Analysis 1

Analysis Series

Chigüiro Collection

Universidad de los Andes
Facultad de Ciencias

Contents

These lecture notes are work in progress. They may be abandoned or changed radically at any moment. If you find mistakes or have suggestions how to improve them, please let me know.

Many students found numerous errors and improved parts of the script considerably. Special thanks for a very long list of errors and many improvements goes to Federico Fuentes.

I started writing these notes while I was teaching an introductory course on analysis at the Universidad de los Andes, Bogotá, Colombia in 2008-2. Since then, there have been changes every time I taught it again and hopefully it converges to an error-free state. The lecture is aimed at pregrade students who are already familiar with calculus.

The script is very much influenced by the lecture notes on Analysis 1 of my teacher C. Tretter and by the book *Analysis 1* by T. Bröcker [Brö92]. In addition, I was mainly using the books by Rudin [Rud76] and Dieudonné [Die69] besides several other books as sources to prepare the lecture.

An important part of any mathematics lecture are exercises. For each week there is a problem sheet with exercises (stolen from various books) which hopefully help to understand the material presented in the lecture.

Corrections, comments and remarks on the text are most welcome!

Bogotá, August 2017,
Monika Winklmeier.

Chapter 1

Preliminaries

The following is not intended to serve as an introduction into logic. It only tries to fix the meaning of some symbols the occur frequently.

Statements

A *statement* $\mathcal{A}$ is either true or false.

Examples: "The inner angles of an equilateral triangle are all equal." "Bogotá has more inhabitants than Berlin." "The sun is closer to the earth than the moon." "Today is Monday."

Non-Example: "This sentence is false."

Statements $\mathcal{A}$ and $\mathcal{B}$ can be negated and connected:

$\neg\mathcal{A}$	"not $\mathcal{A}$"; the statement $\mathcal{A}$ is not true.
$\mathcal{A} \wedge \mathcal{B}$	"$\mathcal{A}$ and $\mathcal{B}$"; both $\mathcal{A}$ and $\mathcal{B}$ are true.
$A \vee B$	"$\mathcal{A}$ or $\mathcal{B}$"; at least one of the statements $\mathcal{A}$ and $\mathcal{B}$ is true.
$\mathcal{A} \implies \mathcal{B}$	"$\mathcal{A}$ implies $\mathcal{B}$";
	or: "$\mathcal{A}$ is sufficient for $\mathcal{B}$."
	or: "$\mathcal{B}$ is necessary for $\mathcal{A}$."
$\mathcal{A} \iff \mathcal{B}$	$(\mathcal{A} \implies \mathcal{B}) \wedge (\mathcal{B} \implies \mathcal{A})$"
	"$\mathcal{A}$ is true if and only if $\mathcal{B}$ is is true."
	or: "$\mathcal{A}$ and $\mathcal{B}$ are equivalent."
	or: "$\mathcal{A}$ is necessary and sufficient for $\mathcal{B}$."

For convenience, sometimes the notation $\mathcal{A} \Longleftarrow \mathcal{B}$ is used instead of $\mathcal{B} \Longrightarrow \mathcal{A}$.
Obviously, $\mathcal{A} \iff \neg(\neg\mathcal{A})$ and $(\mathcal{A} \implies \mathcal{B}) \iff (\neg\mathcal{B} \implies \neg\mathcal{A})$.

Sets and Quantors

Let M be a set and x an object. Then exactly one of the following statements is true:

$x \in M$	"x is an element of M" or "x lies in M".
$x \notin M$	"x is not an element of M" or "x does not lie in M".

Hence, $\neg(x \in M) \iff (x \notin M)$.

Let M be a set and $\mathcal{A}$ a statement. Then we have the following statements:

$\forall\, x \in M : \mathcal{A}$	"For all elements x of M the statement $\mathcal{A}$ is true".
$\exists\, x \in M : \mathcal{A}$	"There exists at least one element x of M for which the statement $\mathcal{A}$ is true."
$\exists!\, x \in M : \mathcal{A}$	"There exists exactly one element x of M for which the statement $\mathcal{A}$ is true."
$\nexists\, x \in M : \mathcal{A}$	"There exists no element x of M for which the statement $\mathcal{A}$ is true."

It is easy to see that $\neg(\exists\, x \in M : \mathcal{A})$ is equivalent to $\nexists\, x \in M : \mathcal{A}$.
Instead of "$\forall x \in M : \mathcal{A}$", also the notation "$\mathcal{A},\ x \in M$," is used.

Definitions

For definitions, the symbols $:=$ and $:\Longleftrightarrow$ are used. The left hand side is defined by the right hand side.

Examples:

- A triangle Δ is called *equilateral*. $:\Longleftrightarrow$ The length of all sides of Δ are equal.
- $\mathcal{A} :=$ "The inner angles of an equilateral triangle are all equal".
- $M := \{1, 2, 3\}$.

More on sets

Let X be a set and $\mathcal{A}(x)$ a statement depending on the object x. The set

$$\{x \in X \,:\, \mathcal{A}(x)\} \quad \text{or} \quad \{x \in X \,|\, \mathcal{A}(x)\}$$

is the set of all $x \in X$ such that $\mathcal{A}$ is true.
The set which contains no elements is called the *empty set*. It is denoted by $\emptyset$.

Let M and N be sets. Then

$$M \subseteq N \quad :\Longleftrightarrow \quad \forall\, x \in M : x \in N,$$
$$M = N \quad :\Longleftrightarrow \quad (M \subseteq N) \wedge (N \subseteq M).$$

A set M with $M \subseteq N$ is called a *subset* of N. In this case N is called a *superset* of M and we write $N \supseteq M$.

Each set N has the *trivial subsets* $\emptyset$ and N. A subset $M \subseteq N$ is called a *proper subset of* N if M is not a trivial subset of N.

Other useful definitions are

$N \setminus M$	$:= \{x : x \in N \wedge x \notin M\}$	*difference*,
$N \cup M$	$:= \{x : x \in N \vee x \in M\}$	*union*,
$N \cap M$	$:= \{x : x \in N \wedge x \in M\}$	*intersection*,
$\emptyset$	$:= \{\}$	*empty set* (note: $\emptyset \neq \{0\}$),
$\mathbb{P}(N)$	$:= \{M : M \subseteq N\}$	*power set*,
$N \times M$	$:= \{(x, y) : x \in N,\ y \in M\}$	*Cartesian product*.

The sets M and N are called *disjoint* if and only if $M \cap N = \emptyset$. Often the union of disjoint sets is denoted by $M \mathbin{\dot\cup} N$ or $M \sqcup N$.

Relations

A *relation* R on a set M is a subset of $M \times M$. Instead of $(x, y) \in R$ and $(x, y) \notin R$ we write $x\,R\,y$ and $x \not\!R y$, respectively.

A relation R on a set M is called

- *reflexive* if and only if $x\,R\,x$ for all $x \in M$.
- *symmetric* if and only if for all $x, y \in M$ the relation $x\,R\,y$ implies $y\,R\,x$.
- *transitive* if and only if for all $x, y, z \in M$ the relations $x\,R\,y$ and $y\,R\,z$ imply $x\,R\,z$.

Examples for relations are $=,\ \subseteq,\ \perp,\ \leq,\ \neq$.

Ordered sets

$(M, <)$ is called a (*totally*) *ordered set by the relation* $<$ if the relation $<$ is transitive and for $x, y \in M$ exactly one of the following statements holds:

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

We use the following notations:

$$\begin{aligned} x > y \; &:\Longleftrightarrow \; y < x, \\ x \leq y \; &:\Longleftrightarrow \; x = y \vee x < y, \\ x \geq y \; &:\Longleftrightarrow \; x = y \vee x > y. \end{aligned}$$

Definition 1.1. Let $(M, <)$ be a totally ordered set, $N \subseteq M$ and $x \in M$.

$$\begin{aligned} x \text{ is a } \textit{lower bound of } N \quad &:\Longleftrightarrow \quad x \leq n,\ n \in N, \\ x \text{ is an } \textit{upper bound of } N \quad &:\Longleftrightarrow \quad x \geq n,\ n \in N. \end{aligned}$$

We say that

$$\begin{aligned} N \text{ is bounded from below} \quad &:\Longleftrightarrow \quad N \text{ has a lower bound}, \\ N \text{ is bounded from above} \quad &:\Longleftrightarrow \quad N \text{ has an upper bound}, \\ N \text{ is bounded} \quad &:\Longleftrightarrow \quad N \text{ has an upper and a lower bound}. \end{aligned}$$

The *infimum of* N, denoted by $\inf N$, is the *greatest lower bound of* N, i. e., $\inf N$ is a lower bound of N and for every lower bound x' of N we have $\inf N \geq x'$. If an element n of N is a lower bound of N, then it is called *minimum of* N, denoted by $\min N$.

The *supremum of* N, denoted by $\sup N$, ist the *least upper bound of* N, i. e., $\sup N$ is an upper bound of N and for every upper bound x of N it follows that $\sup N \leq x$. If an element n of N is an upper bound of N, then it is called *maximum of* N, denoted by $\max N$.

Remark. Neither the infimum nor the minimum need to exist. If they exist, then they are unique. If the minimum exists, then also the infimum exists and $\min N = \inf N$. The same assertions hold for the supremum and the maximum.

Functions

Let M and $N \neq \emptyset$ be sets. A *function* (or a *mapping*) *from* M *to* N

$$f : M \to N, \quad x \mapsto f(x),$$

assigns to each element $x \in X$ a *unique* $f(x) \in N$. M is called the *domain* of f.

$G(f) :=$ *graph* of $f := \big\{\big(x, f(x)\big) : x \in M\big\} \subseteq M \times N$,

$\mathrm{R}(f) :=$ *range* of $f := \big\{f(x) : x \in M\big\}$.

A function $f : M \to N$ is called

injective (or *one-to-one*) $:\iff \forall\, x, x' \in M : \big(f(x) = f(x') \implies x = x'\big)$,

surjective $:\iff R(f) = N$,

bijective $:\iff f$ injective $\wedge$ f surjective.

Since

$$f \;\text{ bijective} \iff \forall\, y \in N\ \exists!\, x \in M : f(x) = y,$$

a bijective function $f : M \to N$ defines a function $N \to M$, the so-called *inverse function of* f:

$$f^{-1} : N \to M, \quad y \mapsto x =: f^{-1}(y).$$

Remark 1.2. A function is not only the rule how to assign an element y to some x. The sets M and N are also part of the function. For example, the following functions are all different:

$$\begin{aligned} &f_1 : \mathbb{R} \to \mathbb{R}, && x \mapsto x^2,\\ &f_2 : (0,\infty) \to \mathbb{R}, && x \mapsto x^2,\\ &f_3 : (0,\infty) \to (0,\infty), && x \mapsto x^2,\\ &f_4 : (0,\infty) \to (1,\infty), && x \mapsto x^2. \end{aligned}$$

The function f_1 is neither injective nor surjective, the function f_2 is injective, the function f_3 is bijective, and the function f_4 is not well-defined.

Let M and N be sets and $f : M \to N$ a function. Given a subset $A \subseteq M$ we define the *restriction of* f *to* A by

$$f|_A : A \to N, \quad f|_A(x) = f(x).$$

f is then called *extension of* $f|_A$. Another notation for the restriction of f is $f \restriction A$.
The *image of* A *under* f is

$$f(A) := R(f|_A) := \{y \in N \;:\; \exists x \in A : f(x) = y\} = \{f(x) \;:\; x \in A\}.$$

For a subset $B \subseteq N$ the set

$$f^{-1}(B) := \{x \in M \;:\; f(x) \in B\}$$

is called the *preimage of* B *under* f.

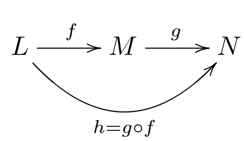
Two functions f, $g : M \to N$ are equal, denoted by $f \equiv g$, if and only if

$$f(x) = g(x), \quad x \in M.$$

For example, in Remark 1.2 the functions $f_1|_{(0,\infty)}$ and f_2 are equal, f_2 and f_3 are not equal.
For sets L, M, N and functions $f : L \to M$, $g : M \to N$ we define the *composition of* f *and* g

$$h = g \circ f : L \to N, \quad x \mapsto h(x) := (g \circ f)(x) := g(f(x)).$$

As a diagram:



Example. Let $f : M \to N$ bijective. Then

$$f^{-1} \circ f : M \to M, \quad (f^{-1} \circ f)(x) = x,$$
$$f \circ f^{-1} : N \to N, \quad (f \circ f^{-1})(y) = y.$$

Proofs

Usually, there are several ways to prove statements like $\mathcal{A} \implies \mathcal{B}$. The end of proofs are usually indicated by the symbol $\square$. The most common types of proofs are the following:

Direct proof

A direct proof of a statement $\mathcal{C}$ starts with a set of axioms that are agreed upon. Using a chain of conclusions, finally $\mathcal{C}$ is established.

Example 1.3. For all $n \in \mathbb{N}$ the following holds:

$$n \text{ is even} \quad \implies \quad n^2 \text{ is even.}$$

Proof.
$$\begin{aligned} n \text{ even} \quad &\implies \quad \exists\, m \in \mathbb{N} : n = 2m \\ &\implies \quad n^2 = (2m)^2 = 2 \cdot 2m^2 \\ &\implies \quad n^2 = 2m' \text{ for } m' = 2m^2 \in \mathbb{N} \\ &\implies \quad n^2 \text{ is even.} \end{aligned}$$
$\square$

(The natural numbers $\mathbb{N}$ are introduced in Section 2.1.)

Proof by transposition

Often it is simpler to proof $\neg\mathcal{B} \implies \neg\mathcal{A}$ than the equivalent statement $\mathcal{A} \implies \mathcal{B}$.

Example 1.4. For all $n \in \mathbb{N}$ the following holds:

$$n^2 \text{ is even} \quad \implies \quad n \text{ is even.}$$

Proof. The implication above is equivalent to the implication:

For all $n \in \mathbb{N}$ the following holds: n is odd $\implies$ n^2 is odd.

The proof of the latter statement is similar to the proof of example 1.3:

$$\begin{aligned} n \text{ odd} \quad &\implies \quad \exists\, m \in \mathbb{N}_0 : n = 2m + 1 \\ &\implies \quad n^2 = (2m+1)^2 = (2m)^2 + 2 \cdot 2m + 1 \\ &\implies \quad n^2 = 2m' + 1 \text{ for } m' = 2m^2 + 2m \in \mathbb{N} \\ &\implies \quad n^2 \text{ is odd.} \end{aligned}$$
$\square$

Proof by contradiction

In order to proof a statement $\mathcal{A} \implies \mathcal{B}$ it is assumed that both $\mathcal{A}$ and $\neg\mathcal{B}$ are true. Then it is shown that this leads to a contradiction, indicated in these notes by ✠.

Example 1.5. For all $n \in \mathbb{N}$ the following holds:

$$n^2 \text{ is odd} \quad \implies \quad n \text{ is odd.}$$

Proof. Assume that the implication is wrong. Then there exists a $n \in \mathbb{N}$ such that n^2 is odd but n is even. This contradicts the assertion of Example 1.3. □

Example 1.6. $a, b \in \mathbb{R} \implies 2ab \leq a^2 + b^2$.

Proof. Assume that the implication is wrong. Then there exist $a, b \in \mathbb{R}$ such that $2ab > a^2 + b^2$. It follows that

$$0 > a^2 + b^2 - 2ab = (a - b)^2 \geq 0. \quad ✠$$

□

Remark. The fact that statement $\mathcal{A}$ implies statement $\mathcal{B}$ and that $\mathcal{B}$ is true, does not imply that also $\mathcal{A}$ is true. For example, the implication

$$-1 = 1 \implies (-1)^2 = 1^2$$

is true and the statement on the right hand side is true, but this does not imply that the initial statement $-1 = 1$ is true.

Proof by Induction

The idea of proof by induction is to show a statement $\mathcal{A}(1)$ (base of induction). When the implication $\mathcal{A}(n) \implies \mathcal{A}(n+1)$ $n \in \mathbb{N}$, is shown, then the statement $\mathcal{A}(n)$ is true for all $n \in \mathbb{N}$. The induction principle is discussed in Section 2.2.

Chapter 2

Natural numbers

The sets

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

can be introduced axiomatically, see [Lan51].
In this chapter the natural numbers

$$\mathbb{N}_0 = \{0, 1, 2, \dots\} \quad \text{and} \quad \mathbb{N} = \mathbb{N}_0 \setminus \{0\} = \{1, 2, \dots\}$$

are defined by the Peano axioms.

2.1 Peano axioms

The set of the natural numbers $\mathbb{N}_0$ satisfies the following axioms:

(P1) $0 \in \mathbb{N}_0$,

(P2) there exists a mapping $\nu : \mathbb{N}_0 \to \mathbb{N}$,

(P3) ν is injective, that is: $m, n \in \mathbb{N},\ m \neq n \implies \nu(m) \neq \nu(n)$,

(P4) **Axiom of Induction:** for all $M \subseteq \mathbb{N}_0$ the following implication holds: $\big(0 \in M \wedge (n \in M \implies \nu(n) \in M)\big) \implies M = \mathbb{N}_0$.

Remark.

- For $n \in \mathbb{N}$ the number $\nu(n)$ is called the *successor of* n.
- (P1) implies that the natural numbers are not the empty set.
- (P2) implies that 0 is not the successor of any natural number.
- (P4) implies that ν is surjective.

As usual, we write

$$\nu(0) = 1, \quad \nu(\nu(0)) = 2, \quad \underbrace{\nu(\nu(\dots\nu}_{n \text{ times}}(0)\dots)) = n.$$

The operations + (addition) and $\cdot$ (multiplication) are introduced by using the function ν:

Definition 2.1. For $n, m \in \mathbb{N}_0$ let

$$n + 0 := n, \qquad n + 1 := \nu(n), \qquad n + \nu(m) := \nu(n + m)$$
$$n \cdot 0 := 0, \qquad n \cdot 1 := n, \qquad n \cdot \nu(m) := n \cdot m + n.$$

Remark 2.2. (i) It can be shown that $+$ and $\cdot$ are commutative, associative and distributive (see Section 3.1).

(ii) For $m, n \in \mathbb{N}$ exactly one of the following relations holds:

(a) $m = n$.

(b) There exists exactly one $x \in \mathbb{N}$ such that $n = m + x$.

(c) There exists exactly one $x \in \mathbb{N}$ such that $m = n + x$.

In case (b) the number $x =: n - m$ is called the *difference* of m and n.

Definition 2.3. Let $m, n \in \mathbb{N}_0$. Then

$$m < n :\iff \exists x \in \mathbb{N} : n = m + x.$$

Remark. It can be shown that $(\mathbb{N}_0, <)$ is a totally ordered set.

Theorem 2.4 (well-ordering principle). *Every non-empty subset of* $\mathbb{N}_0$ *has a smallest element.*

Proof. We have to show:

$$M \subseteq \mathbb{N}_0,\ M \neq \emptyset \implies \exists\, m_0 \in M\ :\ \big(\forall\, m \in M : m_0 \leq m\big).$$

Let $M \subseteq \mathbb{N}_0$, $M \neq \emptyset$. Then there exists an $m_0 \in M$. Let

$$A = \{k \in \mathbb{N}_0\ :\ k \leq m,\ m \in M\}.$$

Obviously, $0 \in A$ and $m_0 + 1 \notin A$, hence $A \neq \mathbb{N}_0$. Therefore there exists an $a \in A$ such that $a + 1 \notin A$ (otherwise $A = \mathbb{N}_0$ by the axiom (P4)). Hence there must be an element $m \in M$ such that $a \leq m < a + 1$, hence $m = a$ and a is the minimum of M. □

Without proof we cite the following important facts:

Theorem 2.5. *The natural numbers are not bounded from above, i. e.*

$$\nexists\, N \in \mathbb{N}_0 :\ \big(\forall\, n \in \mathbb{N}_0 : n \leq N\big).$$

Theorem 2.6 (Euclidean algorithm, Division with remainder). *For every* $m \in \mathbb{N}$ *and* $n \in \mathbb{N}_0$ *there exist uniquely determined* $k, l \in \mathbb{N}_0$*,* $l < m$*, such that*

$$n = k \cdot m + l. \tag{2.1}$$

2.2 The induction principle

Principle of induction. Let $(\mathcal{A}(n))_{n\in\mathbb{N}}$ be a family of statements, such that

(i) *Basis*: $\mathcal{A}(0)$ is true.

(ii) *Inductive step*: For all $n \in \mathbb{N}_0$ the implication $\mathcal{A}(n) \implies \mathcal{A}(n+1)$ is true.

Then $\mathcal{A}(n)$ is true for all $n \in \mathbb{N}_0$.

Proof. Let $M := \{n \in \mathbb{N}_0 : \mathcal{A}(n) \text{ is true}\} \subseteq \mathbb{N}_0$. Then $0 \in M$ by (i) and with every $m \in M$ also $\nu(m) \subseteq M$ by (ii). By (P4) it follows that $M = \mathbb{N}_0$. □

A variation of the induction principle is the following: Let $n_0 \in \mathbb{N}_0$ and assume that

(i) *Basis*: $\mathcal{A}(n_0)$ is true.

(ii) *Inductive step*: For all $n \in \mathbb{N}_0$, $n \geq n_0$, the implication $A(n) \implies A(n+1)$ is true.

Then $\mathcal{A}(n)$ is true for all $n \in \mathbb{N}_0$, $n \geq n_0$.

Remark (Complete induction principle). Assume that for all $n \in \mathbb{N}$ the implication

$$\mathcal{A}(k) \quad \text{is true for all} \quad k < n \qquad \implies \qquad \mathcal{A}(n) \quad \text{is true}$$

holds. Then $\mathcal{A}(n)$ holds for all $n \in \mathbb{N}_0$.

Proof. Assume that there exists an $n \in \mathbb{N}_0$ such that $\mathcal{A}(n)$ is not true. Then the set $B := \{n \in \mathbb{N}_0 : \mathcal{A}(n) \text{ is not true}\}$ is not empty. By the well-ordering principle (Theorem 2.4) B has a minimum $n_0 := \min B$ and by definition of B the statement $\mathcal{A}(m)$ is true for all $m < n_0$. Hence the induction assumption implies that $\mathcal{A}(n_0)$ is true. ✠ □

The principle of induction can be used for definitions. For example:

Definition 2.7. Let $k_0 \in \mathbb{N}_0$ and $a_k \in \mathbb{N}$, $k \in \mathbb{N}_0,\ k \geq k_0$. Then the symbols

$$\sum_{k=k_0}^{n} a_k, \qquad \prod_{k=k_0}^{n} a_k, \qquad n \in \mathbb{N},\ n \geq k_0,$$

are defined by

$$\sum_{k=k_0}^{k_0} a_k := a_{k_0}, \qquad \prod_{k=k_0}^{k_0} a_k := a_{k_0},$$

and

$$\sum_{k=k_0}^{n+1} a_k := \left(\sum_{k=k_0}^{n} a_k\right) + a_{n+1}, \qquad \prod_{k=k_0}^{n+1} a_k := \left(\prod_{k=k_0}^{n} a_k\right) \cdot a_{n+1}, \qquad n \geq k_0.$$

For $n < k_0$ we define the *empty sum* and *empty product* by

$$\sum_{k=k_0}^{n} a_k = 0, \qquad \prod_{k=k_0}^{n} a_k = 1.$$

Theorem 2.8. $$\sum_{k=1}^{n} k = 1 + 2 + \cdots + n = \frac{1}{2}n(n+1), \quad n \in \mathbb{N}.$$

Proof by induction on n.

