, y_1) - f(x, y_2)| = |y_1 + y_2|\,|y_1 - y_2|.$$

Hence the Lipschitz condition is satisfied if and only if

$$\sup_{|y_1|,|y_2| \le B} |y_1 + y_2| < \infty,$$

which is satisfied if and only if $B \in \mathbb{R}$, i. e., $B < \infty$.
We have $A = c$ and

$$M = \sup_{|y| \le B} |1 + y^2| = 1 + B^2.$$

We must satisfy $AM \le B$, i. e.,

$$c(1 + B^2) \le B \quad \text{or} \quad c \le \frac{B}{1 + B^2}.$$

But

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

has a maximum of 1 for $B = 1$, so that the optimal c is $c = \dfrac{1}{2}$.

b) The solution of the IVP is $y(x) = \tan x$. So the maximal interval where a solution exists is $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$.

Let $M_n(\mathbb{F})$ be the set of $n \times n$ matrices with values in $\mathbb{F}$. Then the norm $\|M\|$ of $M \in M_n(\mathbb{F})$ is the norm of $M : \mathbb{F}^n \to \mathbb{F}^n$. It is easy to see that

$$\max_{i,j=1}^{n} |m_{ij}| \le \|M\| \le \sum_{i,j=1}^{n} |m_{ij}|, \quad \text{where } M = (m_{ij})_{i,j=1}^{n}.$$

Corollary 7.14. *Let I be an interval, $a \in I$, $b \in \mathbb{F}^n$, and consider the linear system of differential equations*

$$y' = F(x)y + g(x),$$

where $F : I \to M_n(\mathbb{F})$ and $g : I \to \mathbb{F}^n$ are continuous.
Then the initial value problem

$$y' = F(x)y + g(x),$$

has a unique solution $y \in C^1(I, \mathbb{F}^n)$.

Proof. First assume that $I = [c, d]$. Then, for $y_1, y_2 \in \mathbb{F}^n$, $x \in I$,

$$\begin{aligned} \|(F(x)y_1 + g(x)) - (F(x)y_2 - g(x))\| &\le \|F(x)\| \, \|y_1 - y_2\| \\ &\le \left(\max_{x \in I} \|F(x)\|\right) \|y_1 - y_2\|. \end{aligned}$$

Hence the Lipschitz condition is satisfied with $B = \infty$. Now apply Theorem 7.13.
If I is arbitrary, we can choose a decreasing sequence (c_k) and an increasing sequence (d_k) such that $c_1 \le a \le d_1$ and

$$I = \bigcup_{k=1}^{\infty} [c_k, d_k].$$

[If, e. g., the right endpoint d of I belongs to I, choose $d_k = d$.]
By the first part of the proof, the IVP has a unique solution y_k on $[c_k, d_k]$. Therefore, $y_{k+m}(x) = y_k(x)$ for all $m > 0$ and $x \in [c_k, d_k]$, so that $y(x) = y_k(x)$ for $x \in [c_k, d_k]$ is well-defined and solves the IVP. Because this solution is unique on $[c_k, d_k]$ for each k, y is unique. □

For $\eta^{(n)} = f(x, \eta, \eta', \dots, \eta^{(n-1)})$ define

$$y = \begin{pmatrix} \eta \\ \eta' \\ \vdots \\ \eta^{(n-1)} \end{pmatrix}, \qquad F(x, y) = \begin{pmatrix} y_2 \\ \vdots \\ y_n \\ f(x, y) \end{pmatrix}.$$

Then

$$\eta^{(n)} = f(x, \eta, \eta', \dots, \eta^{(n-1)}) \Leftrightarrow y' = F(x, y).$$

Let $b_0, \dots, b_{n-1} \in \mathbb{F}$. Then

$$\eta^{(j)}(a) = b_j \ (j = 0, \dots, n-1) \Leftrightarrow y(a) = b,$$

where

$$b = \begin{pmatrix} b_0 \\ \vdots \\ b_{n-1} \end{pmatrix}.$$

Clearly,

$$\|F(x, y) - F(x, z)\| \le (n-1)\|y - z\| + |f(x, y) - f(x, z)|$$

and

$$|f(x, y) - f(x, z)| \le \|F(x, y) - F(x, z)\|,$$

so that F satisfies a Lipschitz condition if and only if f satisfies a Lipschitz condition.