(i) *Basis* $\underline{n=1}$: $\displaystyle\sum_{k=1}^{1} k = 1 = \frac{1}{2}1(1+1).$ ✓

(ii) *Induction step:* $\underline{n \curvearrowright n+1}$ for arbitrary $n \in \mathbb{N}$.

We write down the *induction hypothesis* and what we want to prove:

induction hypothesis: $$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1). \qquad \text{(statement } \mathcal{A}(n))$$

we want to show: $$\sum_{k=1}^{n+1} k = \frac{1}{2}(n+1)(n+2). \qquad \text{(statement } \mathcal{A}(n+1))$$

The implication $\mathcal{A}(n) \implies \mathcal{A}(n+1)$ follows from

$$\begin{aligned}\sum_{k=1}^{n+1} k &= \Big(\sum_{k=1}^{n} k\Big) + (n+1) \overset{\text{ind.hyp.}}{=} \frac{1}{2}n(n+1) + (n+1)\\ &= \Big(\frac{1}{2}n+1\Big)(n+1) = \frac{1}{2}(n+2)(n+1).\end{aligned}$$ □

Proposition 2.9. $2^n > n^2, \quad n \in \mathbb{N}, n \geq 5$.

Proof by induction on n, $n_0 = 5$.

$\underline{n=5}$: $2^5 = 32 > 25 = 5^2$. ✓

$\underline{n \curvearrowright n+1}$: $2^{n+1} = 2 \cdot 2^n \overset{\text{ind.hyp.}}{>} 2\,n^2$. Since for $n \geq 3$

$$n^2 = n(n-2+2) = n(n-2) + 2n \geq 1 + 2n$$

it follows that

$$2^{n+1} = 2 \cdot 2^n \overset{\text{ind.hyp}}{>} 2n^2 = n^2 + n^2 \geq n^2 + 2n + 1 = (n+1)^2.$$ □

2.3 Countable sets

Definition. The sets M and N *have the same cardinal number* if and only if there exists a bijection $\varphi : M \to N$. In this case we write $M \sim N$.

Obviously, the relation $\sim$ is

(i) *reflexive*: $M \sim M$,
(ii) *symmetric*: $M \sim N \implies N \sim M$,
(iii) *transitive*: $M \sim N \wedge N \sim P \implies M \sim P$.

Remark. For $m, n \in \mathbb{N}$ the following equivalence holds:

$$\{1, 2, \ldots, n\} \sim \{1, 2, \ldots, m\} \iff n = m.$$

Definition 2.10. A set M is called

(i) *finite* if $M = \emptyset$ or if $M \sim \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$,
(ii) *infinite* if it is not finite,
(iii) *countable* (or *denumerable*) if $M \sim \mathbb{N}$,
(iv) *uncountable* if it is neither countable nor finite,
(v) *at most countable* if it is either countable or finite.

If $M \sim \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, then

$$\#M := |M| := n = \text{number of elements in } M.$$

Examples. The sets $\mathbb{N}$, $\mathbb{N}_0$, $-\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ are countable, see Corollary 2.14 (Exercise 2.5).
The set of all real numbers $\mathbb{R}$ is uncountable (Corollary 4.57).

The proofs of the following facts can be found, e. g., in [Rud76, Chapter 2].

Proposition 2.11. *Any subset of a countable set is at most countable.*

Proposition 2.12. *The finite union of finite sets is finite.*
The countable union of finite sets is at most countable.
The finite union of countable sets is countable.
The countable union of countable sets is countable.

Proposition 2.13. *If M and N are countable, then $M \times N$ is countable.*

Corollary 2.14. *The sets $\mathbb{Z}$ and $\mathbb{Q}$ are countable.*

Proof. $\mathbb{Z} = \mathbb{N}_0 \cup (-\mathbb{N})$ is countable by Proposition 2.12.
Since every element of $\mathbb{Q}$ is of the form $\frac{p}{q}$ with some $p \in \mathbb{Z}$ and $q \in \mathbb{N}$, the set $\mathbb{Q}$ can be identified with a subset of the countable set $\mathbb{Z} \times \mathbb{N}$, so $\mathbb{Q}$ is at most countable. On the other hand, $\mathbb{Q}$ contains the subset $\{\frac{1}{n} : n \in \mathbb{N}\}$ which evidently has the same cardinality as $\mathbb{N}$. Therefore $\mathbb{Q}$ cannot be finite by Proposition 2.12. In conclusion, $\mathbb{Q}$ is countable. □

2.4 Binomial coefficients

Definition 2.15. For $n \in \mathbb{N}_0$ we define $n! \in \mathbb{N}$ ("n factorial") recursively by

(i) $0! := 1$,

(ii) $(n+1)! := (n+1) \cdot n!, \quad n \in \mathbb{N}_0$.

Remark. $n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \cdots \cdot n, \quad n \in \mathbb{N}_0.$
$1! = 1,\ 2! = 2,\ 3! = 6,\ 4! = 24,\ \ldots,\ 10! = 3\,628\,800$ is growing very fast.

Theorem 2.16. *Let M, N finite sets with $\#M = \#N = n \in \mathbb{N}$ elements. Then there are exactly $n!$ bijections $M \to N$.*

Proof by induction on n. (i) *Basis*: For $n = 1$, the assertion is clear.

(ii) *Induction step* $\underline{n \curvearrowright n+1}$: Let M and N be sets with $n+1$ elements. To define a bijection $f : M \to N$ we fix an arbitrary element $x \in M$. Then there are $n+1$ possible values of $f(x) \in N$. The mapping $f : M \setminus \{x\} \to N \setminus \{f(x)\}$ must also be a bijection, hence by induction hypothesis there are $n!$ possibilities to extend $f : \{x\} \to f(x)$ to a bijection $M \to N$. In summary, there are exactly $(n+1) \cdot n! = (n+1)!$ different bijections $M \to N$. □

Definition 2.17. Let M be a set. A *permutation of* M is a bijection $M \to M$.

By Theorem 2.16 a set M with n elements has exactly $n!$ permutations. Moreover, since an order on M is equivalent to a bijection $\{1, 2, \ldots, n\} \to M$, there are exactly $n!$ order of M.

Example.

- $M = \{a\}$ has only one order and only the permutation $a \mapsto a$.
- $M = \{a, b\}$ with $a \neq b$ has
  2 order $1 \mapsto a$, $2 \mapsto b$ and $1 \mapsto b$, $2 \mapsto a$,
  2 permutations: $a \mapsto a$, $b \mapsto b$ and $a \mapsto b$, $b \mapsto a$.

Remark. For the rest of this section the rational numbers $\mathbb{Q}$ are used (for the definition of the binomial coefficients). This could be avoided by combining the definition and theorem 2.20 to something like: *Given $k, n \in \mathbb{N}_0$, $k \leq n$ there exists a natural number x such that $xk!(n-k)! = n!$. The number x is then denoted by $x = \binom{n}{k}$.*

Definition 2.18. For k, $m \in \mathbb{N}_0$ we define the *binomial coefficients*

$$(k,\, m) := \binom{k+m}{k} := \frac{(k+m)!}{k!\, m!}.$$

("$k+m$ choose k"). We set $\binom{n}{k} = 0$ if $k > n$.

Remark. Let $k, n \in \mathbb{N}_0$, $k \leq n$. It follows immediately from the definition that

$$\binom{n}{k} = \binom{n}{n-k} \quad \text{and} \quad \binom{n}{0} = \binom{n}{n} = 1.$$

Proposition 2.19. $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}, \quad k,\, n \in \mathbb{N},\, k \leq n-1.$

Proof. Using Definition 2.18 a straightforward calculation yields

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \frac{(n-1)!}{(k-1)!\,(n-1-(k-1))!} + \frac{(n-1)!}{k!\,(n-1-k)!}$$

$$= \frac{(n-1)!}{k!\,(n-k)!}\,(k+n-k) = \frac{n!}{k!\,(n-k)!} = \binom{n}{k}.$$ □

Theorem 2.20. *Let M be a set with $\#M = n < \infty$. Then there are exactly $\binom{n}{k}$ different subsets of M with cardinality $k \leq n$.*

Corollary 2.21. $\binom{n}{k} \in \mathbb{N}$ *for all k, $n \in \mathbb{N}_0$, $k \leq n$.*

Proof of Theorem 2.20. We prove the claim by induction on n.
$\underline{n = 0}$: In this case $M = \emptyset$ and necessarily $k = 0$. Since $\emptyset$ is the only subset of M the number of all subsets with zero elements is $1 = \binom{0}{0}$.
$\underline{n \curvearrowright n+1}$: Let $\#M = n+1$, $k \in \mathbb{N}_0$, $k \leq n+1$, and define

$$C(n,k) := \#\{N \subseteq M \,:\, \#N = k\}.$$

We have to show

$$C(n,k) = \binom{n+1}{k}.$$

Fix $x \in M$. Since by induction hypothesis and the defintion of $\binom{n}{n+1}$ there are exactly $\binom{n}{k}$ subsets of $M \setminus \{x\}$ with cardinality k, there are exactly $\binom{n}{k}$ subsets of M with cardinality k which do not contain x. Again by induction hypothesis, there exist exactly $\binom{n}{k-1}$ subsets of $M \setminus \{x\}$ with cardinality $k-1$, therefore there exist exactly $\binom{n}{k-1}$ subsets of M with cardinality k containing x. Since an arbitrary subset of M either contains or does not contain the element x, the number of all subsets with cardinality k is

$$C(n,k) = \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}. \qquad \square$$

Theorem 2.22 (Binomial expansion). *For numbers x, y and $n \in \mathbb{N}_0$ the following holds:*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{\substack{0 \le k \le n \\ k+m=n}} (k,m) x^k y^m.$$

Remark. The formula holds for all x, y in any commutative ring R with the canonical actions of $\mathbb{N}$ on R, for example for real numbers, matrices, functions, etc.

Proof. The second equality is clear. We prove the first equality by induction on n.

$\underline{n=0}$: $(x+y)^0 = 1 = \binom{0}{0} x^0 y^0$. ✓

$\underline{n \curvearrowright n+1}$: Using the induction hypothesis we find

$$\begin{aligned}(x+y)^{n+1} &= (x+y)\cdot(x+y)^n = (x+y)\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{k+1}y^{n-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k+1}.\end{aligned}$$

An index shift $k \curvearrowright k-1$ in the first sum yields

$$\begin{aligned}(x+y)^{n+1} &= \sum_{k=1}^{n+1}\binom{n}{k-1}x^k y^{n-(k-1)} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-(k-1)} \\ &= x^{n+1} + \sum_{k=1}^{n}\underbrace{\left(\binom{n}{k-1}+\binom{n}{k}\right)}_{=\binom{n+1}{k} \text{ by Prop. } 2.19} x^k y^{n+1-k} + y^{n+1} \\ &= \sum_{k=0}^{n+1}\binom{n+1}{k}x^k y^{n+1-k}. \qquad \square\end{aligned}$$

In the special cases $x = y = 1$ and $x = -y = 1$, Theorem 2.22 yields

Corollary 2.23. $\displaystyle\sum_{k=0}^{n}\binom{n}{k} = 2^n$ *for* $n \in \mathbb{N}_0$ *and* $\displaystyle\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$ *for* $n \in \mathbb{N}$.

Proof. The formulae follow from the binomial expansion (Theorem 2.22):

$$\sum_{k=0}^{n}\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k}1^k 1^{n-k} = (1+1)^n = 2^n, \qquad n \in \mathbb{N}_0,$$

$$\sum_{k=0}^{n}(-)^k\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k}(-1)^k 1^{n-k} = (1-1)^n = 0^n = 0, \qquad n \in \mathbb{N}. \qquad \square$$

Corollary 2.24. *Let M be a set with $\#M = n < \infty$.*

(i) *By Theorem 2.20 and Corollary 2.23, M has 2^n subsets, i. e., $\#\mathbb{P}M = 2^n$.*

(ii) *If $n \neq 0$, then M has as many subsets with an even number of elements as subsets with an odd number of elements by Theorem 2.20 and Corollary 2.23.*

Chapter 3

Real and complex numbers

Integers

The ring of *integer numbers* $\mathbb{Z}$ is the smallest extension of the natural numbers such that for each $n \in \mathbb{N}$ the equation

$$x + n = 0$$

has a solution in $\mathbb{Z}$ and $(\mathbb{Z}, +, \cdot)$ is a commutative ring with identity, that is, the associativity and commutativity laws hold. The solution of $x+n = 0$ is denoted by $-n$ and we write $m+(-n) = m-n$.

Rational numbers

The *rational numbers* $\mathbb{Q}$ are the field of fractions of $\mathbb{Z}$, that is, the smallest field containing $\mathbb{Z}$ such that for each $n \in \mathbb{Z} \setminus \{0\}$ the equation

$$x \cdot n = 1$$

has a solution in $\mathbb{Q}$. The elements of $\mathbb{Q}$ are equivalence classes of the form $\frac{p}{q}$ with $p, q \in \mathbb{Z}$, $q \neq 0$.

The order relation $<$ on $\mathbb{N}$ in Definition 2.3 can be extended to $\mathbb{Z}$ and $\mathbb{Q}$.

The field $\mathbb{Q}$ is still not sufficient:

(i) Not all equations have solution, e. g., $x^2 = 2$ has no solution in $\mathbb{Q}$.

(ii) Not every bounded subset of $\mathbb{Q}$ has a supremum, for instance $\{x \in \mathbb{Q} : x^2 < 2\}$ has no supremum in $\mathbb{Q}$ (see Exercise 3.3).

Proof of (i). Assume that there exist $p, q \in \mathbb{Z}$, $q \neq 0$, without common divisor such that $\left(\frac{p}{q}\right)^2 = 2$. Since $p^2 = 2q^2$, there exists an $p' \in \mathbb{Z}$ such that $p = 2p'$. Since $2q^2 = 4p'^2$, 2 divides also q, in contradiction to the assumption that p and q have no common divisors. □

3.1 Ordered fields

Definition 3.1. A *field* $(K, +, \cdot)$ is a set K together with operations

$$
\begin{aligned}
+ &: K \times K \to K, \quad (x, y) \mapsto x + y, && \text{(addition)}\\
\cdot &: K \times K \to K, \quad (x, y) \mapsto x \cdot y, && \text{(multiplication)}
\end{aligned}
$$

satisfying the following axioms:

Axioms of addition

(A1) $x + (y + z) = (x + y) + z, \quad x, y, z \in K$ (*associativity*),

(A2) $x + y = y + x, \quad x, y \in K$ (*commutativity*),

(A3) $\exists\, 0 \in K :\ x + 0 = x, \quad x \in K$ (*identity element* 0 *of addition*),

(A4) $\forall\, x \in K\ \exists\, -x \in K :\ x + (-x) = 0$ (*additive inverse element*).

Axioms of multiplication

(M1) $x \cdot (y \cdot z) = (x \cdot y) \cdot z, \quad x, y, z \in K$ (*associativity*),

(M2) $x \cdot y = y \cdot x, \quad x, y \in K$ (*commutativity*),

(M3) $\exists\, 1 \in K \setminus \{0\} :\ x \cdot 1 = x, \quad x \in K$ (*multiplicative identity element* 1),

(M4) $\forall\, x \in K,\ x \neq 0,\ \exists\, x^{-1} \in K :\ x \cdot x^{-1} = 1$ (*multiplicative inverse element*).

Law of Distribution

(D) $x \cdot (y + z) = x \cdot y + x \cdot z, \quad x, y, z \in K.$

If it is clear what the operations $+$ and $\cdot$ on K are, then one writes usually simply K instead of $(K, +, \cdot)$.

Notation 3.2. The following notation is commonly used:

$$
\begin{aligned}
x - y &:= x + (-y), \quad xy := x \cdot y, && x,\, y \in K,\\
\frac{x}{y} &:= x \cdot y^{-1}, && x,\, y \in K,\ y \neq 0,\\
xy + z &:= (x \cdot y) + z, \quad \text{etc.} && x,\, y,\, z \in K.
\end{aligned}
$$

Remark 3.3.

- $(K, +)$ satisfying the axioms (A1) – (A4) is called a *commutative group*.
- $(K, +, \cdot)$ satisfying the axioms (A1) – (A4), (M1), (M2) and (D) is called a *commutative ring*.

Examples.

- $(\mathbb{Q}, +, \cdot)$ and $(\mathbb{R}, +, \cdot)$ are fields,
- $(\mathbb{Z}, +, \cdot)$ is a ring but not a field because (M4) is not satisfied,
- $(\mathbb{N}, +, \cdot)$ is not a ring because (A4) is not satisfies,
- $(\mathbb{F}_2, +, \cdot)$ with $\mathbb{F}_2 = \{0, 1\}$ and $+, \cdot$ defined by

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

is the only field with 2 elements.

The following corollary follows immediately from the axioms:

Corollary 3.4. *For a field* $(K, +, \cdot)$ *the following is true:*

(i) *The identity elements* 0 *and* 1 *are uniquely determined.*

(ii) *For* $x \in K$ *and* $y \in K \setminus \{0\}$ *the inverse elements* $-x$ *and* y^{-1} *are uniquely determined.*

(iii) $-0 = 0, \ 1^{-1} = 1$.

(iv) $-(-x) = x, \quad x \in K$.

(v) $-(x + y) = -x + (-y), \quad x, y \in K$.

(vi) *The equation* $a + x = b$ *for* $a, b \in K$ *has the unique solution* $x = b - a$ *in* K.

(vii) $(x^{-1})^{-1} = x, \quad x \in K, \ x \neq 0$.

(viii) $(x \cdot y)^{-1} = x^{-1} \cdot y^{-1}, \quad x, y \in K, \quad x, y \neq 0$.

(ix) *The equation* $a \cdot x = b$ *for* $a, b \in K$, $a \neq 0$ *has the unique solution* $x = ba^{-1}$ *in* K.

(x) $x \cdot 0 = 0 \cdot x = 0, \quad x \in K$.

(xi) $\big(x \cdot y = 0 \iff x = 0 \ \vee \ y = 0\big), \quad x, y \in K$.

(xii) $(-x) \cdot y = -(x \cdot y), \quad x, y \in K;$ *in particular,* $-y = -1 \cdot y$.

(xiii) $(-x) \cdot (-y) = x \cdot y, \quad x, y \in K$.

Proof. We prove only (i), (vi), (x) and (xi).
(i) *Uniqueness of the additive identity*: Let $0, 0' \in K$ be additive identity elements. We have to show $0 = 0'$. This follows from

$$0' \overset{(A3)}{=} 0' + 0 \overset{(A2)}{=} 0 + 0' \overset{(A3)}{=} 0.$$

The uniqueness of the multiplicative identity element can be proved analogously.
(vi) *Existence of the solution:* Let $x = b - a$. Then x is a solution of $a + x = b$ since

$$a + (b - a) \overset{(A2)}{=} a + (-a + b) \overset{(A1)}{=} \underbrace{(a + (-a))}_{=0 \text{ by } (A4)} + b = 0 + b \overset{(A2)}{=} b + 0 \overset{(A3)}{=} b.$$

Uniquness of the solution: Let $s, s' \in K$ be solutions of $a + x = b$. Then

$$s = s + b - b = s + (s' + a) - b \underset{(A2)}{\overset{(A1)}{=}} (s + a) - b + s' = b - b + s' = 0 + s' \overset{(A2)}{=} s' + 0 \overset{(A3)}{=} s'.$$

(x) Since the solution of $x \cdot 0 + x = x \cdot 0$ is unique by (vi) and since (use (D) in the first line and (A3) in the second line)

$$x \cdot 0 + x \cdot 0 = x(0 + 0) = x \cdot 0,$$
$$x \cdot 0 + 0 = x \cdot 0.$$

it follows that $x \cdot 0 = 0$. the commutativity (M2) yields $0 \cdot x = x \cdot 0 = 0$.
(xi) "$\Longrightarrow$": Let $x \cdot y = 0$. If $x = 0$, then the assertion is clear. Now assume $x \neq 0$.

$$y \overset{(M3)}{=} y \cdot 1 \overset{(M4)}{=} y \cdot (x \cdot x^{-1}) \overset{(M1)}{=} (y \cdot x) \cdot x^{-1} = 0 \cdot x^{-1} \overset{(x)}{=} 0.$$

"$\Longleftarrow$": Follows from (x). □

Definition 3.5. Let $(K, +, \cdot)$ be a field. For $x, y \in K$, $y \neq 0$, define

$$\begin{aligned} nx := xn &:= \sum_{j=1}^{n} x, & x^n &:= \prod_{j=1}^{n} x, & n &\in \mathbb{N}_0, \\ nx := xn &:= -n(-x), & y^n &:= (y^{-n})^{-1}, & n &\in \mathbb{Z} \setminus \mathbb{N}_0. \end{aligned}$$

Proposition 3.6. *Let $(K, +, \cdot)$ a field and x, $y \in K$, $n, m \in \mathbb{Z}$. Then*

(i) $nx + mx = (n+m)x$, (ii) $n(mx) = (nm)x$, (iii) $n(x+y) = nx + ny$,

If $x, y \neq 0$, then also

(iv) $x^n \cdot x^m = x^{n+m}$, (v) $(x^n)^m = x^{n \cdot m}$, (vi) $(x \cdot y)^n = x^n \cdot y^n$.

The statements (iv), (v) and (vi) hold for $x = 0$ or $y = 0$ if $m, n > 0$.

Proof. We prove only the last statement. The other ones can be proved similarly.
First, let $n \in \mathbb{N}_0$.
$\underline{n = 0}$: By Definition 3.5 and (M3): $(xy)^0 = 1 = 1 \cdot 1 = x^0 \cdot y^0$. ✓
$\underline{n \curvearrowright n+1}$: By Definition 3.5 and the induction hypothesis:

$$\begin{aligned} x^{n+1} \cdot y^{n+1} &= (x \cdot x^n) \cdot (y \cdot y^n) \overset{\text{(M1)}}{\underset{\text{(M2)}}{=}} (x \cdot y) \cdot \underbrace{(x^n \cdot y^n)}_{=(x \cdot y)^n} = (x \cdot y) \cdot (x \cdot y)^n \\ &= (x \cdot y)^{n+1}. \end{aligned}$$

Now let $n \in \mathbb{Z} \setminus \mathbb{N}_0$, i. e., $-n \in \mathbb{N}$. By Definition 3.5 and what we have already shown it follows that

$$x^n \cdot y^n = \big(x^{-1}\big)^{-n} \cdot \big(y^{-1}\big)^{-n} = \big(x^{-1} \cdot y^{-1}\big)^{-n} \overset{\text{Cor. 3.4(viii)}}{=} \big((x \cdot y)^{-1}\big)^{-n} = (x \cdot y)^n. \qquad \square$$

Notation 3.7. Let $(K, +, \cdot)$ a field, $A, B \subseteq K$ and $x \in K$. Then

$$\begin{aligned} x + A &:= \{x + a \,:\, a \in A\}, \\ xA &:= \{xa \,:\, a \in A\}, \\ A + B &:= \{a + b \,:\, a \in A,\ b \in B\}. \end{aligned}$$

Ordered fields

Definition 3.8. A field $(K, +, \cdot, >)$ is an *ordered field* if $(K, +, \cdot)$ is a field and the property "> 0" (positivity) is compatible with $+$ and $\cdot$, i. e., the *order axioms* hold:

(OA1) For all $x \in K$ *exactly one* of the following properties holds:

$$x > 0, \quad x = 0, \quad -x > 0,$$

(OA2) $x,\ y \in K,\ x > 0 \ \wedge\ y > 0 \implies x + y > 0.$

(OA3) $x,\ y \in K,\ x > 0 \ \wedge\ y > 0 \implies x \cdot y > 0.$

Let x, $y \in K$. Then

x is *positive* $:\iff x > 0$,
x is *negative* $:\iff x < 0$,
x is *non-negative* $:\iff x \geq 0 \ :\iff (x > 0 \ \vee\ x = 0)$
x is *non-positive* $:\iff x \leq 0 \ :\iff (x < 0 \ \vee\ x = 0)$.

$$\begin{aligned} x < y \ &:\iff x - y < 0, & x \leq y \ &:\iff x - y \leq 0, \\ x > y \ &:\iff x - y > 0, & x \geq y \ &:\iff x - y \geq 0. \end{aligned}$$

Usually the ordered field $(K, +, \cdot, >)$ is denoted by K.

Examples. $(\mathbb{Q}, +, \cdot, <)$ and $(\mathbb{R}, +, \cdot, <)$ are ordered fields.

The following rules are immediate consequences of the order axioms:

Corollary 3.9. *For elements a, x, x', y, y' in an ordered field $(K, +, \cdot, >)$ the following holds:*

(i) *Exactly one of the following holds: $x < y$, $x = y$, $x > y$.*

(ii) $x < y \ \wedge\ y < a \implies x < a$,

(iii) $x < y \implies x + a < y + a$,

(iv) $x < y \ \wedge\ x' < y' \implies x + x' < y + y'$,

(v) $x < y \ \wedge\ a > 0 \implies a \cdot x < a \cdot y$,
$x < y \ \wedge\ a < 0 \implies a \cdot x > a \cdot y$,

(vi) $0 \le x < y \ \wedge\ 0 \le x' < y' \implies 0 \le x' \cdot x < y' \cdot y$,

(vii) $x^2 > 0,\ x \neq 0$,

(viii) $x > 0 \implies x^{-1} > 0$,

(ix) $0 < x < y \implies 0 < y^{-1} < x^{-1}$,

(x) $1 > 0$.

Properties (i) and (ii) show that $(K, >)$ is a totally ordered set.

Proof. Property (i) is clear.
Proof of (ii): By assumption $y - x > 0$ and $a - y > 0$, therefore, by axiom (OA2), $a - x = (a - y) + (y - x) > 0$ which is equivalent to $x < a$.
Proof of (vii): If $x > 0$ then the assertion follows from axiom (OA3). If $x < 0$ then $x^2 = \underbrace{(-x)}_{>0}\underbrace{(-x)}_{>0} > 0$.
Proof of (viii): $x^{-1} = x \cdot x^{-1} \cdot x^{-1} > 0$ by (vii) and axiom (OA3).
Proof of (x): Follows from $1 = 1 \cdot 1$ and (vii).
For the proof of the other properties, see Exercise 3.1. □

Corollary 3.9 shows that the field $\mathbb{F}_2$ is not an ordered field since in $\mathbb{F}_2$ we have that $1 + 1 = 0 \not> 0$, in contradiction to property (iv). Actually, every ordered field must have infinitely many elements. Indeed, assume $(K, +, \cdot, >)$ is a finite ordered field. Since $1 > 0$ by (x), property (iv) yields (use induction): $0 < \sum_{j=1}^{n} 1,\ n \in \mathbb{N}$. On the other hand, since K is finite, there is an $m \in \mathbb{N}$ such that $\sum_{j=1}^{m} 1 = 0$. ✠

Definition 3.10. Let $(K, +, \cdot, >)$ be an ordered field.

$$K_+ := \{x \in K \,:\, x > 0\}, \quad K_+^0 := \{x \in K \,:\, x \ge 0\},$$
$$K_- := K \setminus K_+^0 = \{x \in K \,:\, x < 0\}.$$

For $x \in K$ define the *absolute value* (or *modulus*) of x by

$$\operatorname{abs}(x) := |x| := \begin{cases} x, & \text{if } x \ge 0, \\ -x, & \text{if } x < 0, \end{cases}$$

and the *sign of* x by

$$\operatorname{sign}(x) := \begin{cases} x/|x|, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

The functions $\operatorname{abs} : K \to K,\ x \mapsto \operatorname{abs}(x)$ and $\operatorname{sign} : K \to K,\ x \mapsto \operatorname{sign}(x)$ are called the *absolute value function* and *sign function*.

The next proposition states some elementary properties of the absolute value.

Proposition 3.11. *Let* $(K, +, \cdot, >)$ *be an ordered field and* $x, y \in K$*. Then:*

(i) $|x| \geq 0$, *and* $|x| = 0 \iff x = 0$,

(ii) $|-x| = |x|$,

(iii) $x \leq |x|$ *and* $-x \leq |x|$,

(iv) $|xy| = |x| \cdot |y|$,

(v) $\operatorname{sign}(x) \in \{-1, 0, 1\}$ *and* $\operatorname{sign}(x) \cdot \operatorname{abs}(x) = x$.

Fundamental is the so-called *triangle inequality*.

Theorem 3.12. *Let* $(K, +, \cdot, >)$ *be an ordered field and* $x, y \in K$*. Then*

$$|x + y| \leq |x| + |y|. \tag{3.1}$$

Proof. Since substituting x by $-x$ and y by $-y$ does not change the formula, we can assume without restriction that $x + y \geq 0$. Now

$$|x + y| = x + y \leq |x| + |y|. \qquad \square$$

Corollary 3.13. *Let* $(K, +, \cdot, >)$ *be an ordered field and* $x, y, z \in K$*. Then*

$$|x - y| \leq |x| + |y|, \quad |x + y| \geq \big||x| - |y|\big| \quad \text{and} \quad |x - y| \geq \big||x| - |y|\big|.$$

Proof. The first inequality follows directly from the triangle inequality

$$|x - y| = |x + (-y)| \leq |x| + |-y| = |x| + |y|.$$

In the same way, the third inequality follows from the second inequality. To prove the second inequality we not that $|x| = |x + y - y| \leq |x + y| + |y|$. Without restriction we can assume that $|x| \geq |y|$ because the assertion in symmetric in x and y. Therefore $|x+y| \geq |x| - |y| = \big||x| - |y|\big|$. $\square$

Theorem 3.14 (Bernoulli's inequality). *Let* $(K, +, \cdot, >)$ *be an ordered field. For* $x \in K$*,* $x \geq -1$ *and* $n \in \mathbb{N}_0$ *Bernoulli's inequality holds:*

$$(1 + x)^n \geq 1 + nx. \tag{3.2}$$

Proof by induction on n.
$\underline{n = 0}$: $(1 + x)^0 = 1 = 1 + 0 \cdot x$.
$\underline{n \curvearrowright n + 1}$: The induction hypothesis yields

$$\begin{aligned} (1 + x)^{n+1} &= \underbrace{(1 + x)}_{\geq 0} \overbrace{(1 + x)^n}^{\geq 1 + nx \text{ by ind.hyp.}} \geq (1 + x)(1 + nx) = 1 + (n + 1)x + \underbrace{nx^2}_{\geq 0} \\ &\geq 1 + (n + 1)x. \end{aligned} \qquad \square$$

Definition 3.15. (Least-upper-bound-property) An ordered field $(K, +, \cdot, >)$ has the *least-upper-bound-property* if every non-empty subset which has an upper bound has a supremum (least upper bound).

The least upper bound property is also called the (*Dedekind*) *completeness.*

Definition and Theorem 3.16. *Every ordered field* $(K, +, \cdot, >)$ *with the least-upper-bound-property has the so-called* Archimedean property

(AP) *For* $x, y \in K$*,* $x > 0$*, there exists an* $n \in \mathbb{N}_0$ *such that*

$$nx > y.$$

Proof. Assume there is no such n. Then the non-empty set

$$M = \{nx : n \in \mathbb{N}_0\}$$

is bounded by y hence, by the completeness assumption, $\sup M =: s$ exists. Since $0 < s - x < s$ there exists an $m_0 \in \mathbb{N}$ such that $m_0 x > s - x$ because s is the supremum of M (otherwise $s - x$ would be an upper bound of M which is smaller than s). This leads to $s < (m_0 + 1)x \in M$. ✠ □

Example. $(\mathbb{R}, +, \cdot, >)$ und $(\mathbb{Q}, +, \cdot, >)$ have the Archimedean property. There exist ordered fields without the Archimedean property.

The Archimedean property implies the following theorem.

Theorem 3.17. *Let* $(K, +, \cdot, >)$ *be an ordered field with the Archimedean property and* $x, \varepsilon \in K$*,* $\varepsilon > 0$.

(i) *If* $x > 1$*, then there exists an* $N \in \mathbb{N}$ *such that*

$$x^N > \varepsilon.$$

(ii) *If* $x < 1$*, then there exists an* $N \in \mathbb{N}$ *such that*

$$x^N < \varepsilon.$$

Proof. (i) Since $x > 1$ there exists an $y > 0$ such that $x = 1 + y$. Bernoulli's inequality yields

$$x^n = (1 + y)^n \geq 1 + ny, \quad n \in \mathbb{N}_0.$$

By the Archimedean property (AP) there exists an $N \in \mathbb{N}_0$ such that $Ny > \varepsilon$, also

$$x^N \geq 1 + Ny > 1 + \varepsilon > \varepsilon.$$

(ii): If $x \leq 0$, we may choose $N = 1$. Now let us assume that $0 < x < 1$. Since $x^{-1} > 1$ and $\varepsilon^{-1} > 0$ there exists an $N \in \mathbb{N}_0$ such that $0 < \varepsilon^{-1} < (x^{-1})^N = x^{-N}$ by (i), hence $x^N < \varepsilon$. □

3.2 The real numbers

It can be shown that up to isomorphy there is exactly one field $\mathbb{R}$ containing the natural numbers $\mathbb{N}$ satisfying

- the field axioms (A1)–(A4), (M1)–(M4), (D),

- the order axioms (OA1)–(OA3),
- least-upper-bound-property (Definition 3.15).

The existence of the real numbers will be shown in Section 4.6 using Cauchy sequences. Another way to construct the real numbers uses the so-called Dedekind cuts, see for instance [Rud76, Appendix to Chapter 1].

Remark. $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$, and the restriction of the order on $\mathbb{R}$ to $\mathbb{N}$ coincides with the order defined in Definition 2.3.

Special subsets of $\mathbb{R}$ are the intervals:

Definition 3.18. For $a,\, b \in \mathbb{R},\ a < b$, we define the sets

$$\begin{aligned} [a,\, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} && (\textit{closed interval}),\\ (a,\, b) &:= \{x \in \mathbb{R} : a < x < b\} && (\textit{open interval}),\\ [a,\, b) &:= \{x \in \mathbb{R} : a \leq x < b\} && (\textit{right half-open interval}),\\ (a,\, b] &:= \{x \in \mathbb{R} : a < x \leq b\} && (\textit{left half-open interval}). \end{aligned}$$

Remark. Let $a < b \in \mathbb{R}$. Then $\sup((a,\, b)) = \sup((a,\, b]) = \max((a,\, b]) = b$, (a, b) has no maximum.

Proposition 3.19. *For every $x \in \mathbb{R}_+$ there exists an $n \in \mathbb{N}_0$ with $n \leq x < n+1$.*

Proof. Exercise 3.3 □

Proposition 3.20. *Every interval in $\mathbb{R}$ contains a rational number.*

Proof. Exercise 3.3 □

For the proof of the next theorem we use the following

Remark. Let $0 < a < b \in \mathbb{R}$ and $n \in \mathbb{N}$, $n \geq 2$. Then

$$b^n - a^n < (b-a)nb^{n-1}. \tag{3.3}$$

Proof. Using $b - a > 0$ and $a^k < b^k$, $k \in \mathbb{N}_0$, we obtain

$$b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \cdots + ba^{n-2} + a^{n-1}) < (b-a)nb^{n-1}$$ □

Definition and Theorem 3.21. *For every real $x > 0$ and every $n \in \mathbb{N}$ there is exactly one real $y > 0$ such that $y^n = x$.*
This y is called the nth root of x denoted by $y =: \sqrt[n]{x} =: x^{\frac{1}{n}}$. In addition we define the nth root of 0 to be 0.

Proof. Uniqueness: Follows from Corollary 3.9 (vi): $0 \leq y_1 < y_2 \implies y_1^n < y_2^n$.
Existence: For $n = 1$ choose $y = x$. Now let $n \geq 2$. Let $A := \{t \in \mathbb{R} : t > 0,\ t^n \leq x\}$. The set A is not empty since it contains $t_0 := \frac{x}{1+x}$. (because $0 < t_0 < 1$ and therfore $t_0^n < t_0 < x$). Moreover, the set is bounded from above since $(1+x)^2 > 1 + x > x$. Since $\mathbb{R}$ has the least-upper-bound-property,

$$y := \sup A$$

exists. We want to show that $y^n = x$.

Step 1: Show $y^n \geq x$. Assume that $y^n < x$. Then there exists an $h \in \mathbb{R}$ such that

$$0 < h < \min\left\{1,\ \frac{x - y^n}{n(y+1)^{n-1}}\right\}.$$

The inequality (3.3) (with $a = y$ and $b = y + h$) yields

$$(y+h)^n - y^n < hn(y+h)^{n-1} < hn(y+1)^{n-1} < x - y^n.$$

Since $(y+h)^n < x$ it follows that $y + h \in A$ in contradiction to y being an upper bound of A.

Step 2: Show $y^n \leq x$. Assume that $y^n > x$. Then

$$k := \frac{y^n - x}{ny^{n-1}}.$$

satisfies $0 < k < y$. Inequality (3.3) yields that for all $t \geq y - k$

$$y^n - t^n \leq y^n - (y-k)^n < nky^{n-1} = y^n - x.$$

Therefore, $[y - k, \infty) \cap A = \emptyset$. Since y is an upper bound of A, also $y - k$ is an upper bound of A, in contradiction to y being the least upper bound of A.

Since we have shown that $y^n \leq x$ and $y^n \geq x$ it follows that $y^n = x$. □

The extended real line

Definition 3.22. The extended real line is $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ with the convention that $-\infty < x < \infty$ for all $x \in \mathbb{R}$.

$\overline{\mathbb{R}}$ is not a field but for $x \in \mathbb{R}$ we define

$$x + \infty = \infty + x = \infty, \quad x - \infty = -\infty + x = -\infty, \qquad \frac{x}{\infty} = \frac{x}{-\infty} = 0,$$
$$\infty \cdot x = x \cdot \infty = \begin{cases} \infty & \text{if } x > 0, \\ -\infty & \text{if } x < 0, \end{cases} \qquad -\infty \cdot x = x \cdot (-\infty) = \begin{cases} -\infty & \text{if } x > 0, \\ \infty & \text{if } x < 0. \end{cases}$$

For $a,\, b \in \mathbb{R}$ let

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

Definition 3.23. Let $A \subset \mathbb{R}$. We define

$$\sup A = \infty \qquad \text{if } A \text{ has no upper bound,}$$
$$\inf A = -\infty \qquad \text{if } A \text{ has no lower bound,}$$
$$\sup \emptyset = -\infty, \qquad \inf \emptyset = \infty.$$

3.3 The field of complex numbers

Definition 3.24. A *complex number* is an element $(a, b) \in \mathbb{R} \times \mathbb{R}$. The set of all complex numbers is denoted by $\mathbb{C}$. Two complex numbers $z_1 = (a_1, b_1)$ and $z_2 = (a_2, b_2)$ are equal if and only if $a_1 = a_2$ and $b_1 = b_2$.

On $\mathbb{C}$ the operations $+$ and $\cdot$ are defined by

$$+ : \mathbb{C}\times\mathbb{C}\to\mathbb{C},\quad (a,b)+(c,d) := (a+c, b+d),$$
$$\cdot : \mathbb{C}\times\mathbb{C}\to\mathbb{C},\quad (a,b)\cdot(c,d) := (ac-bd, ad+bc).$$

The *absolute value* (or *modulus*) of a complex number $z=(a,b)$ is

$$|z| := (a^2+b^2)^{\frac{1}{2}} \in \mathbb{R}.$$

The *complex conjugation* is the map

$$\mathbb{C}\to\mathbb{C},\quad z=(a,b)\mapsto \overline{z} := (a,-b).$$

For a complex number $z=(a,b)$ we set

$$\operatorname{Re} z := a \quad (\textit{real part of } z), \qquad \operatorname{Im} z := a \quad (\textit{imaginary part of } z).$$

Straightforward calculations show:

Proposition 3.25. $(\mathbb{C}, +, \cdot)$ *is a field with additive identity* $(0,0)$ *and multiplicative identity* $(1,0)$. *For* $z=(a,b)$ *the additive inverse is* $-z=(-a,-b)$. *If* $z\neq(0,0)$, *then its multiplicative inverse is* $z^{-1}=(a/|z|^2, -b/|z|^2)$.

Since $(a,0)+(b,0)=(a+b,0)$ and $(a,0)\cdot(b,0)=(a\cdot b,0)$ for all $a,b\in\mathbb{R}$, the field $\mathbb{R}$ can be identified with the subfield $\{(a,0) : a\in\mathbb{R}\}\subset\mathbb{C}$ via the field homomorphism $\mathbb{R}\to\mathbb{C},\ a\mapsto(a,0)$. Note that $|(a,0)|=\sqrt{a^2}=|a|$ in agreement with the definition of the absolute value of real numbers.

Definition 3.26. $\mathrm{i} := (0,1)$.

Theorem 3.27. (i) $\mathrm{i}^2=-1$,

(ii) $(a,b)=a+\mathrm{i}b$, $a,b\in\mathbb{R}$.

Proof. (i) $\mathrm{i}^2=(0,1)^2=(0,1)(0,1)=(0-1,0+0)=(-1,0)=-1$.

(ii) $a+\mathrm{i}b=(a,0)+(0,1)(b,0)=(a,0)+(0,b)=(a,b)$. □

The preceding theorem shows that calculations with complex numbers can be carried out as for real numbers if we take into account that $\mathrm{i}^2=-1$.

The next proposition collects often used properties of complex numbers.

Proposition 3.28. *Let* $z,w\in\mathbb{C}$.

(i) $\overline{\overline{z}}=z$, $z=\overline{z} \iff \operatorname{Im} z=0$, $z=-\overline{z} \iff \operatorname{Re} z=0$,

(ii) $\overline{z+w}=\overline{z}+\overline{w}$,

(iii) $\overline{zw}=\overline{z}\,\overline{w}$, *in particular* $\overline{z}^{-1}=\overline{z^{-1}}$ *for* $z\neq0$,

(iv) $z+\overline{z}=2\operatorname{Re} z$, $\quad z-\overline{z}=2\mathrm{i}\operatorname{Im} z$,

(v) $z\overline{z}=|z|^2\geq0$ *and* $|z|=0 \iff z=0$,

(vi) $|z|=|\overline{z}|$,

(vii) $|zw|=|z|\,|w|$, *in particular* $|z^{-1}|=|z|^{-1}$ *for* $z\neq0$.

(viii) $|\operatorname{Re} z| \leq |z|, \quad |\operatorname{Im} z| \leq |z|$,

(ix) $|z+w| \leq |z| + |w|$.

Proof. (i), (ii), (iii), (iv), (v) and (vi) are easy to check.
For the proof of (vii), let $z = a + \mathrm{i}b$, $w = c + \mathrm{i}d$ with $a, b, c, d \in \mathbb{R}$. Then

$$\begin{aligned} |zw|^2 &= |(a+\mathrm{i}b)(c+\mathrm{i}d)|^2 = |(ac-bd) + \mathrm{i}(ad+bc)|^2 = (ac-bd)^2 + (ad+bc)^2 \\ &= a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 = (a^2+b^2)(c^2+d^2) = |z|^2\,|w|^2. \end{aligned}$$

Taking the square root yields the assertion.
The assertion about the real part of z in (viii) follows from

$$|\operatorname{Re} z|^2 = |a|^2 = a^2 \leq a^2 + b^2 = |z|^2.$$

The assertion about the imaginary part is proved analogously.
In order to prove the triangle inequality in (ix), note that $w\overline{z} = \overline{\overline{w}z}$, hence

$$\begin{aligned} |z+w|^2 &= (z+w)(\overline{z}+\overline{w}) = z\overline{z} + w\overline{w} + z\overline{w} + w\overline{z} = |z|^2 + |w|^2 + z\overline{w} + \overline{z\overline{w}} \\ &= |z|^2 + |w|^2 + 2\operatorname{Re}(z\overline{w}) \leq |z|^2 + |w|^2 + 2|\operatorname{Re}(z\overline{w})| \\ &\leq |z|^2 + |w|^2 + 2|z\overline{w}| = |z|^2 + |w|^2 + 2|z|\,|w| = (|z| + |w|)^2. \end{aligned}$$

The assertion follows by taking the square root. □

Remark. Note that $\mathbb{C}$ cannot be ordered because $\mathrm{i}^2 = -1 < 0$ (cf. Corollary 3.9 (vii)).

Chapter 4

Sequences and Series

In this chapter, the notion of convergence is introduced, one of the most important concepts in analysis. To this end, it is necessary to consider the distance between points in a given set which leads to the definition of metric spaces. Next we deal with sequences in metric spaces and give criteria for convergence and divergence. If, in addition, the metric space is equipped with the structure of a vector space compatible with the given metric, then relationship between the arithmetics and properties of sequences can be established.