Corollary 7.15. *Let I be a bounded closed interval, $a \in I$, $b_0, \dots, b_{n-1} \in \mathbb{F}$, $0 < B \leq \infty$. Let $U = \{y \in \mathbb{F}^n : \|y - b\| \leq B\}$, $R = I \times U$. Let $f : R \to \mathbb{F}$ be continuous and assume there is $L > 0$ such that $|f(x,y) - f(x,z)| \leq L\|y - z\|$ for all $x \in I$ and $y, z \in U$. Then there is $A > 0$ such that the initial value problem*

$$\eta^{(n)} = f(x, \eta, \eta', \dots, \eta^{(n-1)}), \qquad \eta^{(j)}(a) = b_j \ (j = 0, \dots, n-1)$$

has a unique solution on $I \cap [a - A, a + A]$.

Note. According to Theorem 7.13 one uses $AM \leq B$ where M is the supremum of F on R. Finally, a combination of Corollaries 7.14 and 7.15 gives

Corollary 7.16. *Let I be an interval, $a \in I$, $b_0, \dots, b_{n-1} \in \mathbb{F}$, and consider the linear differential equation*

$$\eta^{(n)} = f_{n-1}(x)\eta^{(n-1)} + f_{n-2}(x)y^{(n-2)} + \dots f_1(x)y' + f_0(x)y + g(x),$$

with initial conditions

$$y^{(j)}(a) = b_j, \quad (j = 0, \dots, n-1),$$

where $f_j : I \to \mathbb{F}$ $(j = 0, \dots, n-1)$ and $g : I \to \mathbb{F}$ are continuous. Then the initial value problem has a unique solution $\eta \in C^n(I, \mathbb{F})$.

Tutorial 7.4.1. 1. Let $n \in \mathbb{N}^*$ and $M \in M_n(\mathbb{F})$. Prove that

$$\max_{i,j=1}^{n} |m_{ij}| \leq \|M\| \leq \sum_{i,j=1}^{n} |m_{ij}|, \quad \text{where } M = (m_{ij})_{i,j=1}^{n}.$$

2. Let $M_1, M_2 \in M_n(\mathbb{F})$.
a) Prove that $\|M_1 + M_2\| \leq \|M_1\| + \|M_2\|$.
b) Let $d(M_1, M_2) = \|M_1 - M_2\|$. Prove that d is a metric on $M_n(\mathbb{F})$.
3. Prove that $M_n(\mathbb{F})$ is complete.

7.5 Power Series and the Exponential Function in $M_n(\mathbb{F})$

In the following let B be either $\mathbb{F}^n$ or $M_n(\mathbb{F})$. We know that B is complete, see Theorem 7.3 and Tutorial 7.4.1, 3.

Definition 7.8. Let (a_j) be a sequence in B. Then $\sum_{j=0}^{\infty} a_j$ is called a series in B.

The series is said to converge in B if the sequence of its partial sums $\sum_{j=0}^{k} a_j$ converges in B as $k \to \infty$.

The series is said to converge absolutely if $\sum_{j=0}^{\infty} \|a_j\| < \infty$.

Since B is complete, it follows as in Theorems 5.5 and 5.8:

Theorem 7.17. *Let $\sum_{j=0}^{\infty} a_j$ be a series in B.*

1. The series $\sum_{j=0}^{\infty} a_j$ converges in B if and only if for all $\varepsilon > 0$ there is $K \in \mathbb{R}$ such that for all $k \geq m \geq K$,

$$\left\| \sum_{j=m}^{k} a_j \right\| < \infty.$$

2. If $\sum_{j=0}^{\infty} a_j$ converges absolutely, then it converges in B.

Theorem 7.18. *Let* $\sum_{j=0}^{\infty} a_j$ *be a series in* B*, assume that* $\rho = \limsup_{j\to\infty} \sqrt[j]{\|a_j\|} < \infty$ *and put* $R = \frac{1}{\rho}$*. Let* $a, x \in \mathbb{R}$*. Then*
1. The series