4.1 Metric spaces

A metric space is a set of points X together with a function $X \times X \to \mathbb{R}$ that measures the distance between two points in X and satisfies the properties that are expected from a distance.

Definition 4.1. Let X be a set. A *metric on X* is a function

$$d : X \times X \to \mathbb{R}, \quad (x, y) \mapsto d(x, y),$$

such that

(i) $d(x, y) = 0 \iff x = y$,

(ii) $d(x, y) = d(y, x),\ x, y \in X$ (*symmetry*),

(iii) $d(x, y) \leq d(x, z) + d(z, y),\ x, y, z \in X.$ (*triangle inequality*).

Then, (X, d) is called a *metric space* and $d(x, y)$ is the *distance* between the points $x, y \in X$.

Note that the definition of d implies

$$d(x, y) \geq 0, \quad x, y \in X,$$

since the triangle inequality yields $0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$.

Examples. (i) Any set X with $|X| \leq 1$,

(ii) $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ with $d(x, y) = |x - y|$ are metric spaces. If not stated explicitely otherwise, we always consider $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ as equipped with this metric.

(iii) Let $\mathbb{F} = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ and $n \in \mathbb{N}$. Then $\mathbb{F}^n$ with the *Euclidean metric*

$$d(x,y) = \sqrt{|x_1 - y_1|^2 + \cdots + |x_n - y_n|^2}, \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{F}^n,$$

is a metric space. Note that for $n = 1$ the euclidian metric coincides with the metric defined in (iii).

(iv) The set X with the *discrete metric*

$$d(x,y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

Remark. Let (X,d) be a metric space and $Y \subseteq X$ a subset. Then $(Y, d|_{Y \times Y})$ is also a metric space.

Special subsets of metric spaces are open and closed balls.

Definition 4.2. Let (X,d) be a metric space. For $a \in X$ and $r \in \mathbb{R}_+$ we define

$$B_r(a) := \{x \in X \, : \, d(a,x) < r\} =: \textit{open ball with centre at } a \textit{ and radius } r,$$
$$K_r(a) := \{x \in X \, : \, d(a,x) \leq r\} =: \textit{closed ball with centre at } a \textit{ and radius } r.$$

Example. In the special case of $\mathbb{R}$ the open balls are exactly the open intervals, and the closed balls are the closed intervals.

Definition 4.3. Let (X,d) be a metric space. For a subset $M \subseteq X$

$$\operatorname{diam} M := \sup\{d(x,y) \, : \, x, y \in M\}$$

is the *diameter of* M. M is called *bounded* if $\operatorname{diam} M < \infty$.

Remark. (Exercise 4.1)

- M bounded $\iff \exists\, a \in X,\ r > 0 : \ M \subset B_r(a)$.
- A subset $M \subseteq \mathbb{R}$ is bounded in the sense of Definition 1.1 (as a subset of an ordered set) if and only if it is bounded in the sense of Definition 4.3 (as a subset of a metric space).
- A subset $M \subseteq \mathbb{R}$ is bounded if and only if there exist $a, \in \mathbb{R}$ and $r > 0$ such that $M \subseteq B_r(a)$.

4.2 Sequences in metric spaces

Definition 4.4. Let (X,d) be a metric space. A *sequence in* X is a map

$$\mathbb{N} \to X, \quad n \mapsto x_n \in X.$$

The sequence is usually denoted by

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

The x_n are called *terms* of the sequence.

The important properties of the domain of a sequence are that it is countable and ordered. Therefore, instead of the *index set* $\mathbb{N}$ any subset $M = \{m, m+1, m+2, \dots\} \subseteq \mathbb{Z}$ can be used as domain of a sequence (note that there is an order preserving bijection between $\mathbb{N}$ and M).

Since a sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X,d) defines the set $\{x_n : n \in \mathbb{N}\} \subseteq X$, one writes $(x_n)_{n\in\mathbb{N}} \subseteq X$.

Examples (Sequences in $\mathbb{R}$).

- $x_n := a$, $n \in \mathbb{N}$ for some $a \in \mathbb{R}$: $(a, a, a, \dots)$ (constant sequence),
- $x_n = \frac{1}{n}$, $n \in \mathbb{N}$: $(1, \frac{1}{2}, \frac{1}{3}, \dots)$,
- $x_n = x^n$, $n \in \mathbb{N}$, for a fixed $x \in \mathbb{R}$: $(x, x^2, x^3, \dots)$,
- $x_n = (-1)^n$, $n \in \mathbb{N}$: $(-1, 1, -1, \dots)$,
- $x_0 = 0$, $x_1 = 1$, $x_{n+1} = x_{n-1} + x_n$, $n \in \mathbb{N}$: $(0, 1, 1, 2, 3, 5, \dots)$ (*Fibonacci sequence*).

Definition 4.5. Let (X,d) be a metric space. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is said to be *convergent* if and only if

$$\exists\, a \in X \quad \forall \varepsilon > 0 \quad \exists\, N \in \mathbb{N} \quad \forall n \geq N \quad d(x_n, a) < \varepsilon.$$

The sequence is then said to *converge to* a, and a is called the *limit* of $(x_n)_{n\in\mathbb{N}}$ denoted by

$$\lim_{n\to\infty} x_n = a, \quad \text{or} \quad x_n \xrightarrow{n\to\infty} a, \quad \text{or} \quad x_n \to a,\ n \to \infty. \tag{4.1}$$

A sequence is said to be *divergent* if it does not converge.
The sequence $(x_n)_{n\in\mathbb{N}}$ is said to be *bounded* if $\{x_n : n \in \mathbb{N}\}$ is bounded in X.

(Here and in the following, a statement like $\varepsilon > 0$ always means $\varepsilon \in \mathbb{R},\ \varepsilon > 0$.)

The definition says that a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges to $a \in X$ if and only if for every $r > 0$ *almost all* (i. e.all with exception of finitely many) x_n lie in $B_r(a)$.

The next theorem justifies the notation $a = \lim_{n\to\mathbb{N}} x_n$ in (4.1).

Theorem 4.6 (Uniqueness of the limit). *The limit of a convergent sequence in a metric space is unique.*

Proof. Let (X,d) be a metric space and $(x_n)_{n\in\mathbb{N}}$ a convergent sequence in X. Let $a, b \in X$ such that $x_n \to a$ and $x_n \to b$ for $n \to \infty$ and $a \neq b$. Then $d(a,b) > 0$ and there exist N_a, $N_b \in \mathbb{N}$ such that

$$d(x_n, a) < \frac{d(a,b)}{2},\ n \geq N_a, \qquad d(x_n, b) < \frac{d(a,b)}{2},\ n \geq N_b.$$

Let $N = \max\{N_a, N_b\}$. Then the triangle inequality yields for $n \geq N$ the contradiction

$$d(a,b) \leq d(x_n, a) + d(x_n, b) < \frac{d(a,b)}{2} + \frac{d(a,b)}{2} = d(a,b). \qquad \square$$

Examples 4.7. Consider some of the sequences in $\mathbb{R}$ of the example at the beginning of this section:

(i) $x_n = a$, $n \in \mathbb{N}$, for some $a \in \mathbb{R}$: $(x_n)_{n\in\mathbb{N}}$ is bounded and $\lim_{n\to\infty} x_n = a$.

(ii) $x_n = \frac{1}{n}$, $n \in \mathbb{N}$. The sequence $(x_n)_{n\in\mathbb{N}}$ is bounded and converges to 0.

Proof. The sequence is bounded because $\{\frac{1}{n} \,:\, n \in \mathbb{N}\} \subseteq B_2(0)$. To prove the convergence, let $\varepsilon > 0$. By the Archimedean property, there exists an $N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon}$. It follows that

$$|x_n - 0| = \frac{1}{n} \le \frac{1}{N} < \varepsilon, \quad n \ge N. \qquad \square$$

(iii) $x_n = (-1)^n,\ n \in \mathbb{N}$. The sequence $(x_n)_{n\in\mathbb{N}}$ is bounded and divergent.

Proof. The sequence is bounded since $\{(-1)^n \,:\, n \in \mathbb{N}\} = \{-1, 1\} \subseteq B_2(0)$. Let $0 < \varepsilon < \frac{1}{2}$ and assume that $(x_n)_{n\in\mathbb{N}}$ converges to some $a \in \mathbb{R}$. Then there exists an $N \in \mathbb{N}$ such that $d(x_n, a) < \varepsilon$, $n \ge N$. By the triangle inequality, it follows that

$$2 = d(x_N, x_{N+1}) \le d(x_N, a) + d(x_{N+1}, a) < \frac{1}{2} + \frac{1}{2} = 1. \qquad \maltese$$

Therefore $((-1)^n)_{n\in\mathbb{N}}$ does not converge in $\mathbb{R}$. $\square$

(iv) $\displaystyle\lim_{n\to\infty} \frac{n}{n+1} = 1$. (Exercise)

(v) $\displaystyle\lim_{n\to\infty} \frac{n}{2^n} = 0$. (see Exercise 4.3)

Theorem 4.8. *Every convergent sequence in a metric space is bounded.*

Note that not every bounded sequence converges as Example 4.7 (iii) shows.

Proof of Theorem 4.8. Let (X, d) be a metric space, $(x_n)_{n\in\mathbb{N}} \subseteq X$ a convergent sequence and let $a := \lim_{n\to\mathbb{N}} x_n$. Then there exists an $N \in \mathbb{N}$ such that

$$d(x_n, a) < 1, \quad n \ge N.$$

Let $R := \max\{d(a, x_1),\, d(a, x_2),\, \dots,\, d(a, x_{N-1})\} + 1$. Then $d(a, x_n) < R$ for all $n \in \mathbb{N}$, hence $(x_n)_{n\in\mathbb{N}} \subseteq B_R(a)$. $\square$

Definition 4.9. Let (X, d) be a metric space. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is called a *Cauchy sequence in X* if and only if

$$\forall \varepsilon > 0 \quad \exists\, N \in \mathbb{N} \quad \forall n, m \ge N \quad d(x_n, x_m) < \varepsilon.$$

Theorem 4.10. *Every convergent sequence in a metric space is a Cauchy sequence.*

Proof. Let (X, d) be a metric space, $(x_n)_{n\in\mathbb{N}} \subseteq X$ a convergent sequence and $a := \lim_{n\to\infty} x_n$. Let $\varepsilon > 0$. Then there exists a $N \in \mathbb{N}$ such that

$$d(x_n, a) < \frac{\varepsilon}{2}, \qquad n \ge N.$$

Therefore, by the triangle inequality,

$$d(x_n, x_m) \le d(x_n, a) + d(x_m, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \qquad n, m \ge N. \qquad \square$$

Note that the converse is not true.

Example. Consider the metric spaces $(\mathbb{R}, d)$ and the subspace $((0,1), d|_{(0,1)})$ where d is the usual metric on $\mathbb{R}$. Let $x_n = \frac{1}{n}$, $n \in \mathbb{N}$. We already showed that $(x_n)_{n\in\mathbb{N}}$ converges to 0 in $\mathbb{R}$, hence it is also a Cauchy sequence. Since $(x_n)_{n\in\mathbb{N}} \subseteq (0,1)$ and the metric on $(0,1)$ is a restriction of the metric on $\mathbb{R}$, the sequence is also a Cauchy sequence in $((0,1), d|_{(0,1)})$. But the sequence does not converge in $((0,1), d|_{(0,1)})$. Indeed, if it would converge to some $a \in (0,1)$, then it would converge to a also in the metric space $(\mathbb{R}, d)$. The uniqueness of the limit would imply $a = 0$, in contradiction to $a \notin (0,1)$.

Definition 4.11. A metric space in which every Cauchy sequence converges is called a *complete metric space.*

Examples 4.12.

- $\mathbb{Q}$ is not a complete metric space.
- The metric spaces $\mathbb{R},\ \mathbb{C},\ \mathbb{R}^n,\ \mathbb{C}^n$ are complete.

The completeness of $\mathbb{R}$ is equivalent to the least-upper-bound-property on $\mathbb{R}$.

Theorem 4.13. *Every Cauchy sequence in a metric space is bounded.*

Proof. Let (X, d) be a metric space and $(x_n)_{n\in\mathbb{N}} \subseteq X$ a Cauchy sequence. Then there exists an $N \in \mathbb{N}$ such that

$$d(x_n, x_m) < 1, \qquad n, m \geq N.$$

Hence, by the triangle inequality,

$$d(x_1, x_n) \leq d(x_1, x_N) + d(x_N, x_n) < d(x_1, x_N) + 1.$$

Let $R := \max\{d(x_1, x_2),\ d(x_1, x_3),\ \dots,\ d(x_1, x_N)\} + 1$. Then $(x_n)_{n\in\mathbb{N}} \subseteq B_R(x_1)$ which implies the assertion. $\square$

Definition 4.14. Let (X, d) be a metric space, $(x_n)_{n\in\mathbb{N}} \subseteq X$ a sequence in X and $\rho : \mathbb{N} \to \mathbb{N}$ such that $\rho(n) < \rho(n+1)$, $n \in \mathbb{N}$. Then $(x_{\rho(n)})_{n\in\mathbb{N}}$ is called a *subsequence* of $(x_n)_{n\in\mathbb{N}}$.

Theorem 4.15. *Let (X, d) be a metric space and $(x_n)_{n\in\mathbb{N}} \subseteq X$.*

(i) *If $(x_n)_{n\in\mathbb{N}}$ converges, then every subsequence converges and has the same limit.*

(ii) *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and contains a convergent subsequence, then it converges.*

Proof. (i): Let $(x_{\rho(n)})_{n\in\mathbb{N}}$ be a subsequence of the convergent sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ and let $a := \lim_{n\to\infty} x_n$. Let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $d(x_n, a) < \varepsilon$, $n \geq N$. Now choose $M \in \mathbb{N}$ such that $\rho(M) \geq N$. Since $\rho(n) \geq N$ for all $n \geq M$, it follows that $d(x_{\rho(n)}, a) < \varepsilon$, $n \geq M$.

(ii) Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a Cauchy sequence with the convergent subsequence $(x_{\rho(n)})_{n\in\mathbb{N}}$. Let $a := \lim_{n\in\mathbb{N}} x_{\rho(n)}$ and $\varepsilon > 0$. By assumption, there exists an $K \in \mathbb{N}$ such that $d(x_{\rho(k)}, a) < \frac{\varepsilon}{2}$ for $k \geq K$ and an $M \in \mathbb{N}$ such that $d(x_n, x_m) < \frac{\varepsilon}{2}$, $n, m > M$.
Let $N := \max\{K, M\}$. Then, using that $\rho(k) \geq k$ for all $k \in \mathbb{N}$, we obtain

$$d(x_n, a) \leq d(x_n, x_{\rho(N)}) + d(x_{\rho(N)}, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \qquad n \geq \rho(N).$$ $\square$

4.3 Sequences in normed spaces

Next we want to consider metric spaces with the additional structure of a vector space such that the metric is compatible with the algebraic structure.

Definition 4.16. Let $\mathbb{F}$ be a field. A set V is called an $\mathbb{F}$-*vector space* if there are operations

$$\begin{aligned} +: V \times V \to V, \quad (x, y) \mapsto x + y, \quad x, y \in V \qquad & (\textit{Addition}), \\ \cdot\ : \mathbb{F} \times V \to V, \quad (\lambda, x) \mapsto \lambda \cdot x, \quad \lambda \in \mathbb{F},\ x \in V \qquad & (\textit{scalar multiplication}), \end{aligned}$$

satisfying the following axioms:

Axioms of vector space addition

(VS1) $x + (y + z) = (x + y) + z, \quad x, y, z \in V$,

(VS2) $x + y = y + x, \quad x, y \in V$,

(VS3) $\exists\, 0_V \in V :\ \forall\, x \in V\ x + 0_V = x$,

(VS4) $\forall\, x \in V\ \exists - x \in V :\ x + (-x) = 0_V$.

Axioms of scalar multiplication

(VS5) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y, \quad \lambda \in \mathbb{F},\ x, y \in V$,

(VS6) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x, \quad \lambda, \mu \in \mathbb{F},\ x \in V$,

(VS7) $\lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x, \quad \lambda, \mu \in \mathbb{F},\ x \in V$,

(VS8) $1 \cdot x = x, \quad x \in V$.

The elements of V are called *vectors*, the elements of $\mathbb{F}$ are called *scalars*. It is custom to write λx instead of $\lambda \cdot x$ for $\lambda \in \mathbb{F}$ and $x \in V$.

Corollary 4.17. *Let V be a $\mathbb{F}$-vector space. Then:*

(i) *0_V and $-x$ are uniquely determined,*

(ii) *$0 \cdot x = 0_V, \quad x \in V$,*

(iii) *$(-1) \cdot x = -x, \quad x \in V$.*

Proof. (i) Analogously to the proof of uniqueness of the additive identity in fields (Corollary 3.4).
(ii) Let $\lambda = 1,\ \mu = 0 \in \mathbb{F}$ and $x \in V$ arbitrary. By axiom (VS6) it follows that

$$x \overset{\text{(VS8)}}{=} 1 \cdot x = (1 + 0) \cdot x \overset{\text{(VS6)}}{=} 1 \cdot x + 0 \cdot x \overset{\text{(VS8)}}{=} x + 0 \cdot x.$$

Therefore $0 \cdot x = 0_V$ by the uniqueness of 0_V shown in (i).
(iii) Let $\lambda = 1,\ \mu = -1$ and $x \in V$ arbitrary. Then

$$0_V \overset{\text{(ii)}}{=} 0 \cdot x = (1 - 1) \cdot x \overset{\text{(VS6)}}{=} 1 \cdot x + (-1) \cdot x, \overset{\text{(VS8)}}{=} x + (-1) \cdot x.$$

Therefore $(-1) \cdot x = -x$ by the uniqueness of $-x$ shown in (i). ☐

Examples.
- Every field $\mathbb{F}$ is an $\mathbb{F}$-vector space.

- Let $\mathbb{F}$ be a field and $n = 1$. For $x = (x_j)_{j=1}^n,\ y = (y_j)_{j=1}^n \in \mathbb{F}^n = \mathbb{F} \times \cdots \times \mathbb{F}$ and $\lambda \in \mathbb{F}$ let

$$\begin{aligned} x + y &:= (x_j + y_j)_{j=1}^n = (x_1 + y_1, \ldots, x_n + y_n), \\ \lambda \cdot x &:= (\lambda x_j)_{j=1}^n = (\lambda x_1, \ldots, \lambda x_n). \end{aligned}$$

  It is easy to check that $\mathbb{F}^n$ is a $\mathbb{F}$-vector space. For $n = 1$ this vector space coincides with the vector space above.

- $\mathbb{C}$ is a $\mathbb{R}$-vector space; $\mathbb{R}$ is a $\mathbb{Q}$-vector space.

- Let $\mathbb{F}$ be a field, $X \neq \emptyset$ a set and denote the set of all functions $f : X \to \mathbb{F}$ by $\mathbb{F}^X$. For $f, g \in \mathbb{F}^X$ and $\lambda \in \mathbb{F}$ define

$$\begin{aligned} f + g : X \to \mathbb{F}, &\quad (f+g)(x) = f(x) + g(x), && x \in X, \\ \lambda \cdot f : X \to \mathbb{F}, &\quad (\lambda \cdot f)(x) = \lambda f(x), && x \in X. \end{aligned}$$

  Then $\mathbb{F}^X$ is a $\mathbb{F}$-vector space.

Now we want to equip a vector space V with a metric that is compatible with the algebraic structure on V.

Definition 4.18. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and V a $\mathbb{F}$-vector space. A *norm* on V is a map

$$\|\cdot\| : V \to \mathbb{R}, \quad x \mapsto \|x\|$$

such that

(i) $\|x\| = 0 \iff x = 0, \quad x \in V,$

(ii) $\|\lambda x\| = |\lambda|\,\|x\|, \quad \lambda \in \mathbb{F},\ x \in V,$

(iii) $\|x + y\| \leq \|x\| + \|y\|, \quad x, y \in V.$

Then $(V, \|\cdot\|)$ is called a *normed space*.

Remark. Instead of $\mathbb{R}$ or $\mathbb{C}$, $\mathbb{F}$ can be any field with a norm in the sense above. Usually we always deal with $\mathbb{R}$- or $\mathbb{C}$-vector spaces.

Immediately from the definition of a normed spaces follows the following proposition.

Proposition 4.19. *Every normed space $(V, \|\cdot\|)$ is a metric space with the metric*

$$d(x, y) := \|x - y\|, \quad x, y \in V.$$

In particular, $\|x\| = d(x, 0) \geq 0$, $x \in V$.

Using the proposition above, convergent sequences and Cauchy sequences are also defined in normed spaces. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in the normed space $(V, \|\cdot\|)$. Then

- $(x_n)_{n\in\mathbb{N}}$ converges to $a \in V :\iff \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\ \ \|x_n - a\| < \varepsilon.$

- $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in V

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

Definition 4.20. A normed space in which every Cauchy sequence converges is called a *Banach space*.

In the following, $\mathbb{F}$ is always assumed to be $\mathbb{R}$ or $\mathbb{C}$.

Examples.
- Every ordered field is a normed space with the absolute value as norm.
- $\mathbb{Q}$ with the norm $\|x\| = |x|$ is a normed space but it is not complete.
- $\mathbb{R}$ and $\mathbb{C}$ with the norm $\|x\| = |x|$ are Banach spaces.
- If $V = \mathbb{R}^n$ or $\mathbb{C}^n$, $n \in \mathbb{N}$ and

$$\|x\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \quad x = (x_j)_{j=1}^n \in \mathbb{F}^n,$$

then $(\mathbb{F}^n, \|\cdot\|)$ are Banach spaces. The norm $\|\cdot\|$ is called the *Euclidean norm* on $\mathbb{F}^n$.

Analogously to Corollary 3.13 for ordered fields the following lemma can be shown:

Lemma 4.21. *Let $(V, \|\cdot\|)$ be a normed space. Then*

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \leq \|x\| + \|y\|, \quad x, y \in V.$$

Proposition 4.22. *Let $(V, \|\cdot\|)$ be a normed space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.*

(i) *$(x_n)_{n\in\mathbb{N}}$ Cauchy sequence in V $\implies$ $(\|x_n\|)_{n\in\mathbb{N}}$ Cauchy sequence in $\mathbb{R}$.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ converges in V $\implies$ $(\|x_n\|)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$.*
In this case: $\|\lim_{n\to\infty} x_n\| = \lim_{n\to\infty} \|x_n\|$.

Proof. (i) Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, there exists an $N \in \mathbb{N}$ such that

$$\big|\|x_n\| - \|x_m\|\big| \leq \|x_n - x_m\| < \varepsilon, \quad m, n \geq N.$$

(ii) Let $\varepsilon > 0$. Let $\varepsilon > 0$ and let $\lim_{n\to\infty} x_n := a$. Then there is an $N \in \mathbb{N}$ such that

$$\big|\|x_n\| - \|a\|\big| \leq \|x_n - a\| < \varepsilon, \quad n \geq N.$$

□

Note that in both cases the converse direction is wrong as the example $((-1)^n)_{n\in\mathbb{N}}$ shows. Moreover, any normed space over a non-complete field $\mathbb{F}$ is not complete.

Proposition 4.23. *Let $(V, \|\cdot\|)$ be a normed space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $(x_n)_{n\in\mathbb{N}}$ a sequence in V.*

(i) $\lim\limits_{n\to\infty} x_n = 0 \iff \lim\limits_{n\to\infty} \|x_n\| = 0,$

(ii) $\lim\limits_{n\to\infty} x_n = a \iff \lim\limits_{n\to\infty} (x_n - a) = 0, \iff \lim\limits_{n\to\infty} \|x_n - a\| = 0,$

(iii) *If there exists a sequence $(\lambda_n)_{\lambda\in\mathbb{N}} \subseteq \mathbb{F}$ and an $N_0 \in \mathbb{N}$, such that $\lambda_n \to 0$, $n \to \infty$ and*

$$\|x_n\| \leq |\lambda_n|, \quad n \geq N_0,$$

then $\lim\limits_{n\to\infty} x_n = 0.$

Proof. (i) and (ii) are immediate consequences of Proposition 4.22. For the proof of (iii) fix an $\varepsilon > 0$. Then there is an $N \in \mathbb{N}$ such that $|\lambda_n| < \varepsilon$, $n \geq N$. Hence

$$\|x_n\| \leq |\lambda_n| < \varepsilon, \quad n \geq \max\{N_0, N\}.$$

□

Next we show that the algebraic operations on a normed space and taking limits are compatible.

Theorem 4.24. *Let $(V, \|\cdot\|)$ be a normed space over a field $\mathbb{F}$ and let $\lambda \in \mathbb{F}$.*

(i) *If $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ are Cauchy sequences in V, then so are*

$$(a_n + b_n)_{n\in\mathbb{N}} \quad \textit{and} \quad (\lambda a_n)_{n\in\mathbb{N}}.$$

(ii) *If $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ are convergent sequences in V, then so are*

$$(a_n + b_n)_{n\in\mathbb{N}} \quad \textit{and} \quad (\lambda a_n)_{n\in\mathbb{N}},$$

and

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n, \qquad \lim_{n\to\infty} \lambda a_n = \lambda \lim_{n\to\infty} a_n.$$

Proof. (i) Let $\varepsilon > 0$. Then there exist $N_a \in \mathbb{N}$ and $N_b \in \mathbb{N}$ such that

$$\begin{aligned} \|a_n - a_m\| &< \frac{\varepsilon}{2}, \qquad m, n \geq N_a, \\ \|b_n - b_m\| &< \frac{\varepsilon}{2}, \qquad m, n \geq N_b. \end{aligned}$$

For $m, n \geq \max\{N_a, N_b\}$ it follows that

$$\|(a_n + b_n) - (a_m + b_m)\| \leq \|a_n - a_m\| + \|b_n - b_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

hence $(a_n + b_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.
If $\lambda = 0$, then $(\lambda a_n)_{n\in\mathbb{N}} = (0)_{n\in\mathbb{N}}$ a constant sequence and therefore a Cauchy sequence. Now let $\lambda \neq 0$ and $\varepsilon > 0$. Then there exist $N \in \mathbb{N}$ such that $\|a_n - a_m\| < \frac{\varepsilon}{|\lambda|}$ for all $m, n \geq N$, hence

$$\|\lambda a_n - \lambda a_m\| = |\lambda| \|a_n - a_m\| < |\lambda| \frac{\varepsilon}{|\lambda|} = \varepsilon, \qquad m, n \geq N.$$

(ii) is proved similarly. □

Example. The sequence $(x_n)_{n\in\mathbb{N}}$ where $x_n := \frac{n+1}{n}$, $n \in \mathbb{N}$, converges to 1.

Proof. Since the constant sequence $(1)_{n\in\mathbb{N}}$ and the sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ converge, we have that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{1}{n} = 1 + 0 = 1.$$ □

Theorem 4.25. *Let $(V, \|\cdot\|)$ be a normed space over the field $\mathbb{F}$ with norm $|\cdot|$.*

(i) *If $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ and $(x_n)_{n\in\mathbb{N}} \subseteq V$ are Cauchy sequences, then so is*

$$(\lambda_n x_n)_{n\in\mathbb{N}} \subseteq V.$$

(ii) *If $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ and $(x_n)_{n\in\mathbb{N}} \subseteq V$ are convergent, then so is $(\lambda_n x_n)_{n\in\mathbb{N}} \subseteq V$ and*

$$\lim_{n\to\infty}(\lambda_n x_n) = (\lim_{n\to\infty} \lambda_n)(\lim_{n\to\infty} x_n).$$

(iii) *If the sequences $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ and $(x_n)_{n\in\mathbb{N}} \subseteq V$ are bounded and at least one of them converges to 0, then sequence of the products converges to 0.*

Proof. (i) Since the sequences $(\lambda_n)_{n\in\mathbb{N}}$ and $(x_n)_{n\in\mathbb{N}}$ are Cauchy sequences, they are bounded (Theorem 4.13). Let $R_x, R_\lambda \in \mathbb{R}$ such that

$$\|x_n\| \le R_x, \quad |\lambda_n| \le R_\lambda, \qquad n \in \mathbb{N}.$$

For $\varepsilon > 0$ choose $N_x, N_\lambda \in \mathbb{N}$ such that

$$\|x_n - x_m\| < \frac{\varepsilon}{2R_\lambda}, \quad m, n \ge N_x, \qquad \text{and} \qquad |\lambda_n - \lambda_m| < \frac{\varepsilon}{2R_x}, \quad m, n \ge N_\lambda.$$

For all $m, n \ge \max\{N_x,\, N_\lambda\}$ it follows that

$$\begin{aligned}\|\lambda_n x_n - \lambda_m x_m\| &= \|\lambda_n(x_n - x_m) - (\lambda_m - \lambda_n)x_m\| \\ &\le |\lambda_n|\,\|x_n - x_m\| + |\lambda_m - \lambda_n|\,\|x_m\| < R_\lambda \frac{\varepsilon}{2R_\lambda} + R_x \frac{\varepsilon}{2R_x} = \varepsilon.\end{aligned}$$

(ii) is proved similarly.

(iii) is proved similarly. Let R_x, R_λ as in the proof of (i). Without restriction we assume that $x_n \to 0$, $n \to \infty$. Therefore

$$\begin{aligned} & \|\lambda_n x_n\| = |\lambda_n|\,\|x_n\| \le R_\lambda\,\|x_n\| \to 0, && n \to \infty, \\ \implies\ & \|\lambda_n x_n\| \to 0, && n \to \infty, \quad \text{(by Prop. 4.23 (iii))} \\ \implies\ & \lambda_n x_n \to 0, && n \to \infty, \quad \text{(by Prop. 4.23 (i))} \end{aligned}$$ □

Example. Let $x_n := (-1)^n \frac{n!}{n^n} \in \mathbb{R}$, $n \in \mathbb{N}$. Since $0 < \frac{m}{n} < 1$ for $m = 1, \dots, n-1$ it follows that

$$|x_n| = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{1}{n}.$$

Since $(\frac{1}{n})_{n\in\mathbb{N}}$ converges to 0 by Example 4.7 (ii), Theorem 4.25 (iii) yields that $(x_n)_{n\in\mathbb{N}}$ converges to zero.

Theorem 4.26. *Let $(V, \|\cdot\|)$ be a normed space over a field $\mathbb{F}$ with norm $|\cdot|$. Let $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ and $(x_n)_{n\in\mathbb{N}} \subseteq V$ be convergent sequences such that $\lim_{n\to\infty} \lambda_n \ne 0$. Then there exists an $N_0 \in \mathbb{N}$ such that $\lambda_n \ne 0$, $n \ge N_0$ and the sequence $(\frac{1}{\lambda_n} x_n)_{n=N_0}^\infty$ converges with*

$$\lim_{n\to\infty} \frac{1}{\lambda_n} x_n = \Big(\frac{1}{\lim\limits_{n\to\infty} \lambda_n}\Big)\big(\lim_{n\to\infty} x_n\big).$$

Proof. Let $a := \lim_{n\to\infty} \lambda_n \ne 0$. Then there exists an $N_0 \in \mathbb{N}$ such that $|\lambda_n - a| < \frac{|a|}{2}$, $n \ge N_0$, hence, by the triangle inequality,

$$|\lambda_n| \ge |a| - |\lambda - a| \ge |a| - \frac{|a|}{2} = \frac{|a|}{2} > 0, \qquad n \ge N_0.$$

Let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that

$$|\lambda_n - a| \le \frac{\varepsilon |a|^2}{2}, \qquad n \ge N.$$

Therefore we have for all $n \ge \max\{N_0, N\}$

$$\Big|\frac{1}{\lambda_n} - \frac{1}{a}\Big| = \frac{1}{|a||\lambda_n|}\big|a - \lambda_n\big| \le \frac{1}{|a|\frac{|a|}{2}} \frac{\varepsilon |a|^2}{2} = \varepsilon.$$

This shows

$$\lim_{n\to\infty} \frac{1}{\lambda_n} = \frac{1}{\lim_{n\to\infty} \lambda_n}.$$

The assertion of the theorem follows now by Theorem 4.25(ii). □

Remark. Important special cases of Theorem 4.24, Theorem 4.25 and Theorem 4.26 are when $V = \mathbb{R}$ or $V = \mathbb{C}$. For example, Theorem 4.25 shows that

$$(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \subseteq \mathbb{C} \text{ convergent} \implies (a_n b_n)_{n\in\mathbb{N}} \subseteq \mathbb{C} \text{ convergent and } \lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n).$$

Example. Let $x_n := \dfrac{n^3+n^2}{7n^3+12n-1}$, $n \in \mathbb{N}$. Then $\lim\limits_{n\to\infty} x_n = \dfrac{1}{7}$.

Proof. Since $x_n = \frac{1+n^{-1}}{7+12n^{-2}-n^{-3}}$, $n \in \mathbb{N}$, and the limits $\lim_{n\to\infty} \frac{1}{n}$, $\lim_{n\to\infty} \frac{1}{n^2}$ and $\lim_{n\to\infty} \frac{1}{n^3}$ exist and are equal to zero, Theorem 4.24, Theorem 4.25 and Theorem 4.26 yield

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{1+n^{-1}}{7+12n^{-2}-n^{-3}} = \frac{1+0}{7+0-0} = \frac{1}{7}. \qquad \square$$

Theorem 4.27. (i) *Let $m \in \mathbb{N}$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $(\mathbb{F}^m, \|\cdot\|)$ with the Euclidean norm $\|\cdot\|$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{F}^m$ with $x_n = (x_{1,n}, \dots, x_{m,n})$, $n \in \mathbb{N}$. Then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{F}^m$ if and only if for all $j = 1, \dots, m$ the sequences $(x_{j,n})_{n\in\mathbb{N}}$ are Cauchy sequences in $\mathbb{F}$. The sequence $(x_n)_{n\in\mathbb{N}}$ is convergent in $\mathbb{F}^m$ if and only if for all $j = 1, \dots, m$ the sequences $(x_{j,n})_{n\in\mathbb{N}}$ are convergent in $\mathbb{F}$. In this case*

$$\lim_{n\to\mathbb{N}} x_n = (\lim_{n\to\mathbb{N}} x_{1,n}, \dots, \lim_{n\to\mathbb{N}} x_{m,n}).$$

(ii) *Let $(z_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$ and $x_n := \operatorname{Re} z_n$, $y_n := \operatorname{Im} z_n$, $n \in \mathbb{N}$. Then $(z_n)_{n\in\mathbb{N}}$ is a Cauchy sequence if and only if both $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are Cauchy sequences in $\mathbb{R}$ and $(z_n)_{n\in\mathbb{N}}$ is convergent if and only if both $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are convergent in $\mathbb{R}$. In this case*

$$\lim_{n\to\mathbb{N}} z_n = \lim_{n\to\mathbb{N}} x_n + \mathrm{i} \lim_{n\to\mathbb{N}} y_n.$$

In particular, it follows that $\mathbb{C}$ and $\mathbb{F}^m$ are complete since $\mathbb{R}$ is complete.

4.4 Sequences in an ordered field

Let $\mathbb{F}$ be an ordered field. As in Definition 3.23 we can extend the order on $\mathbb{F}$ to an order on $\overline{\mathbb{F}} = \mathbb{F} \cup \{-\infty, \infty\}$. The most important example is, of course, $\mathbb{F} = \mathbb{R}$.

Definition 4.28. Let $\mathbb{F}$ be an ordered field. We say that a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ *diverges to* ∞, in formula $\lim_{n\to\infty} x_n = \infty$, if and only if

$$\forall R \in \mathbb{F} \;\; \exists N \in \mathbb{N} \;\; \forall n \geq N \;\; x_n > R.$$

The sequence $(x_n)_{n\in\mathbb{N}}$ *diverges to* $-\infty$ if and only if $(-x_n)_{n\in\mathbb{N}}$ diverges to ∞, in formula $\lim_{n\to\infty} x_n = -\infty$.

Remark 4.29.

- $\lim_{n\to\mathbb{N}} x_n = \infty \implies \{x_n : n \in \mathbb{N}\} \subseteq \mathbb{F}$ is not bounded from above.
- $\lim_{n\to\mathbb{N}} x_n = -\infty \implies \{x_n : n \in \mathbb{N}\} \subseteq \mathbb{F}$ is not bounded from below.
- The converse is not true: For example, the sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ with $x_n = (1+(-1)^n)n$ does not diverge to ∞ but $\{x_n : n \in \mathbb{N}\} = 2\mathbb{N}_0$ is unbounded above.
- $\lim_{n\to\infty} x_n = -\infty \iff \forall R \in \mathbb{F} \;\; \exists N \in \mathbb{N} \;\; \forall n \geq N \;\; x_n < R.$

Proposition 4.30. *Let $\mathbb{F}$ be an ordered field and $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$.*

(i) *If $\lim_{n\to\infty} |x_n| = \infty$, then there exists an $N \in \mathbb{N}$ such that $x_n \neq 0$, $n \geq N$ and $\lim_{n\to\infty} x_n^{-1} = 0$.*

(ii) *If $x_n > 0,\ n \in \mathbb{N},$ and $\lim_{n\to\infty} x_n = 0,$ then $\lim_{n\to\infty} x_n^{-1} = \infty$.*
If $x_n < 0,\ n \in \mathbb{N},$ and $\lim_{n\to\infty} x_n = 0,$ then $\lim_{n\to\infty} x_n^{-1} = -\infty$.
If $x_n \neq 0,\ n \in \mathbb{N},$ and $\lim_{n\to\infty} x_n = 0,$ then $\lim_{n\to\infty} |x_n^{-1}| = \infty$.

(iii) *If there exists a sequence $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ such that $\lim_{n\to\mathbb{N}} \lambda_n = \infty$ and an $N \in \mathbb{N}$ such that $x_n > \lambda_n,\ n \geq N,$ then $\lim_{n\to\mathbb{N}} x_n = \infty$.*

Theorem 4.31. *Let $\mathbb{F}$ be and ordered field and $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ convergent sequences. Assume that there exists an $N_0 \in \mathbb{N}$ such that*

$$x_n \leq y_n, \quad n \geq N_0. \tag{4.2}$$

Then $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.

Proof. Assume $\lim_{n\to\infty} x_n > \lim_{n\to\infty} y_n$. Then

$$0 < \lim_{n\to\infty} x_n - \lim_{n\to\infty} y_n = \lim_{n\to\infty} (x_n - y_n).$$

and there exists an $N \in \mathbb{N}$ such that

$$x_n - y_n \geq \frac{1}{2} \lim_{n\to\infty} (x_n - y_n) > 0, \quad n \geq N, \tag{4.3}$$

(see proof of Theorem 4.26). Hence we obtain the contradiction

$$0 \overset{(4.2)}{\geq} x_{N+N_0} - y_{N+N_0} \overset{(4.3)}{>} 0.$$ □

Remark. Even if condition (4.2) is substituted by $x_n < y_n$, $n \in \mathbb{N}$, we cannot conclude $\lim_{n\to\infty} x_n < \lim_{n\to\infty} y_n$, as the example $x_n = \frac{1}{n}$ and $y_n = 0$, $n \in \mathbb{N}$, shows.

Corollary 4.32. *Let $\mathbb{F}$ be an ordered field, $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ a convergent sequence. Assume that there exist $N_0 \in \mathbb{N}$ and $\alpha, \beta \in \mathbb{F}$ such that $\alpha \leq x_n \leq \beta,\ n \geq N_0$. Then*

$$\alpha \leq \lim_{n\to\infty} x_n \leq \beta.$$

Corollary 4.33 (Sandwich lemma). *Let $\mathbb{F}$ be an ordered field, $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ convergent sequences in $\mathbb{F}$ with*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = a.$$

Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ and $N \in \mathbb{N}$ such that

$$a_n \leq x_n \leq b_n, \quad n \geq N.$$

Then also $(x_n)_{n\in\mathbb{N}}$ converges and $\lim_{n\to\infty} x_n = a$.

Definition 4.34. Let $\mathbb{F}$ be an ordered field. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ is called

(i) *monotonically increasing* $\iff$ $x_{n+1} \geq x_n,\ n \geq 1,$

(ii) *strictly monotonically increasing* $\iff$ $x_{n+1} > x_n,\ n \geq 1.$

(iii) (*strictly*) *monotonically decreasing*, if $(-x_n)_{n\in\mathbb{N}}$ is (strictly) monotonically increasing.

(iv) (*strictly*) *monotonic*, if it is either (strictly) monotonically increasing or (strictly) monotonically decreasing.

Not every convergent sequence is monotonic, and not every monotonic sequence is convergent or bounded.

Theorem 4.35. *Let $\mathbb{F}$ be a complete ordered field and $(x_n)_{n\in\mathbb{N}}$ a monotonic sequence in $\mathbb{F}$. Then*

$$(x_n)_{n\in\mathbb{N}} \text{ is convergent} \quad \iff \quad (x_n)_{n\in\mathbb{N}} \text{ is bounded.}$$

Proof. "$\Longrightarrow$" is shown in Theorem 4.8 (every convergent sequence in a metric space is bounded).

"$\Longleftarrow$" Without restriction, we may assume that $(x_n)_{n\in\mathbb{N}}$ is increasing. Since $(x_n)_{n\in\mathbb{N}}$ is bounded and $\mathbb{F}$ is complete, it follows that

$$a := \sup\{x_n : n \in \mathbb{N}\} < \infty.$$

Let $\varepsilon > 0$. By Exercise 3.4 it follows that there exists an $N \in \mathbb{N}$ such that $a - \varepsilon < x_N \le a$. Using that $x_n \le a$, $n \in \mathbb{N}$, and the monotonicity of the sequence we find

$$|x_n - a| = a - x_n \le a - x_N < \varepsilon, \qquad n \ge N,$$

which implies the convergence of $(x_n)_{n\in\mathbb{N}}$. □

Corollary 4.36. *Let $\mathbb{F}$ be a complete ordered field and $(x_n)_{n\in\mathbb{N}}$ a bounded monotonic sequence in $\mathbb{F}$. Then $(x_n)_{n\in\mathbb{N}}$ converges and*

$$\lim_{n\to\infty} x_n = \begin{cases} \sup\{x_n : n \in \mathbb{N}\}, & \text{if } (x_n)_{n\in\mathbb{N}} \text{ is increasing,} \\ \inf\{x_n : n \in \mathbb{N}\}, & \text{if } (x_n)_{n\in\mathbb{N}} \text{ is decreasing.} \end{cases}$$

In Theorem 3.21 we showed that for $x > 0$ and $k \in \mathbb{N}$ there exists exactly one solution of the equation $y^k = x$ but the proof is not constructive, i. e., it gives no rule how to find y. The following example gives a constructive proof.

Example 4.37 (kth root in $\mathbb{R}$). Let $x > 0$ and $k \in \mathbb{N}$. Define the sequence $(x_n)_{n\in\mathbb{N}}$ recursively by

$$x_0 = x + 1, \quad x_{n+1} = x_n \left(1 - \frac{x_n^k - x}{k x_n^k}\right), \quad n \in \mathbb{N}.$$

(i) The sequence $(x_n)_{n\in\mathbb{N}}$ converges.

(ii) $\lim_{n\to\infty} x_n = \sqrt[k]{x}$.

Proof. (i) We show the following by induction on n:

(a) $x_n > 0,\ n \in \mathbb{N}$, (b) $x_n \le x_{n-1},\ n \in \mathbb{N}$, (c) $x_n^k \ge x,\ n \in \mathbb{N}$.

<u>$n = 0$:</u> (a) and (c) are clearly satisfied, and for (b) there is nothing to prove.

<u>$n \curvearrowright n+1$:</u>

(a) $x_{n+1} > 0$ because $x_n > 0$ and $kx_n^k - (x_n^k - x) = (k-1)x_n^k + x > 0$.

(b) $x_n^k \ge x$ by induction hypothesis. Therefore $1 - \frac{x_n^k - x}{kx_n} < 1$ and hence $x_{n+1} = x_n\big(1 - \frac{x_n^k - x}{kx_n^k}\big) \le x_n$.

(c) Note that $-\frac{x_n^k - x}{kx_n} = \frac{x - x_n^k}{kx_n} > -\frac{x_n^k}{kx_n^k} = -\frac{1}{k} > -1$. Therefore, Bernoulli's inequality (3.2) shows

$$x_{n+1}^k = x_n^k \left(1 - \frac{x_n^k - x}{kx_n^k}\right)^k \geq x_n^k \left(1 - k\frac{x_n^k - x}{kx_n^k}\right) = x_n^k \left(1 - k\frac{x_n^k - x}{kx_n^k}\right) = x.$$

(a) and (b) imply that $(x_n)_{n\in\mathbb{N}}$ is a bounded monotonic sequence, hence it converges by Theorem 4.35.

(ii) From the definition of the x_n it follows that

$$kx_n^{k-1}x_{n+1} = (k-1)x_n^k + x. \tag{4.4}$$

Let $y := \lim_{n\to\infty} x_n$. Taking the limit on both sides in (4.4) shows that

$$ky^{k-1}y = (k-1)y^k + x,$$

hence $y^k = x$. □

Another important example is the definition of Euler's number.

Example 4.38. The limit

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

exists and is called *Euler's number* (e = 2,71828182...).