$$\sum_{j=0}^{\infty} a_j (x-a)^j$$

converges absolutely if $|x-a| < R$*, and it diverges if* $|x-a| > R$ *in case* $R < \infty$*.*
2. Let $I = \mathbb{R}$ *if* $R = \infty$ *and* $I = (a-R, a+R)$ *if* $R \in \mathbb{R}$*. The series* $\sum_{j=0}^{\infty} a_j (x-a)^j$ *is differentiable for* $x \in I$*, and the derivative is represented by the absolutely convergent series*

$$\sum_{j=1}^{\infty} j a_j (x-a)^{j-1}$$

Proof. The proof of part 1 is similar to that of Theorem **??**.
For the proof of part 2 define for $x \in I$:

$$f(x) = \sum_{j=0}^{\infty} a_j (x-a)^j,$$
$$g(x) = \sum_{j=1}^{\infty} j a_j (x-a)^{j-1}.$$

It is straightforward to show that

$$\limsup_{j\to\infty} \sqrt[j]{\|(j+1)a_{j+1}\|} = \limsup_{j\to\infty} \sqrt[j]{\|j a_j\|} = \lim_{j\to\infty} \sqrt[j]{\|j\|} \limsup_{j\to\infty} \sqrt[j]{\|a_j\|} = \limsup_{j\to\infty} \sqrt[j]{\|a_j\|} = \rho,$$

so that g is an absolutely convergent series.
Now let $x \in I$ and choose $h_0 > 0$ such that $[x-h_0, x+h_0] \subset I$. Then $r = \max\{|x-h_0-a|, |x+h_0-a|\} < R$. Let $h \in \mathbb{R}$ with $0 < |h| < h_0$. Since $f(x)$, $f(x+h)$ and $g(x)$ are represented by (absolutely) convergent series, we have

$$\begin{aligned} f(x+h) - f(x) - g(x)h &= \lim_{k\to\infty} \sum_{j=1}^{k} a_j \left[(x+h-a)^j - (x-a)^j - j(x-a)^{j-1}h\right] \\ &= \lim_{k\to\infty} \sum_{j=2}^{k} a_j \left[(x+h-a)^j - (x-a)^j - j(x-a)^{j-1}h\right]. \end{aligned}$$

By the Mean Value Theorem, for each $j \ge 2$ there is $h_j \in \mathbb{R}$ with $0 < |h_j| < |h|$ such that

$$(x+h-a)^j - (x-a)^j = j(x+h_j-a)^{j-1}h.$$

Another application of the MVT gives $\hat{h}_j \in \mathbb{R}$ with $0 < |\hat{h}_j| < |h_j|$ such that

$$\begin{aligned} (x+h-a)^j - (x-a)^j - j(x-a)^{j-1}h &= jh\left[(x+h_j-a)^{j-1} - (x-a)^{j-1}\right] \\ &= j(j-1)h(x+\hat{h}_j-a)^{j-2}h_j, \end{aligned}$$

so that

$$\begin{aligned} \left\| \sum_{j=2}^{k} a_j \left[(x+h-a)^j - (x-a)^j - j(x-a)^{j-1}h\right] \right\| &\le \sum_{j=2}^{k} j(j-1)|h|\,|h_j|\,|x+\hat{h}_j-a|^{j-2}\|a_j\| \\ &\le h^2 \sum_{j=2}^{k} j(j-1)r^{j-2}\|a_j\|. \end{aligned}$$

With a reasoning as above, we have

$$\limsup_{j\to\infty} \sqrt[j]{j(j-1)r^{j-2}\|a_j\|} = r\rho < 1.$$

By the Root Test,

$$c = \sum_{j=2}^{\infty} j(j-1)r^{j-2}\|a_j\| < \infty,$$

and therefore

$$\|f(x+h) - f(x) - g(x)h\| \le h^2 c,$$

which proves that f is differentiable at x with

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

□

Theorem 7.19. *Let $A \in M_n(\mathbb{F})$. Then:*
1. *The matrix exponential*

$$\exp(A) = \sum_{j=0}^{\infty} \frac{1}{j!} A^j$$

converges absolutely in $M_n(\mathbb{F})$.
2. $\exp(A)$ *is invertible and* $(\exp(A))^{-1} = \exp(-A)$.
3. *If $A, C \in M_n(\mathbb{F})$ commute, then*