Proof. See Exercise 4.12. □

Not every sequence in an ordered field is monotonic, but every sequence contains a monotonic subsequences.

Theorem 4.39. *In an ordered field $\mathbb{F}$ every sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$ contains a monotonic subsequence.*

Proof. We call an x_n a "low" if $x_n \leq x_m$ for all $m \geq n$. There are two possible cases:

Case 1: The sequence contains infinitely many low terms. Then the subsequence which consists of all low terms is monotonically increasing.

Case 2: The sequence contains only finitely many low terms. Then there exists a $N \in \mathbb{N}$ such for all $n \geq N$ the term x_n is not low. Hence for every $n \geq N$ there exists an $m > n$ such that $x_m < x_n$ because x_n is not low. Let $n_1 := N$. Since x_{n_1} is not a low term of the sequence, there exists an $n_2 > n_1$ such that $x_{n_2} < x_{n_1}$. Inductively, we can find $n_1 < n_2 < n_3 < \dots$ such that $x_{n_1} > x_{n_2} > x_{n_3} > \dots$. The sequence $(x_{n_k})_{k\in\mathbb{N}}$ is a monotonically decreasing subsequence of $(x_n)_{n\in\mathbb{N}}$. □

Theorem 4.40 (Bolzano-Weierstraß).

(i) *Every bounded sequence in $\mathbb{R}$ contains a convergent subsequence.*

(ii) *Every bounded sequence in $\mathbb{C}$ contains a convergent subsequence.*

Proof. (i) By Theorem 4.39 every sequence contains a monotonic subsequence. Since $\mathbb{R}$ is complete and every subsequence of a bounded sequence is bounded, this subsequences must converge by Theorem 4.35.

(ii) Let $(z_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{C}$ and let $x_n := \operatorname{Re} z_n$ and $y_n := \operatorname{Im} z_n$. Then $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are bounded sequences in $\mathbb{R}$ (by Proposition 3.28). By (i) there exists a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. Again by (i), $(y_{n_k})_{k\in\mathbb{N}}$ contains a convergent subsequence $(y_{n_{k_m}})_{m\in\mathbb{N}}$. Therefore, $(x_{n_{k_m}} + \mathrm{i}\, y_{n_{k_m}})_{m\in\mathbb{N}}$ is convergent subsequence of $(z_n)_{n\in\mathbb{N}}$. □

Remark. The Bolzano-Weierstraß theorem is equivalent to the completeness of $\mathbb{R}$.

Definition 4.41. Let $(x_n)_{n\in\mathbb{N}}$ be sequence in a metric space X. A value $a \in X$ is called a *cluster value of* $(x_n)_{n\in\mathbb{N}}$ if there exists a subsequence that converges to a.

In addition, for an ordered field ∞ is a cluster value of a sequence $(x_n)_{n\in\mathbb{N}}$ in the field, if it contains a subsequence that diverges to ∞ and $-\infty$ is a cluster value of the sequence if it contains a subsequence that diverges to $-\infty$.

Remark. The Bolzano-Weierstraß Theorem implies

(i) that every sequence in $\mathbb{R}$ contains either a convergent subsequence or a subsequence that diverges to ∞ or $-\infty$,

(ii) that every sequence in $\mathbb{R}$ has a cluster value.

Remark 4.42. (i) If a sequence in a metric space converges, then it has exactly one cluster value. The reverse is not true.

(ii) A bounded sequence in $\mathbb{R}$ or $\mathbb{C}$ is convergent if and only if it has exactly one cluster value.

(iii) Let $x = (x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space. Then a is a cluster value of x if and only if for each $\varepsilon > 0$ the ball $B_\varepsilon(a)$ contains infinitely many terms of the sequence, that is, there are infinitely many $n \in \mathbb{N}$ such that $x_n \in B_\varepsilon(a)$. In formula: $\#\{n \in \mathbb{N} : x_n \in B_\varepsilon(a)\}) = \infty$. Note, however, that $\#(B_\varepsilon(a) \cap \{x_n | n \in \mathbb{N}\}) < \infty$ is possible as the example $((-1)^n)_{n\in\mathbb{N}}$ shows.

Proof. (iii) Let a be a cluster value of x. Then it has a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges to a. For given $\varepsilon > 0$ there exist an $K \in \mathbb{N}$ such that $d(a, x_{n_k}) < \varepsilon$, $k > K$, hence $x_{n_k} \in B_\varepsilon(a)$ for every $k > K$.

Assume now that for every $\varepsilon > 0$ infinitely many x_n lie in $B_\varepsilon(a)$. Then we can choose inductively $n_1 < n_2 < \dots$ such that $x_{n_k} \in B_{\frac{1}{k}}$, $k \in \mathbb{N}$, i. e., $d(x_{n_k}, a) < \frac{1}{k}$, $k \in \mathbb{N}$. Hence a is a cluster value of the sequence because the subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converges to a. □

Definition 4.43. Let $\mathbb{F}$ be a complete ordered field and $(x_n)_{n\in\mathbb{N}}$ a sequence in $\mathbb{F}$. The *limes superior* and *limes inferior*

$$\limsup_{n\to\infty} x_n := \overline{\lim}\, x_n := \inf\{x \in \mathbb{F} : x_n \leq x \text{ for almost all } x_n\},$$
$$\liminf_{n\to\infty} x_n := \underline{\lim}\, x_n := \sup\{x \in \mathbb{F} : x_n \geq x \text{ for almost all } x_n\}.$$

Proposition 4.44. *Let $\mathbb{F}$ be a complete ordered field and $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{F}$. Then $\limsup_{n\to\infty} x_n$ is the greatest cluster value of $(x_n)_{n\in\mathbb{N}}$ and $\liminf_{n\to\infty} x_n$ is the smallest cluster value of $(x_n)_{n\in\mathbb{N}}$.*

Proof. We show only the assertion for $a := \limsup_{n\to\infty} x_n$. If $a = \infty$, then the sequence contains a subsequence which diverges to ∞, hence ∞ is a cluster value. Obviously, it is the largest cluster value. If $a = -\infty$, then the sequence diverges to $-\infty$, hence $-\infty$ is the only cluster value.

Now assume $a \in \mathbb{R}$. First we show that a is the greatest accumulation point. Let $\varepsilon > 0$. Then $x_n \leq a + \frac{\varepsilon}{2}$ for almost all all x_n. Hence, only finitely many x_n lie in $B_{\frac{\varepsilon}{2}}(a + \varepsilon)$. Therefore, by Remark 4.42, $a + \varepsilon$ cannot be a cluster value.

We have to show that a is cluster value of $(x_n)_{n\in\mathbb{N}}$. If a is not a cluster value, then there exists an $\varepsilon > 0$ such that at most finitely many x_n lie in $B_\varepsilon(a)$. Since in addition only finitely many x_n are larger than $a + \epsilon/2$, it follows that $x_n \le a - \frac{\varepsilon}{2}$ for almost all x_n.

To prove the assertion for lim inf, we only need to apply the claim to the sequence $(-x_n)_{n\in\mathbb{N}}$ and observe that $\liminf x_n = -\limsup(-x_n)$ and that the largest accumulation point of $(x_n)_{n\in\mathbb{N}}$ is equal to the negative o f the smallest accumulation point of $(-x_n)_{n\in\mathbb{N}}$. □

Corollary 4.45. *A sequence $(x_n)_{n\in\mathbb{N}}$ is convergent if and only if* $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

Remark. Another characterisation of lim sup and lim inf is the following: For the sequence $(x_n)_{n\in\mathbb{N}}$ define the sequences $(y_n^+)_{n\in\mathbb{N}}$ and $(y_n^-)_{n\in\mathbb{N}}$ in $\mathbb{F}\cup\{\pm\infty\}$ by

$$y_n^+ := \sup\{x_k : k \ge n\}, \qquad y_n^- := \inf\{x_k : k \ge n\}.$$

Then $(y_n^+)_{n\in\mathbb{N}}$ is monotonically decreasing and $(y_n^-)_{n\in\mathbb{N}}$ is monotonically increasing (and therefore convergent in $\mathbb{F}\cup\{\pm\infty\}$) and

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} y_n^+, \qquad \liminf_{n\to\infty} x_n = \lim_{n\to\infty} y_n^-.$$

Proof. Exercise 4.8. Obviously, the sequence $(y_n^+)_{n\in\mathbb{N}}$ is monotonically decreasing and $x_k \le y_n^+$ for $k \ge n$.

Case 1. $(x_k)_{k\in\mathbb{N}}$ unbounded from above. Then $y_n^+ = \infty$, $n \in \mathbb{N}$, and $\limsup_{n\to\infty} x_n = \infty = \lim_{n\to\infty} y_n^+$.

Case 2. $(x_k)_{k\in\mathbb{N}}$ bounded from above and has no cluster value $a \in \mathbb{R}$. Then the sequence $(x_n)_{n\in\mathbb{N}}$ diverges to $-\infty$, i.e., for every $R \in \mathbb{R}$ exists $n_R \in \mathbb{N}$ such that $x_n \le R$ for all $n \ge n_R$. Hence also $y_n^+ \le n_R$ for all $n \ge n_R$ and therefore $\lim_{n\to\infty} y_n^+ = -\infty = \limsup_{n\to\infty} x_n$.

Case 3. $(x_k)_{k\in\mathbb{N}}$ bounded from above and has at least one cluster value $a \in \mathbb{R}$. Then $y_n^+ \ge a$, $n \in \mathbb{N}$, and $(y_n^+)_{n\in\mathbb{N}}$ converges by Theorem 4.35 (since the sequence is bounded and monotonic). Let $y := \lim_{n\to\infty} y_n^+$. First we show that y is the greatest cluster value of $(x_n)_{n\in\mathbb{N}}$. Let $b > y$ and $\varepsilon := b - y$. Since $(y_n^+)_{n\in\mathbb{N}}$ is decreasing, there exists $N \in \mathbb{R}$ such that $y_n^+ < b - \frac{\epsilon}{2}$, $n \ge N$, but then also $x_n \le y_n^+ < b - \frac{\epsilon}{2}$, $n \ge N$. In particular, $B_{\frac{\varepsilon}{2}}(b) \cap \{x_n : n \ge N\} = \emptyset$. This implies that b cannot be cluster value of $(x_n)_{n\in\mathbb{N}}$.

Next we show that y is a cluster value. To this end we construct a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ which converges to y. By the definition of y_1 there exists an n_1 such that $0 \le y_1 - x_{n_1} < \frac{1}{1}$ (use Exercise 3.4). Now assume that n_k, $k = 1, \dots, m$, are chosen such that $n_k < n_{k+1}$ and $0 \le y_{n_{k-1}} - x_{n_k} < \frac{1}{k}$. By the definition of y_{n_m} there exists an n_{m+1} such that $|y_m - x_{n_{m+1}}| < \frac{1}{m+1}$. Since the sequences $(y_{n_k})_{k\in\mathbb{N}}$ and $(y_{n_k} - x_{n_k})_{k\in\mathbb{N}}$ converge, also $(x_{n_k})k \in \mathbb{N}$ converges and

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty}(x_{n_k} - y_{n_k}) + y_{n_k} = \lim_{k\to\infty}(x_{n_k} - y_{n_k}) + \lim_{k\to\infty} + y_{n_k} = 0 + y = y.$$

Alternative proof of Case 3. If $(x_k)_{k\in\mathbb{N}}$ is bounded from above and has at least one cluster value $a \in \mathbb{R}$, then $\alpha := \limsup_{n\to\infty} x_n \in \mathbb{R}$. Since α is an accumulation point of $(x_k)_{k\in\mathbb{N}}$, there exists a subsequence $(x_{k_m})_{m\in\mathbb{N}}$ which converges to α. In particular, α is a lower bound for $(y_k^+)_{k\in\mathbb{N}}$ because $y_k^+ \ge \sup\{x_{k_m} : k_m \ge k\} \ge \lim_{m\to\infty} x_{k_m} = \alpha$ for all $k \in \mathbb{N}$. Since $(y_k^+)_{k\in\mathbb{N}}$ is decreasing, it is convergent and $\lim_{k\to\infty} y_k \ge \alpha$.

Now let $\epsilon > 0$. Since α is the largest accumulation point of $(x_n)_{n\in\mathbb{N}}$, there must be $N \in \mathbb{N}$ such that $x_n \le \alpha + \epsilon$ for all $n \ge N$ (otherwise there would be a subsequence $(x_{k_m})_{m\in\mathbb{N}}$ with $x_{k_m} \ge \alpha + \epsilon$ for all $m \in \mathbb{N}$ and this subsequence must have an accumulation point $\ge \alpha + \epsilon > \alpha$). Consequently, $y_n = \sup\{x_k : k \ge n\} \le \alpha + \epsilon$ para todo $n \ge N$. It follows that $\lim_{n\to\infty} y_n \le \alpha + \epsilon$. Since this is true for all $\epsilon > 0$, we actually have $\lim_{n\to\infty} y_n \le \alpha$.

The claim for lim inf can be proved analogously (or can be deduced from the claim non lim sup applied to $(-x)_{n\in\mathbb{N}}$). □

4.5 Series

4.5.1 Basic criteria of convergence and series in $\mathbb{R}$

Definition 4.46. Let $(V, \|\cdot\|)$ be a normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$. Then we define the *partial sums*

$$s_n := \sum_{k=1}^{n} x_k, \qquad n \in \mathbb{N}.$$

The sequence $(s_n)_{n\in\mathbb{N}}$ is a *series* in V, denoted by $\sum_{k=1}^{\infty} x_n$. The series is *convergent* if the sequence of the partial sums is convergent. In this case, $s := \lim_{n\to\infty} s_n$ exists and we write

$$\sum_{k=1}^{\infty} x_n = s.$$

Otherwise the series is called *divergent*. In the special case where V is an ordered field, we write

$$\sum_{k=1}^{\infty} x_n = \pm\infty$$

if $\lim_{n\to\infty} s_n = \pm\infty$.

Remark.

- $\sum_{k=1}^{\infty} x_n$ is not a sum but the limit of a sequence.
- The symbol $\sum_{k=1}^{\infty} x_n$ has two meanings: it denotes the sequence of the partial sums, and it denotes its limit if is exists.

Example. $\sum_{n=1}^{\infty} 1 + \dfrac{1}{n}$ diverges.

Proof. $s_n := \sum_{k=1}^{n} 1 + \frac{1}{k} \geq n(1 + \frac{1}{n}) \geq n + 1$. Therefore the sequence of the partial sums diverges to ∞. □

Example. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n!}$ converges. (See Exercise 4.12.)

Theorem 4.47. *Let $(V, \|\cdot\|)$ be a normed space over a field $\mathbb{F}$, $\lambda \in \mathbb{F}$ and $\sum_{n=1}^{\infty} x_n$, $\sum_{n=1}^{\infty} y_n$ convergent series in V. Then $\sum_{n=1}^{\infty} \lambda x_n + y_n$ converges.*

Proof. Apply Theorem 4.24 to the sequences of the partial sums. □

Theorem 4.48. *Let $(V, \|\cdot\|)$ be a complete normed space and $(x_n)_{n\in\mathbb{N}} \subseteq V$.*

(i) *Cauchy criterion for series:*

$$\sum_{n=1}^{\infty} x_n \quad \text{converges} \quad \iff \quad \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq m \geq N\ \left\| \sum_{k=m}^{n} x_n \right\| < \varepsilon.$$

(ii) $$\sum_{n=1}^{\infty} x_n \quad \text{converges} \quad \Longrightarrow \quad \lim_{n\to\infty} x_n = 0.$$

Proof. Since V is a complete normed space, the series converges if and only if the sequence of the partial sums $s_n := \sum_{k=1}^{n} x_k$ is a Cauchy sequence. This is the case if and only if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for $m-1, n \geq N$, without restriction $m-1 < n$,

$$\|s_n - s_{m-1}\| = \Big\| \sum_{k=m}^{n} x_n \Big\| < \varepsilon.$$

In particular, it follows that

$$\|x_n\| = \|s_n - s_{n-1}\| < \varepsilon, \qquad n \geq N,$$

so also the second part of the theorem is proved. □

Note that $\lim_{n\to\infty} x_n = 0$ does not implies the convergence of the series $\sum_{n=1}^{\infty} x_n$ as the following example shows:

Example 4.49 (Harmonic series). $$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Proof. Let $s_n := \sum_{k=1}^{n} \frac{1}{k}$. Then $(s_n)_{n\in\mathbb{N}}$ is not a Cauchy sequence since

$$|s_{2n} - s_n| = \sum_{k=n+1}^{2n} \frac{1}{k} \geq \sum_{k=n+1}^{2n} \frac{1}{2n} = \frac{1}{2}.$$

Therefore, the harmonic series diverges. Since it is monotonically increasing, it follows that is diverges to ∞.

□

Example 4.50 (Geometric series). Let $z \in \mathbb{C}$.

(i) $|z| \geq 1 \implies \displaystyle\sum_{n=0}^{\infty} z^n$ diverges.

(ii) $|z| < 1 \implies \displaystyle\sum_{n=0}^{\infty} z^n$ converges and $\displaystyle\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$.

Proof. If $|z| \geq 1$ then $|z|^n = |z^n|$ does not converge to 0, hence the sum cannot converge by Theorem 4.47.
Now let $|z| < 1$ and let $s_n := \sum_{k=0}^{n} z^k$. Then

$$(1-z)s_n = (1-z)\sum_{k=0}^{n} z^k = \sum_{k=0}^{n} z^k + \sum_{k=1}^{n+1} z^k = 1 - z^{n+1}.$$

Since $|z| < 1$, we have that $z \neq 1$ and $\lim_{n\to\infty} z^n = 0$. So we obtain

$$s_n = \frac{1 - z^{n+1}}{1-z} \;\to\; \frac{1}{1-z}, \qquad n \to \infty. \tag{4.5}$$

□

Theorem 4.51. *Let $(x_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$, $x_k \geq 0$, $k \in \mathbb{N}$, and define $s_n := \sum_{k=1}^{n} x_k$, $n \in \mathbb{N}$. Then*

$$\sum_{k=1}^{\infty} x_k \textit{ converges} \quad \iff \quad (s_n)_{n\in\mathbb{N}} \textit{ bounded.}$$

Proof. The theorem follows immediately from Theorem 4.35 since the sequence $(s_n)_{n\in\mathbb{N}}$ is monotonically increasing. □

Example 4.52. The series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^2}$ converges.

Proof. Since $\frac{1}{k^2} \geq 0$ for all $k \in \mathbb{N}$, it suffices to show that the sequence of the partial sums $s_n := \sum_{k=1}^{n}$ is bounded. This follows from

$$\begin{aligned}0 \leq s_n - 1 &= \sum_{k=2}^{n}\frac{1}{k^2} \leq \sum_{k=2}^{n}\frac{1}{k(k-1)} = \sum_{k=2}^{n}\frac{1}{k-1} - \frac{1}{k} = \sum_{k=2}^{n}\frac{1}{k-1} - \sum_{k=2}^{n}\frac{1}{k}\\ &= 1 - \frac{1}{n} \leq 1.\end{aligned}$$

□

Definition 4.53. A series in $\mathbb{R}$ is called *alternating* if it is of the form

$$\pm\sum_{n=0}^{\infty}(-)^n x_n$$

with $x_n \geq 0$, $n \in \mathbb{N}$.

Theorem 4.54 (Leibniz criterion). *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a monotonically decreasing sequence such that $x_n \geq 0$, $n \in \mathbb{N}$, and $\lim_{n\to\infty} x_n = 0$. Then*

$$s := \sum_{n=0}^{\infty}(-)^n x_n$$

exists and $|s - s_n| \leq x_{n+1}$ where $s_n := \sum_{k=0}^{n}(-)^k x_k$, $n \in \mathbb{N}$.

Proof. First we show that the subsequences $(s_{2n})_{n\in\mathbb{N}}$ and $(s_{2n+1})_{n\in\mathbb{N}}$ converge. For all $n \in \mathbb{N}$

$$s_{2n} \geq s_{2n} - x_{2n+1} + x_{2n+2} = s_{2n+2}, \tag{4.6}$$
$$s_{2n+1} \leq s_{2n+1} - x_{2n+2} + x_{2n+3} = s_{2n+3}, \tag{4.7}$$
$$s_{2n} \geq s_{2n} - x_{2n+1} = s_{2n+1} \overset{(4.6)}{\geq} s_1, \tag{4.8}$$
$$s_{2n+1} \leq s_{2n+1} + x_{2n+2} = s_{2n+2} \overset{(4.7)}{\leq} s_0. \tag{4.9}$$

By (4.6) and (4.8) the sequence $(s_{2n})_{n\in\mathbb{N}}$ is monotonically decreasing and bounded from below, hence convergent by Theorem 4.35. Analogously, using (4.7) and (4.9), it follows that the sequence $(s_{2n+1})_{n\in\mathbb{N}}$ is convergent.
Let $a := \lim_{n\to\infty} x_{2n}$.

$$\lim_{n\to\infty} s_{2n+1} = \lim_{n\to\infty}(s_{2n+1} - x_n + x_n) = (\lim_{n\to\infty} s_{2n+1} - x_n) + \lim_{n\to\infty} x_n = a + 0 = a.$$

Since $(s_{2n})_{n\in\mathbb{N}}$ and $(s_{2n+1})_{n\in\mathbb{N}}$ have the same limit, it follows that also $(s_n)_{n\in\mathbb{N}}$ converges and has the same limit.
The error estimate follows from

$$|s_{2n} - s| = s_{2n} - \underbrace{s}_{\geq s_{2n+1}} \leq s_{2n} - s_{2n+1} = x_{2n+1},$$
$$|s_{2n+1} - s| = \underbrace{s}_{\leq s_{2n+2}} - s_{2n+1} \leq s_{2n+2} - s_{2n+1} = x_{2n+2}.$$

□

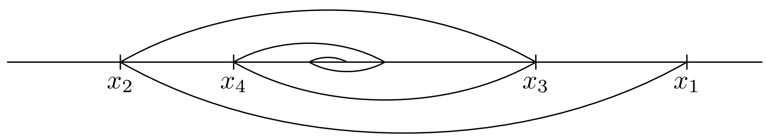


FIGURE 4.1: Leibniz criterion.

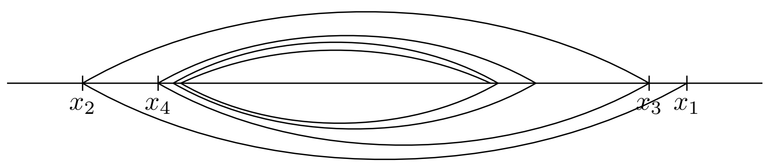


FIGURE 4.2: Leibniz criterion in the case that $\lim x_n \neq 0$.

Remark. If in the preceding theorem $\lim_{n\to\infty} x_n = x$ not necessarily equal to zero, but otherwise all assumptions are satisfied, then the sequence of the partial sums $(s_n)_{n\in\mathbb{N}}$ has exactly two cluster values and

$$\limsup_{n\to\infty} s_n - \liminf_{n\to\infty} s_n = \lim_{n\to\infty} x_n.$$

Examples. $\sum_{k=1}^{\infty} (-)^n \frac{1}{k} = \ln 2, \quad \sum_{k=1}^{\infty} (-)^n \frac{1}{2k-1} = \frac{\pi}{4}.$

Proof of the limits: Example 6.71 and Exercise 7.1.

Definition 4.55. Let $b \in \mathbb{N}$, $b \geq 2$, $\ell \in \mathbb{N}_0$, $(a_k)_{k=-\ell}^{\infty} \subseteq \{0, 1, \dots, b-1\}$. Then the sum

$$\pm \sum_{k=-\ell}^{\infty} a_k b^{-k}$$

is called a *b-adic fraction*[1]. If there exists a $K \in \mathbb{N}$ such that $a_k = 0$, $k \geq K$, then it is called a finite b-adic fraction.

Theorem 4.56. (i) *Each b-adic fraction converges to a real number.*

(ii) *Each real number has a representation as a b-adic fraction. The representation is unique if $a_k \neq b-1$ for almost all $k \geq -\ell$.*

Proof. (i) It suffices to show that $\pm \sum_{k=-\ell}^{\infty} a_k b^{-k}$ is a Cauchy sequence. Let $\varepsilon > 0$. Since $b \geq 2 >$

[1] not to be confused with p-adic fractions from number theory

1, there exists an $N \in \mathbb{N}$ such that $b^{-N} < \varepsilon$. For $n > m \geq N$ it follows that

$$\left|\sum_{k=-\ell}^{n} a_k b^{-k} - \sum_{k=-\ell}^{m} a_k b^{-k}\right| = \sum_{k=m+1}^{n} a_k b^{-k} \leq (b-1) \sum_{k=m+1}^{n} b^{-k}$$
$$= (b-1)b^{-(m+1)} \sum_{k=0}^{n} b^{-k} = (b-1)b^{-(m+1)} \frac{1}{1-\frac{1}{b}} = b^{-m} \leq b^{-N} < \varepsilon.$$

(ii) Let $x \in \mathbb{R}$. Without loss of generality we can assume $x > 0$. By Theorem 3.17 there exists an $N \in \mathbb{Z}$ such that $b^N \leq x < b^{N+1}$. We will construct a sequence $(a_n)_{n=N}^{\infty} \subset \{0, \ldots, b-1\}$ such that for all $n \geq N$:

$$\underbrace{\sum_{k=N}^{n} a_k b^{-k}}_{:=s_n} \leq x < \Big(\sum_{k=N}^{n} a_k b^{-k}\Big) + b^{-(n)} = \Big(\sum_{k=N}^{n} a_k b^{-k}\Big) + bb^{-(n+1)}. \tag{4.10}$$

Let $a_N = \max\{a \in \mathbb{N}_0 : ab^{-N} \leq x\}$; obviously we have $0 \leq a_N \leq b-1$ and $a_N b^{-N}/ \leq x < (a_N+1)b^N$. Let $n \geq N$ and assume that we have already chosen $a_N, \ldots, a_n \in \{0, \ldots, b-1\}$ such that (4.10) holds for n. Let $a_{n+1} = \max\{a \in \mathbb{N}_0 : ab^{-(n+1)} \leq x - s_n\}$. Obviously, $0 \leq a_{n+1} \leq b-1$ and the inequalities (4.10) hold also for $n+1$. Since $|s_n - x| < b^{-n} \to \infty$, $n \to \infty$, it follows that b-adic fraction constructed above converges to x. □

Corollary 4.57 (Cantor). $\mathbb{R}$ *is uncountable.*

Proof. Let $A = \{(a_n)_{n\in\mathbb{N}} : a_n \in \{0,1\}, n \in \mathbb{N}\}$ be the set consisting of all sequences that contain only 0 and 1. Assume that A is countable. Then $A = \{x_n : n \in \mathbb{N}\}$. Each $x_n \in A$ is a sequence $(x_{n,k})_{k\in\mathbb{N}}$. We construct a sequence $y = (y_n)_{n\in\mathbb{N}} \in A$ as follows: Let $y_k = 0$ if $x_{k,k} = 1$ and $y_k = 1$ if $x_{k,k} = 0$. Since $y_k \in \{0,1\}$, $k \in \mathbb{N}$, we hae that $y \in A$. On the other hand, $y \neq x_n$ for all $n \in \mathbb{N}$. Hence the set A is not countable.

Since by Theorem 4.56 the map

$$A \to \mathbb{R}, \quad a = (a_n)_{n\in\mathbb{N}} \mapsto \sum_{n\in\mathbb{N}} a_n 10^{-n}$$

is well-defined and injective, $\mathbb{R}$ contains an uncountable set and therefore it is not countable. □

4.5.2 Series in normed spaces and absolute convergence

Definition 4.58. Let $(V, \|\cdot\|)$ be a normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$. The series $\sum_{n=1}^{\infty} x_n$ is called *absolutely convergent* if the series

$$\sum_{n=1}^{\infty} \|x_n\|$$

converges in $\mathbb{R}$.

Theorem 4.59. *Let* $(V, \|\cdot\|)$ *be a complete normed space and* $(x_n)_{n\in\mathbb{N}} \subseteq V$. *Then*

$$\sum_{n=1}^{\infty} \|x_n\| \ \text{converges} \quad \Longrightarrow \quad \sum_{n=1}^{\infty} x_n \ \text{converges.}$$

Proof. Let $\varepsilon > 0$. Since the series is absolutely convergent there exists an $N \in \mathbb{N}$ such that for all $n \geq m \geq N$

$$\Big\| \sum_{k=m}^{n} x_n \Big\| \leq \sum_{k=m}^{n} \|x_n\| < \varepsilon.$$

Therefore the series $\sum_{k=1}^{\infty} x_n$ converges by the Cauchy criterion (Theorem 4.48). $\square$

Lemma 4.60. *Let $(V, \|\cdot\|)$ be a complete normed space and $(x_n)_{n\in\mathbb{N}} \subseteq V$ such that $\sum_{n=1}^{\infty} \|x_n\|$ converges. Then*

$$\Big\| \sum_{n=1}^{\infty} x_n \Big\| \leq \sum_{n=1}^{\infty} \|x_n\|.$$

Proof. For all $n \in \mathbb{N}$ we have $\Big\| \sum_{k=1}^{n} x_k \Big\| \leq \sum_{k=1}^{n} \|x_k\|$. Taking the limit $n \to \infty$ on both sides proves the assertion. $\square$

Criteria for absolute convergence

Theorem 4.61 (Comparison test). *Let $(V, \|\cdot\|)$ be a complete normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$, $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$, $N_0 \in \mathbb{N}$ such that $a_n \geq \|x_n\|$, $n \geq N_0$. Then*

$$\sum_{n=1}^{\infty} a_n \quad \textit{converges} \quad \Longrightarrow \quad \sum_{n=1}^{\infty} x_n \quad \textit{converges and} \quad \sum_{n=1}^{\infty} \|x_n\| \leq \sum_{n=1}^{\infty} a_n.$$

Proof. Let $\varepsilon > 0$. Since $\sum_{n\in\mathbb{N}} a_n$ converges, there exists an $N \in \mathbb{N}$ such that

$$\sum_{k=m}^{n} \|x_k\| \leq \sum_{k=m}^{n} a_k < \varepsilon, \qquad n \geq m \geq \max\{N_0, N\}.$$

Therefore the series $\sum_{k=1}^{\infty} \|x_n\|$ converges by the Cauchy criterion (Theorem 4.48) which implies that also the series $\sum_{k=1}^{\infty} x_n$ converges (Theorem 4.59). $\square$

Example. The series $\sum_{k=1}^{n} \frac{1}{k^s}$ converges for $s \geq 2$ by the comparison test since $0 < \frac{1}{k^s} \leq \frac{1}{k^2}$, $k \in \mathbb{N}$, and $\sum_{k=1}^{n} \frac{1}{k^2}$ converges by Example 4.52.

Remark. $\zeta(s) = \sum_{k=1}^{n} \frac{1}{k^s}$ defines the so-called *Riemann Zeta function.* .

Theorem 4.62 (Root test). *Let $(V, \|\cdot\|)$ be a complete normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$ and $a = \limsup_{n\to\infty} \sqrt[n]{\|x_n\|}$. Then*

$$a > 1 \quad \Longrightarrow \quad \sum_{n=1}^{\infty} x_n \quad \textit{divergent,}$$

$$a < 1 \quad \Longrightarrow \quad \sum_{n=1}^{\infty} x_n \quad \textit{absolutely convergent.}$$

Proof. Assume that $a > 1$. Then there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $\sqrt[n_k]{x_{n_k}} \geq 1$, $k \in \mathbb{N}$. Hence $(x_{n_k})_{k\in\mathbb{N}}$ does not converge to 0, therefore the series does not converge (Theorem 4.48).

Now let $a < 1$ and fix a q such that $a < q < 1$. Since a is the greatest cluster value of $(\sqrt[n]{\|x_n\|})_{n\in\mathbb{N}}$ there exists a $K \in \mathbb{N}$ such that $q > \sqrt[k]{\|x_k\|})$ for all $k \geq K$. Since $\sum_{k=K}^{\infty} q^k$ is a convergent harmonic series and $q^k > \|x_k\|$, $k \geq K$. it follows by the comparison test that also $\sum_{k=K}^{\infty} \|x_k\|$ converges. $\square$

Similarly, the ratio test is proved.

Theorem 4.63 (Ratio test). *Let $(V, \|\cdot\|)$ be a complete normed space and let $(x_n)_{n\in\mathbb{N}} \subseteq V$. If there exists an $a > 1$ such that $\|x_{n+1}\| \geq a\|x_n\|$ for almost all $n \in \mathbb{N}$, then the series $\sum_{n=1}^{\infty} x_n$ diverges.*
If there exists an $0 < a < 1$ such that $\|x_{n+1}\| \leq a\|x_n\|$ for almost all $n \in \mathbb{N}$, then the series $\sum_{n=1}^{\infty} x_n$ converges absolutely.

For $a = 1$ in Theorem 4.62 or Theorem 4.63 then the root respectively ratio test gives no information about convergence of the series as the examples $\sum_{n=1}^{\infty} \frac{1}{n}$ and $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ show.

Examples 4.64. (i) $\sum_{n=0}^{\infty} \dfrac{z^n}{n!}$ converges absolutely for every $z \in \mathbb{C}$.

Proof. The assertion is clear for $z = 0$. For $z \in \mathbb{C} \setminus \{0\}$ the assertion follows from the ratio test since, for $n > 2|z|$

$$\frac{|z^{n+1}|}{(n+1)!}\Big(\frac{|z^n|}{n!}\Big)^{-1} = \frac{|z|}{n+1} < \frac{|z|}{2|z|+1} < \frac{1}{2} < 1. \qquad \square$$

(ii) The ratio and root tests give no information about convergence of $\sum_{n=0}^{\infty} \dfrac{1}{k^2}$ since $\limsup_{n\to\infty} \frac{1}{k^n} = 1 = \lim_{n\to\infty} \frac{1}{(k+1)^2}\big(\frac{1}{k^n}\big)^{-1}$.

(iii) $\sum_{n=0}^{\infty} 2^{-k+(-)^k} = \dfrac{1}{2} + 1 + \dfrac{1}{8} + \dfrac{1}{4} + \cdots$ converges absolutely.

Proof. Since $\lim_{k\to\infty} 2^k = \lim_{k\to\infty} 2^{-k} = 1$ by Exercise 4.5 the root test yields

$$\limsup_{n\to\infty} \sqrt[k]{2^{-k+(-)^k}} = \limsup_{n\to\infty}\big(2^{-1}\, 2^{\frac{(-)^k}{k}}\big) = 2^{-1} \limsup_{n\to\infty} 2^{\frac{(-)^k}{k}} = \frac{1}{2} < 1. \qquad \square$$

Note that in the last example the ratio test is not applicable. In general, whenever the ratio test shows convergence, so does the root criterion. Indeed, if there is an $0 < a < 1$ such that $\|x_{n+1}\| \leq a\|x_n\|$ for all $n \in \mathbb{N}$, then $\|x_n\| \leq a^n\|x_0\|$ for all $n \in \mathbb{N}$. Therefore $\sqrt[n]{\|x_n\|} \leq a^n \sqrt[n]{\|x_0\|}$ for all $n \in \mathbb{N}$. Since $\sqrt[n]{\|x_0\|} \to 1$ for $n \to \infty$, the root test also shows convergence.

Rearrangement of series

Definition 4.65. Let $(V, \|\cdot\|)$ be a normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$ and $\sigma : \mathbb{N} \to \mathbb{N}$ a permutation. Then $\displaystyle\sum_{n=1}^{\infty} x_{\sigma(n)}$ is a *rearrangement* of $\displaystyle\sum_{n=1}^{\infty} x_n$.

Definition 4.66. Let $(V, \|\cdot\|)$ be a normed space and $(x_n)_{n\in\mathbb{N}} \subseteq V$. The series is called *unconditionally convergent* if for every permutation $\sigma : \mathbb{N} \to \mathbb{N}$ the series $\displaystyle\sum_{n=1}^{\infty} x_{\sigma(n)}$ converges and has the same limit as $\displaystyle\sum_{n=1}^{\infty} x_n$. The series is called *conditionally convergent* if it converges but is not unconditionally convergent.

Theorem 4.67 (Rearrangement theorem). *Let $(V, \|\cdot\|)$ be a normed space, $(x_n)_{n\in\mathbb{N}} \subseteq V$ such that $\sum_{n=1}^{\infty}$ is absolutely convergent. Then every rearrangement converges absolutely and has the same limit.*

Proof. Let $\sigma : \mathbb{N} \to \mathbb{N}$ be a permutation and let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $\sum\limits_{k=n}^{\infty} \|x_k\| < \varepsilon$ for all $n \geq N$. Since σ is a permutation, there exists an $K \in \mathbb{N}$ such that $\sigma(k) \geq N$ for all $k \geq K$. Obviously, $K \geq N$. Define the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ by

$$a_n := \sum_{k=0}^{n} x_k, \quad b_n := \sum_{k=0}^{n} x_{\sigma(k)}, \qquad n \in \mathbb{N}_0.$$

For $n \in \mathbb{N}$, $n \geq K$, is follows that

$$\begin{aligned} \|a_n - b_n\| &= \Big\| \sum_{k=0}^{n} x_k - \sum_{k=0}^{n} x_{\sigma(k)} \Big\| = \Big\| \sum_{k=N+1}^{n} x_k - \sum_{\substack{k=0,\dots,n \\ \sigma(k)>N}} x_{\sigma(k)} \Big\| \\ &\leq \sum_{\substack{k=N+1,\dots,n \\ \sigma(k)>N}} \Big\| x_{\sigma(k)} \Big\| \leq \sum_{k=N}^{\infty} \|x_k\| < \varepsilon. \end{aligned}$$

This shows that the sequence $(b_n)_{n\in\mathbb{N}}$ converges and has the same limit as $(a_n)_{n\in\mathbb{N}}$.
The absolute convergence of the rearranged series follows when the above proof is applied to the series $\sum_{k=0}^{n} \|x_k\|$. □

Theorem 4.68. *Let $(V, \|\cdot\|)$ be a complete normed space, $x_{kl} \in V$, $k, l \in \mathbb{N}_0$, such that*

$$M := \sup\Big\{ \sum_{k=0}^{n}\sum_{l=0}^{n} \|x_{kl}\| : n \in \mathbb{N} \Big\} < \infty. \tag{4.11}$$

Then the series

$$\sum_{k=0}^{\infty}\Big(\sum_{l=0}^{\infty} x_{kl}\Big), \qquad \sum_{l=0}^{\infty}\Big(\sum_{k=0}^{\infty} x_{kl}\Big), \qquad \sum_{n=0}^{\infty}\Big(\sum_{\substack{k,l=0 \\ k+l=n}} x_{kl}\Big)$$

converge absolutely and have the same limits.

Remark. In this case the notation $\sum\limits_{k,l=0}^{\infty} x_{kl}$ is used.

Proof. For each $k \in \mathbb{N}_0$ the series $s_k := \sum_{l=0}^{\infty} x_{kl}$ converges absolutely by theorem 4.51 because $(\|x_{kl}\|)_{l\in\mathbb{N}_0} \subset \mathbb{R}$, $\|x_{kl}\| \geq 0$, and the corresponding sequence of the partial sums $\sum_{l=0}^{m} \|x_{kl}\|$ is bounded by assumption (4.11).
Analogously it follows that for all $l, n \in \mathbb{N}_0$ the series $t_l := \sum_{k=0}^{\infty} x_{kl}$ and $v_n := \sum_{\substack{k,l=0 \\ k+l=n}} x_{kl}$ are absolutely convergent. Therefore, for any $N \in \mathbb{N}$, the partial sums

$$\sum_{k=0}^{N} s_k, \quad \sum_{l=0}^{N} t_l, \quad \text{and} \quad \sum_{n=0}^{N} v_n \tag{4.12}$$

are well-defined. We show that the series $\sum_{k=0}^{\infty} s_k$ is absolutely convergent: For arbitrary $K, L \in \mathbb{N}_0$ we have by the triangle inequality and by assumption (4.11)

$$\sum_{k=0}^{K} \Big\| \sum_{l=0}^{L} x_{kl} \Big\| \leq \sum_{k=0}^{K}\sum_{l=0}^{L} \|x_{kl}\| \leq M < \infty.$$

By Corollary 4.32 this inequality remains true in the limit $L \to \infty$:

$$\sum_{k=0}^{K} \|s_k\| = \sum_{k=0}^{K} \Big\| \sum_{l=0}^{\infty} x_{kl} \Big\| \le M < \infty.$$

The assertion follows again from Theorem 4.51. Analogously it can be shown that the series $\sum_{l=0}^{\infty} t_l$ and $\sum_{n=0}^{\infty} v_n$ are absolutely convergent. Let $S := \sum_{k=0}^{\infty} s_k$ and $V := \sum_{n=0}^{\infty} v_n$. It remains to be shown that $S = V$. For arbitrary $\varepsilon > 0$ there exist $\sigma \in \mathbb{N}$ and $\nu \in \mathbb{N}$ such that $\sum_{k=\sigma}^{\infty} \|s_k\| < \frac{\varepsilon}{4}$, $\sum_{n=\nu}^{\infty} \|v_n\| < \frac{\varepsilon}{4}$ and $\sigma > \nu$. Moreover, there exists $L \in \mathbb{N}$ such that $L > \nu$ and $\sum_{l=L}^{\infty} \|x_{kl}\| < \frac{1}{4\sigma\varepsilon}$, $k = 0, \ldots, \sigma - 1$. Let

$$\begin{aligned} Z :=& \{(k,l) \in \mathbb{N}^2 \,:\, k \le \sigma - 1,\ l \le L-1,\} \setminus \{(k,l) \in \mathbb{N}^2 \,:\, k + l \le \sigma - 1\} \\ &\subset \{(k,l) \in \mathbb{N}^2 \,:\, k + l \ge \sigma\}. \end{aligned}$$

It follows that

$$\begin{aligned} |V - S| \;&=\; \Big| V - \sum_{k=0}^{\sigma-1} s_k - \sum_{k=\sigma}^{\infty} s_k \Big| \;\le\; \sum_{k=\sigma}^{\infty} \|s_k\| + \Big| V - \sum_{k=0}^{\sigma-1} s_k \Big| \\ &\le\; \frac{\varepsilon}{4} + \Big| V - \sum_{k=0}^{\sigma-1} \Big(\sum_{l=0}^{L-1} x_{kl} + \sum_{l=L}^{\infty} x_{kl} \Big) \Big| \\ &\le\; \frac{\varepsilon}{4} + \Big| \sum_{k=0}^{\sigma-1} \sum_{l=L}^{\infty} x_{kl} \Big| + \Big| V - \sum_{k=0}^{\sigma-1} \sum_{l=0}^{L-1} x_{kl} \Big| \\ &\le\; \frac{\varepsilon}{4} + \sum_{k=0}^{\sigma-1} \sum_{l=L}^{\infty} \|x_{kl}\| + \Big| \sum_{n=\nu}^{\infty} v_n \Big| + \Big| \sum_{n=0}^{\nu-1} v_n - \sum_{k=0}^{\sigma-1} \sum_{l=0}^{L-1} x_{kl} \Big| \\ &\le\; \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \Big| \sum_{n=0}^{\nu-1} v_n - \sum_{k=0}^{\sigma-1} \sum_{l=0}^{L-1} x_{kl} \Big| \\ &=\; \frac{3\varepsilon}{4} + \Big| \sum_{(k,l)\in Z} x_{kl} \Big| \le \frac{3\varepsilon}{4} + \sum_{(k,l)\in Z} \|x_{kl}\| \\ &\le\; \frac{3\varepsilon}{4} + \Big| \sum_{k+l \ge \nu} x_{kl} \Big| \;\le\; \frac{3\varepsilon}{4} + \frac{\varepsilon}{4} \;=\; \varepsilon, \end{aligned}$$

Analogously $\sum_{l=0}^{\infty} t_l = \sum_{n=0}^{\infty} v_n$ is shown. □

Theorem 4.69 (Cauchy product). *Let $\mathbb{F}$ be a field $\|\cdot\|$ a norm on $\mathbb{F}$ and $(x_k)_{k\in\mathbb{N}_0}$, $(y_l)_{l\in\mathbb{N}_0} \subset \mathbb{F}$. If the series $\sum_{k=0}^{\infty} x_k$, $\sum_{l=0}^{\infty} y_l$ are absolute convergent, then so is their* Cauchy product

$$\sum_{n=0}^{\infty} z_n := \sum_{n=0}^{\infty} \sum_{k=0}^{n} x_k y_{n-k}$$

and

$$\Big(\sum_{k=0}^{\infty} x_k \Big) \cdot \Big(\sum_{l=0}^{\infty} y_l \Big) = \sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} x_k y_{n-k} \Big).$$

Proof. Let $x_{kl} := x_k \cdot y_l,\ k,l \in \mathbb{N}_0$. Then

$$\begin{aligned}\sum_{k,l=0}^{n} \|x_{kl}\| = \sum_{k,l=0}^{n} \|x_k\| \cdot \|y_l\| &= \Big(\sum_{k=0}^{n} \|x_k\|\Big)\Big(\sum_{l=0}^{n} \|y_l\|\Big) \\ &\leq \underbrace{\Big(\sum_{k=0}^{\infty} \|x_k\|\Big)\Big(\sum_{l=0}^{\infty} \|y_l\|\Big)}_{=:M} < \infty.\end{aligned}$$

and the assertion follows from Theorem 4.68. □

The following theorem shows that the assumption of absolute convergence in Theorem 4.67 is necessary.