$$\exp(A+C) = \exp(A)\exp(C).$$

4. $x \mapsto \exp(Ax)$ *is differentiable on* $\mathbb{R}$ *with*

$$\frac{d}{dx}\exp(Ax) = A\exp(Ax) = \exp(Ax)A.$$

Proof. 1. We calculate

$$\begin{aligned}\limsup_{j\to\infty} \sqrt[j]{\left\|\frac{1}{j!}A^j\right\|} &\le \limsup_{j\to\infty} \sqrt[j]{\frac{1}{j!}\|A\|^j} = \|A\| \limsup_{j\to\infty} \sqrt[j]{\frac{1}{j!}} \\ &= \|A\| \lim_{j\to\infty} \frac{j!}{(j+1)!} = \|A\| \lim_{j\to\infty} \frac{1}{j+1} = 0,\end{aligned}$$

and $\exp(Ax)$ converges for all $x \in \mathbb{R}$ by Theorem 7.18, 1.
4. The differentiability follows from Theorem 7.18, 2, with

$$\begin{aligned}\frac{d}{dx}\exp(Ax) &= \sum_{j=1}^{\infty} \frac{j}{j!} A^j x^{j-1} = \sum_{j=0} \frac{1}{j!} A^{j+1} x^j \\ &= A \sum_{j=0}^{\infty} \frac{1}{j!} A^j x^j = A\exp(A).\end{aligned}$$

Similarly,

$$\frac{d}{dx}\exp(Ax) = \left(\sum_{j=0}^{\infty} \frac{1}{j!} A^j x^j\right) A = \exp(A)A.$$

2 and 3. A direct proof is quite cumbersome. Instead, we consider the matrix function

$$F(x) = \exp((A+C)x)\exp(-Cx)\exp(-Ax)$$

and observe that $\exp(A+C)$, $\exp(A)$ and $\exp(C)$ commute with A and C since $A+C$, A and C commute with A and C. The product rule for derivatives and commutativity lead to

$$\begin{aligned}F'(x) &= \exp((A+C)x)(A+C)\exp(-Cx)\exp(-Ax) \\ &\quad - \exp((A+C)x)C\exp(-Cx)\exp(-Ax) - \exp((A+C)x)\exp(-Cx)A\exp(-Ax) = 0.\end{aligned}$$

Since $F(0) = (\exp(0))^3 = I_n^3 = I_n$, where I_n is the $n \times n$ identity matrix, and since $\frac{d}{dx} I_n = 0$, the uniqueness part of Corollary 7.16 gives $F(x) = I_n$ for all $x \in \mathbb{R}$.
Putting $C = 0$ and $x = 1$, we get

$$\exp(A)\exp(-A) = I_n = \exp(-A)\exp(A),$$

which proves 2. Finally, with 2,

$$\exp(A + C) = F(1)\exp(A)\exp(C) = \exp(A)\exp(C). \qquad \square$$

How can one find $\exp(Ax)$? Recall from Linear Algebra that $\chi(\lambda) = \det(A - \lambda I_n)$ is the characteristic function of A. Note that χ is a polynomial of degree n with leading coefficient $(-\lambda)^n$. In Complex Analysis you will learn about the Fundamental Theorem of Algebra which says for this case that there are n complex numbers $\lambda_1, \dots, \lambda_n$ such that

$$\chi(\lambda) = \prod_{j=1}^{n} (\lambda_j - \lambda).$$

Jordan's canonical form, see Linear Algebra, says that the λ_j can be sorted into k groups of equal number $\mu_1, \dots, \mu_k$, i. e.,

$$\chi(\lambda) = \prod_{j=1}^{k} (\mu_j - \lambda)^{p_j},$$

such there is an invertible $n \times n$ matrix T so that

$$A = T \left(\bigoplus_{j=1}^{k} D_j \right) T^{-1},$$

where

$$D_j = \mu_j I_{p_j} + J_{p_j}$$

and J_ν is the $\nu \times \nu$ matrix which has entries 1 just above the diagonal and 0 elsewhere. If we write, for $l \in \mathbb{N}$, $J_\nu^l = (m_{rs})_{r,s=1}^{\nu}$, then $m_{rs} = 1$ if $s - r = l$ and $m_{rs} = 0$ otherwise. In particular, $J_\nu^l = 0$ if $l \geq \nu$.
Then

$$A^m = T \left(\bigoplus_{j=1}^{k} D_j \right)^m T^{-1} = T \left(\bigoplus_{j=1}^{k} D_j^m \right) T^{-1},$$

which gives

$$\exp(Ax) = T \left(\bigoplus_{j=1}^{k} \exp(D_j x) \right) T^{-1}.$$

It remains to find $\exp(D_j x)$. For convenience, we drop the index and observe

$$(\mu I_\nu + J_\nu)^j = \sum_{l=0}^{j} \binom{j}{l} \mu^{j-l} J_\nu^l.$$

Using

$$\frac{1}{j!}\binom{j}{l} = \frac{j!}{j!\,l!\,(j-l)!} = \frac{1}{l!\,(j-l)!},$$

it follows that

$$\begin{aligned}
\exp(Dx) &= \sum_{j=0}^{\infty} \frac{1}{j!}(\mu I_\nu + J_\nu)^j x^j \\
&= \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{l=0}^{j} \binom{j}{l} \mu^{j-l} J_\nu^l x^j \\
&= \sum_{l=0}^{\nu-1} \sum_{j=l}^{\infty} \frac{1}{j!} \binom{j}{l} \mu^{j-l} J_\nu^l x^j \\
&= \sum_{l=0}^{\nu-1} \frac{1}{l!} \sum_{j=l}^{\infty} \frac{1}{(j-l)!} \mu^{j-l} J_\nu^l x^j \\
&= \sum_{l=0}^{\nu-1} \frac{1}{l!} \sum_{k=0}^{\infty} \frac{1}{k!} (\mu x)^k J_\nu^l x^l \\
&= \sum_{l=0}^{\nu-1} \frac{x^l}{l!} \exp(\mu x) J_\nu^l.
\end{aligned}$$

7.6 Solutions of Linear Differential Equations

In this section let I be an interval, $n \in \mathbb{N}^*$, and $F \in M_n(C(I))$.

Definition 7.9. A matrix function $Y \in M_n(C^1(I))$ is called a fundamental matrix of $y' = Fy$ if $y'(x) = F(x)y(x)$ and if $Y(x)$ is invertible for all $x \in I$.

Proposition 7.20. 1. *$y' = Fy$ has a fundamental matrix.*
2. *Let Y be a fundamental matrix of $y' = Fy$. Then $Z \in M_n(C^1(I))$ is a fundamental matrix of $y' = Fy$ if and only if there is $C \in M_n(\mathbb{F})$ such that $Z = YC$.*

Proof. 1. Let $c \in I$. Identifying $M_n(\mathbb{F})$ with $\mathbb{F}^{n^2}$, it follows from Corollary 7.14 that there is (a unique) $Y \in M_n(C^1(I))$ such that $Y' = FY$ and $Y(c) = I_n$, and there is $Z \in M_n(C^1(I))$ such that $Z' = -F^\top Z$ and $Z(c) = I_n$. Then

$$\begin{aligned}
\left(Z^\top Y\right)' &= -\left(Z^\top\right)' Y + Z^\top Y' \\
&= -Z^\top F^\top Y + Z^\top F^\top Y = 0
\end{aligned}$$

shows that $Z^\top Y$ is constant. So we have

$$Z(x)^\top Y(x) = Z(c)^\top Y(c) = I_n,$$

which implies that $Y(x)$ is invertible for all $x \in I$.
2. If $Z = YC$, then $Z(x) = Y(x)C$ is the product of invertible matrices and hence invertible for all $x \in I$, and

$$Z' = (YC)' = Y'C = FYC = FZ$$

shows that Z is a fundamental matrix of $y' = Fy$.
Conversely, if Z is a fundamental matrix of $y' = Fy$, then $Z(x)$ is invertible for all $x \in I$, and

$$\begin{aligned}
(Y^{-1}Z)' &= -Y^{-1}Y'Y^{-1}Z + Y^{-1}Z' \\
&= -Y^{-1}FYY^{-1}Z + Y^{-1}FZ = 0,
\end{aligned}$$

which implies that $Y^{-1}Z$ is constant. So there is $C \in M_n(\mathbb{F})$ such that $Y(x)^{-1}Z(x) = C$ for all $x \in I$. In particular, C is invertible since $Y(x)^{-1}$ and $Z(x)$ are invertible. Finally, multiplying by $Y(x)$ from the right, we get $Z(x) = Y(x)C$. $\square$

Theorem 7.21. *Assume that $F(x) = F(y)$ for all $x, y \in I$. Let $c \in I$ and define*

$$Y(x) = \exp\left(\int_c^x F(t)\,dt\right).$$

Then Y is a fundamental matrix of $y' = Fy$.

Proof. $Y(c) = \exp(0) = I_n$ is invertible. Since $F(x)$ commutes with $F(t)$ for all $x, t \in I$, it follows that

$$\frac{d}{dx}\left[\left(\int_c^x F(t)\,dt\right)^j\right] = jF(x)\left(\int_c^x F(t)\,dt\right)^{j-1},$$

so that

$$\begin{aligned}
Y'(x) &= \frac{d}{dx}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\int_c^x F(t)\,dt\right)^j \\
&= \sum_{j=0}^{\infty}\frac{1}{j!}\frac{d}{dx}\left[\left(\int_c^x F(t)\,dt\right)^j\right] \\
&= \sum_{j=1}^{\infty}\frac{j}{j!}F(x)\left(\int_c^x F(t)\,dt\right)^{j-1} \\
&= F(x)\exp\left(\int_c^x F(t)\,dt\right) \\
&= F(x)Y(x).
\end{aligned}$$

□

Theorem 7.22. *Let Y be a fundamental matrix of $y' = Fy$, $f \in C(I, \mathbb{F}^n)$, $c \in I$ and $d \in \mathbb{F}^n$. Then the unique solution of the initial value problem $y' = Fy + f$, $y(c) = d$ is*

$$y(x) = Y(x)Y^{-1}(c)d + Y(x)\int_c^x Y^{-1}(t)f(t)\,dt.$$

Proof. We could simply substitute y into the differential equation. However, we give a constructive proof, the so-called method of variation of the constants. It is easy to see that any solution of the homogeneous equation is $y(x) = Y(x)z$ with $z \in \mathbb{F}^n$. (we are not going to prove this separately because it is a special case of this theorem and therefore follows anyway). Now replace the constant vector z by a vector function on I, i. e., $y(x) = Y(x)z(x)$, where z is the solution of the initial value problem. Then

$$\begin{aligned}
y'(x) &= Y'(x)z(x) + Y(x)z'(x) \\
&= F(x)Y(x)z(x) + Y(x)z'(x),
\end{aligned}$$

so that

$$\begin{aligned}
z'(x) &= Y^{-1}(x)Y(x)z'(x) \\
&= Y^{-1}(x)[y'(x) - F(x)y(x)] \\
&= Y^{-1}(x)f(x).
\end{aligned}$$

Since

$$z(c) = Y^{-1}(c)y(c) = Y^{-1}(c)d,$$

it follows that

$$z(x) = \int_c^x Y^{-1}(t)f(t)\,dt + Y^{-1}(c)d.$$

Multiplication by $Y(x)$ completes the proof. □

Note. 1. If $n > 1$ and F is not constant, then, in general, there is no closed form formula for the fundamental matrix.
2. With the procedure outlined before Corollary 7.16, the results of this section also apply to linear n-th order differential equations. In particular, the elements of the first row of a fundamental system are called a fundamental system, say $y_1, \dots, y_n$ of the n-th order differential equation, and any solution of the homogeneous n-th order equation is a linear combination of the fundamental system.
Conversely, if $y_1, \dots, y_n$ is a fundamental system of the n-th order differential equation, then

$$Y = \left(y_j^{(i-1)}\right)_{i,j=1}^n$$

is a fundamental matrix of the corresponding first order system.
3. In principle, if we consider an n-th order differential equation with constant coefficients

$$y^{(n)} = \sum_{j=0}^{n-1} a_j y^{(j)},$$

one can use Jordan's canonical form to find a fundamental system. However, this is very cumbersome. One can proceed as follows: Consider the characteristic polynomial

$$\mu^n - \sum_{j=0}^{n-1} a_j \mu^j$$

and find its distinct (complex) zeros, say $\mu_1, \dots, \mu_k$, with multiplicities $p_1, \dots, p_k$. Then the n functions

$$x^l e^{\mu_j x}, \quad j = 1 \dots, k,\ l = 0, \dots, p_j - 1$$

form a fundamental system of the n-th order linear differential equations.