Theorem 4.70 (Riemann rearrangement theorem). *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ such that $\sum_{n=1}^{\infty} x_n$ is convergent but not absolutely convergent and let $x \in \mathbb{R}$. Then there exists a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

$$\sum_{n=1}^{\infty} x_{\sigma(n)} = x.$$

Proof. There are infinitely many positive and negative terms sequence $(x_n)_{n\in\mathbb{N}}$ because otherwise the series $\sum_{n=1}^{\infty} x_n$ would be absolutely convergent, in contradiction to the assumption that the series is conditionally convergent. Let $(a_n)_{n\in\mathbb{N}}$ be the sequence of all non-negative terms and $(b_n)_{n\in\mathbb{N}}$ be the sequence of all negative terms. These sequences can be chosen by induction: Let $n_1 = \min\{n \in \mathbb{N} : x_n \geq 0\}$ and set $a_1 := x_{n_1}$. Assume that $n_1 < n_1 < \dots n_k$ are already chosen. Let $n_{k+1} := \min\{n \in \mathbb{N} : n > n_k \wedge x_n \geq 0\}$ and set $a_{k+1} := x_{n_{k+1}}$. Since $\sum_{n=1}^{\infty} x_n$ is conditionally convergent, we have

$$\lim_{n\to\infty} a_n = 0, \qquad \lim_{n\to\infty} b_n = 0, \tag{4.13}$$

$$\sum_{n=1}^{\infty} a_n = \infty, \qquad \sum_{n=1}^{\infty} b_n = -\infty. \tag{4.14}$$

Next we define the permutation $\sigma : \mathbb{N} \to \mathbb{N}$ by induction. Assume that $\sigma(1), \dots, \sigma(k)$ are already defined. As $x_{\sigma(k+1)}$ we chose the next not yet chosen term in the sequence

$$\begin{cases} (a_n)_{n\in\mathbb{N}} & \text{if } \sum_{j=1}^{k} x_{\sigma(j)} \leq x, \\ (b_n)_{n\in\mathbb{N}} & \text{if } \sum_{j=1}^{k} x_{\sigma(j)} > x. \end{cases}$$

Note that in the first case there exists a $n \in \mathbb{N}$, $n \geq k$ such that $\sum_{j=1}^{n} x_{\sigma(j)} > x$ by (4.14); analogously in the second case.
Now we prove that the rearranged sum converges to x. The idea is the same as in the proof of the Leibniz criterion (Theorem 4.54).
Let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $|x_{\sigma(n)}| < \varepsilon$, $n \geq N$. Without restriction assume $\sum_{j=1}^{N} x_{\sigma(j)} \leq x$. Choose $K \in \mathbb{N}$ such that $\sum_{j=1}^{N+K} x_{\sigma(j)} > x$. Then

$$\Big|\sum_{j=1}^{n} x_{\sigma(j)} - x\Big| \leq \varepsilon, \qquad n \geq N + K.$$ □

Remark. Theorem 4.67 shows that every absolutely convergent series is conditionally convergent. The Riemann rearrangement theorem (Theorem 4.70) shows that in $\mathbb{R}$ a series is absolutely convergent if and only if it is unconditionally convergent.

For instance, take the space of bounded sequences $V = \big\{(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{C} : \sup\{|x_n| : n \in \mathbb{N}\} < \infty\big\}$. Then $\|x\|_\infty := \sup\{|x_n| : n \in \mathbb{N}\} < \infty\big\}$ defines a norm on V (see Definition 5.38 and Theorem 5.39). We consider the sequence $(\mathrm{e}_n)_{n\in\mathbb{N}} \subseteq V$ where $\mathrm{e}_n = (\delta_{kn})_{k\in\mathbb{N}} \in V$ is the sequence whose nth term is 1 and all other terms are 0. It is not hard to see that $\sum\limits_{n\in\mathbb{N}} \mathrm{e}_{\sigma(n)} = \sum\limits_{n\in\mathbb{N}} \mathrm{e}_n = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots)$ for every bijection $\sigma : \mathbb{N} \to \mathbb{N}$. However, $\sum_{n\in\mathbb{N}} \| \mathrm{e}_n \| = \sum\limits_{n\in\mathbb{N}} 1 = \infty$.

Euler's number e

Theorem 4.71. *The sequence* $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *defined by*

$$x_n := \Big(1 + \frac{1}{n}\Big)^n, \quad n \in \mathbb{N},$$

converges and

$$\mathrm{e} := \lim_{n\to\infty} \Big(1 + \frac{1}{n}\Big)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \tag{4.15}$$

is called Euler's number (e = 2,71828182...).

Proof. Exercise 4.12. We show the assertion in several steps:

(i) $2 \le \Big(1 + \frac{1}{n}\Big)^n \le \sum\limits_{k=0}^{n} \frac{1}{k!} \le 3, \qquad n \ge 4.$

(ii) The sequences $(x_n)_{n\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ where $s_n := \sum_{k=0}^{\infty} \frac{1}{k!}$ converge.

(iii) Finally we show (4.15).

(i) Let us show the second in equality in (4.15). For all $n \in \mathbb{N}$ we have

$$\Big(1 + \frac{1}{n}\Big)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \underbrace{\frac{n!}{n^k (n-k)!}}_{\le 1} \frac{1}{k!} \le \sum_{k=0}^{n} \frac{1}{k!}\,.$$

To show the last inequality in (4.15) let $n \ge 4$. Then, using $2^k < k!$ for all $k \ge 4$, which can be shown easily by induction, we find

$$\begin{aligned}\sum_{k=0}^{n} \frac{1}{k!} &= \underbrace{\sum_{k=0}^{4} \frac{1}{k!}}_{=16} + \sum_{k=4}^{n} \underbrace{\frac{1}{k!}}_{\le 2^{-k}} \le \frac{16}{6} + \sum_{k=4}^{n} 2^{-k} = \frac{16}{6} + 2^{-4} \sum_{k=0}^{n} 2^{-k} \\ &\le \frac{16}{6} + 2^{-4} \sum_{k=0}^{\infty} 2^{-k} = \frac{16}{6} + \frac{2^{-4}}{1 - \frac{1}{2}} = \frac{16}{6} + \frac{1}{2^3} = \frac{16}{6} + 2^{-3} = \frac{67}{24} < 3.\end{aligned}$$

The first inequality holds since $(x_N)_{n\in\mathbb{N}}$ is monotonically increasing as is shown below and $(1 + \frac{1}{2})^2 = 2.25 > 2$.

(ii) Since the sequence $(s_n)_{n\in\mathbb{N}}$ is monotonically increasing and bounded from above by (i), it converges.

To show that the sequence $(s_n)_{n\in\mathbb{N}}$ converges, it suffices to show that is monotonically increasing because it is bounded from above by (i). Since $a_n \geq 1 > 0$, $n \in \mathbb{N}$, the monotonicity follows from

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \Big(\frac{1+\frac{1}{n+1}}{1+\frac{1}{n}}\Big)^n\Big(1+\frac{1}{n+1}\Big) = \Big(\frac{n(n+1)+n}{n(n+1)+n+1}\Big)^n\Big(1+\frac{1}{n+1}\Big)\\ &= \Big(1-\frac{1}{n^2+2n+1}\Big)^n\Big(1+\frac{1}{n+1}\Big) \leq \Big(1-\frac{n}{(n+1)^2}\Big)\Big(1+\frac{1}{n+1}\Big)\\ &= \frac{n^3+3n^2+3n+2}{n^3+3n^2+3n+1} < 1,\end{aligned}$$

where in the second line we used the Bernoulli inequality (3.2).

(iii) It follows from part (i) that

$$\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} s_n = \sum_{k=0}^{\infty}\frac{1}{k!}.$$

To show the converse inequality, we show that for each $n \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ such that $x_m \geq s_n$. For fixed $n \in \mathbb{N}$ and $m > n$ it follows that

$$\begin{aligned}x_m - s_n &= \Big(1+\frac{1}{m}\Big)^m - \sum_{k=0}^{n}\frac{1}{k!} = \sum_{k=0}^{m}\binom{m}{k}\frac{1}{m^k} - \sum_{k=0}^{n}\frac{1}{k!}\\ &= \sum_{k=0}^{m}\frac{m!}{(m-k)!m^k}\frac{1}{k!} - \sum_{k=0}^{n}\frac{1}{k!}\\ &= \sum_{k=0}^{n}\underbrace{\frac{1}{k!}}_{\leq 1}\underbrace{\Big(\frac{m!}{(m-k)!m^k}-1\Big)}_{\leq 0} + \underbrace{\sum_{k=n+1}^{m}\frac{m!}{(m-k)!k!}\frac{1}{m^k}}_{\geq \frac{1}{m^{n+1}}}\\ &\geq \sum_{k=0}^{n}\Big(\frac{m!}{(m-k)!m^k}-1\Big) + \frac{1}{n^{n+1}}\end{aligned}$$

Using that $\frac{m-j}{m} \leq 1$, $0 \leq j \leq m$, we can estimate

$$1 > \frac{m!}{(m-k)!m^k} = \frac{m}{m}\cdot\dots\cdot\frac{m-k+1}{m^k} \geq \frac{m}{m}\cdot\dots\cdot\frac{m-n+1}{m^n} \geq \frac{(m-n+1)^n}{m^n}.$$

Since $\frac{(m-n+1)^n}{m^n} = \Big(1-\frac{n-1}{m}\Big)^n$ tends to 1 for $m \to \infty$, we can find an $M \in \mathbb{N}$ such that $0 < 1-\Big(1-\frac{n-1}{m}\Big)^n < \frac{1}{n^{n+1}(n+1)}$, $m \geq M$. Hence for all $m \geq M$

$$x_m - s_n \geq \sum_{k=0}^{n}\Big(\Big(1-\frac{n-1}{m}\Big)^n - 1\Big) + \frac{1}{n^{n+1}} \geq -\frac{n+1}{n^{n+1}(n+1)} + \frac{1}{n^{n+1}} > 0. \qquad \square$$

4.6 Cantor's construction of $\mathbb{R}$

There are several methods to introduce the real numbers. For the method using Dedekind cuts see for instance [Rud76, Appendix to Chapter 1]. Cantor's construction of $\mathbb{R}$ uses Cauchy sequences on $\mathbb{Q}$.

For a sequence $(q_n)_{n\in\mathbb{N}}$ we say that $\mathbb{Q}-\lim_{n\to\infty} q_n = 0$ if and only if for every $\epsilon \in \mathbb{Q}$ with $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $|q_n| < \epsilon$ for every $n \geq N$.[2]

Definition 4.72. A relation $\sim$ on a set X is called *equivalence relation* it is reflexive, symmetric and transitive.

Let $\mathcal{C}_\mathbb{Q}$ be the set of all Cauchy sequences in $\mathbb{Q}$. On $\mathcal{C}_\mathbb{Q}$ we define the relation

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \quad\iff\quad \mathbb{Q}-\lim_{n\to\infty} |x_n - y_n| = 0.$$

It is easy to see that $\sim$ is an equivalence relation on $\mathcal{C}_\mathbb{Q}$. We define

$$\mathbb{R} := \mathcal{C}_\mathbb{Q}/\sim\, = \{[(x_n)_{n\in\mathbb{N}}] : (x_n)_{n\in\mathbb{N}} \in \mathcal{C}_\mathbb{Q}\}.$$

Note that each $q \in \mathbb{Q}$ is identified with the equivalence class $[(q)_{n\in\mathbb{N}}] \in \mathbb{R}$.
Together with the operations $+$ and $\cdot$

$$\begin{aligned} [(x_n)_{n\in\mathbb{N}}] + [(y_n)_{n\in\mathbb{N}}] &:= [(x_n + y_n)_{n\in\mathbb{N}}], \\ [(x_n)_{n\in\mathbb{N}}] \cdot [(y_n)_{n\in\mathbb{N}}] &:= [(x_n \cdot y_n)_{n\in\mathbb{N}}], \end{aligned} \qquad [(x_n)_{n\in\mathbb{N}}], [(y_n)_{n\in\mathbb{N}}] \in \mathbb{R},$$

$\mathbb{R}$ is a field. We define an order on $\mathbb{R}$ by

$$[(x_n)_{n\in\mathbb{N}}] > 0 \quad :\iff \quad \exists\, r \in \mathbb{Q}_+\ \exists\, N_0 \in \mathbb{N}\ \ \forall n \geq N_0 \quad x_n \geq r.$$

It can be shown that $+, \cdot, <$ are well-defined (i. e., if $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are Cauchy sequences in $\mathbb{Q}$, then so are $(x_n + y_n)_{n\in\mathbb{N}}$ and $(x_n \cdot y_n)_{n\in\mathbb{N}}$, and the definitions above do not depend on the sequence chosen to represent the equivalence classes) and that $\mathbb{R}$ indeed is an ordered field.

It remains to be shown that $\mathbb{R}$ has the least upper bound property.
First we note that in $\mathbb{Q}$ the Archimedean property (Theorem 3.16) holds. Indeed, let $x, y > 0$ in $\mathbb{Q}$. Then there exist $p, q, r, s \in \mathbb{N}$ such that $x = \frac{p}{q}$, $y = \frac{r}{s}$. Then $2(qr)x = 2pr > r \geq \frac{r}{s} = y$. Since every Cauchy sequence in $\mathbb{Q}$ is bounded in $\mathbb{Q}$, also $\mathbb{R}$ has the Archimedean property.
Using the Archimedean property the following proposition is proved (see Exercise 3.3).

Proposition 4.73. *For all every pair of real numbers $a < b$ there exists an $x \in \mathbb{Q}$ such that $a < x < b$.*

Theorem 4.74. *$\mathbb{R}$ has the least upper bound property.*

Proof. Let $M \subseteq \mathbb{R}$ such that $M \neq \emptyset$ and M is bounded from above. Since M is bounded there exists an upper bound $b \in \mathbb{R}$ of M. Since $M \neq \emptyset$, there exists an element $a \in \mathbb{R}$ that is not an upper bound of M (take for example $a = \alpha - 1$ for an arbitrary element $\alpha \in M$. We construct a sequence of intervals $\mathbb{R}$

$$[a, b] =: [a_0, b_0] \supset [a_1, b_1] \supset [a_2, b_2] \ldots$$

as follows: If $c := \frac{b-a}{2}$ is an upper bound of M, then we set $[a_1, b_1] = [a_1, c]$, otherwise $[a_1, b_1] = [c, b_1]$, and so on. For each $n \in \mathbb{N}$, b_n is an upper bound of M, but a_n is not. Moreover, $b_n - a_n = \frac{b-a}{2^n}$, $n \in \mathbb{N}$. By the proposition above, we can choose in each interval $[a_n, b_n]$ some $c_n \in \mathbb{Q}$ such that $a_n < c_n < n < b_n$. Obviously, $(c_n)_{n\in\mathbb{N}}$ is Cauchy sequence in $\mathbb{Q}$ because

$$|c_m - c_n| \leq 2^{-m}, \qquad n \geq m,$$

[2] Note that we cannot use the definition of limit from Section 4.2 since it is based on a metric and $\mathbb{R}$ is already used in the defintion of a metric. A way around that would be to define a $\mathbb{Q}$-metric on a set X as a function $d_\mathbb{Q} : X \times X \to \mathbb{Q}$ which satisfies the conditions in Defintion 4.1 and then us $d_\mathbb{Q}$ to define convergence. Note that all theorems proved in this chapter which do not involve $\mathbb{R}$ explicitly remain valid.

therefore $(c_n)_{n\in\mathbb{N}}$ represents an element $c = [(c_n)_{n\in\mathbb{N}}]$ in $\mathbb{R}$.
Finally, we show that c is the least upper bound of M in $\mathbb{R}$. Let $\alpha \in M$. Since for all $n \in \mathbb{N}$ we have that $\alpha < b_n$, it follows that $\alpha \leq \lim_{n\to\infty} b_n = c$, hence c is an upper bound of M. Let d be an arbitrary upper bound of M. Then $d > a_n$, $n \in \mathbb{N}$. Therefore we have $d \geq \lim_{n\to\infty} a_n = c$, therefore c is the least upper bound of M. □

Chapter 5

Continuous functions

5.1 Continuity

Definition 5.1. Let (X, d_X) and (Y, d_Y) be metric spaces, $\mathcal{D} \subseteq X$ and $x_0 \in \mathcal{D}$. A function $f : \mathcal{D} \to Y$ is called *continuous in* x_0

$$:\iff \quad \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D} :\ \big(d_X(x, x_0) < \delta \implies d_Y(f(x), f(x_0)) < \varepsilon\big).$$

The function f is called *continuous* (*in* $\mathcal{D}$) if it is continuous in every $x_0 \in \mathcal{D}$. If f is called discontinuous in $x_0 \in \mathcal{D}$ if it is not continuous in $x_0 \in \mathcal{D}$.

In other words: f is continuous in $x_0 \in \mathcal{D}$

$$\iff \quad \forall \varepsilon > 0\ \exists \delta > 0\ f(B_\delta(x_0) \cap \mathcal{D}) \subseteq B_\varepsilon(f(x_0)). \tag{5.1}$$

In the special case $X = Y = \mathbb{R}$ or $\mathbb{C}$ with the usual metric, a function $f : \mathbb{R} \supset \mathcal{D} \to \mathbb{R}$ is continuous in $x_0 \in \mathcal{D}$ if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D} :\ \big(|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon\big).$$

Geometric Interpretation. If f is continuous in x_0, then for every strip S_ε at $f(x_0)$ there exists an interval I_δ at x_0 such that the graph of f over I_δ lies in the strip S_ε.
In other words, when x is changed sufficiently little about x_0, then the function value remains as close as we want to the function value $f(x_0)$.

Examples 5.2.

(i) Let $a \in \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R},\ f(x) = a$. Obviously, f is continuous.

(ii) $\mathrm{id} : \mathbb{R} \to \mathbb{R},\ \mathrm{id}(x) = x$, is continuous in $\mathbb{R}$.

Proof. Let $x_0 \in \mathbb{R}$ and $\varepsilon > 0$. Then for all $x \in \mathbb{R}$ such that $|x - x_0| < \delta := \varepsilon$ it follows that $|\,\mathrm{id}(x) - \mathrm{id}(x_0)\,| = |x - x_0| < \varepsilon$. □

(iii) Analogously, for arbitrary metric spaces (X, d_X), (Y, d_Y), $a \in Y$ the functions $f : X \to Y,\ f(x) = a$, and $\mathrm{id}_X : X \to X,\ \mathrm{id}(x) = x$, are continuous.

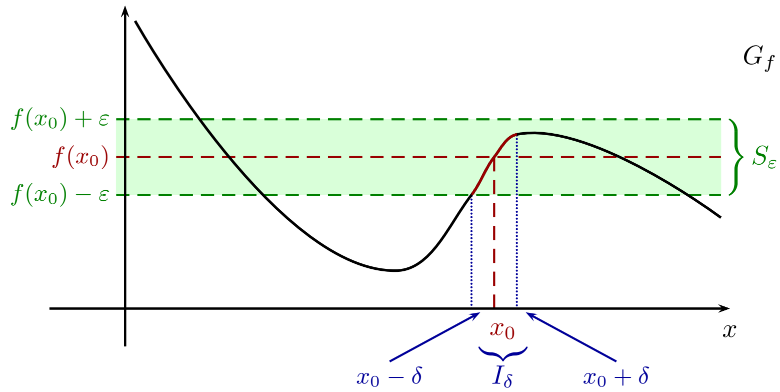


FIGURE 5.1: For every strip S_ε centred at $f(x_0)$ there exists an interval I_δ such that the graph G_f of f above I_δ lies in S_ε.

(iv) Let (X, d_X) be a metric space. For $a \in X$ let Then $f_a : X \to \mathbb{R},\ f(x) = d(a, x)$, is continuous in X.

More generally, let (Y, d_Y) be a metric space and define a metric on $X \times Y$ by

$$d((x_1, y_1), (x_2, y_2)) := \sqrt{d(x_1, x_2)^2 + d(y_1, y_2)^2}, \qquad (x_1, y_1),\ (x_2, y_2), \in X \times Y.$$

For $(a, b) \in X \times Y$ the function $f : X \times Y \to \mathbb{R},\ f(x, y) = d_{X\times Y}((x, y), (a, b))$ is continuous in $X \times Y$.

(v) Let $(X, \|\cdot\|)$ be a normed space. Then $f : X \to \mathbb{R},\ f(x) = \|x\|$, is continuous in X. (This is a special case of (iv) with $a = 0$.)

Proof. Let $x_0 \in X$ and $\varepsilon > 0$. Then for all $x \in X$ we have the implication

$$\|x - x_0\| < \delta \qquad \Longrightarrow \qquad |f(x) - f(x_0)| = |\,\|x\| - \|x_0\|\,| \le \|x - x_0\| < \varepsilon.$$ □

(vi) $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$, is continuous in $\mathbb{R}$.

Proof. Let $x_0 \in \mathbb{R}$ and $\varepsilon > 0$. Let $\delta := \min\{1, \frac{\varepsilon}{1+2x_0}\}$. Then for $x \in \mathbb{R}$ with $\|x - x_0\| < \delta$ it follows that

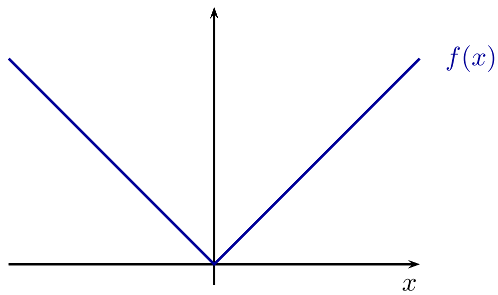
$$|f(x) - f(x_0)| = |x^2 - x_0|^2 = |x - x_0|\,|x + x_0| \le \underbrace{|x - x_0|}_{<\,\delta\,\le\,\frac{\varepsilon}{1+2|x_0|}} \big(\underbrace{|x - x_0|}_{<\delta\le 1} + 2|x_0|\big) < \varepsilon.$$ □

(vii) The Heaviside function

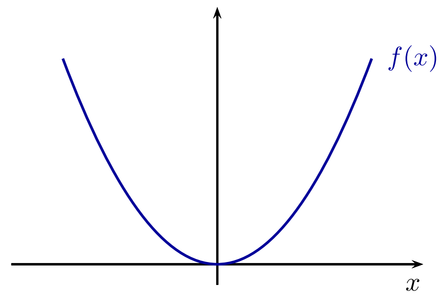
$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} 1, & x \ge 0, \\ 0, & x < 0, \end{cases}$$

is not continuous in $x = 0$.

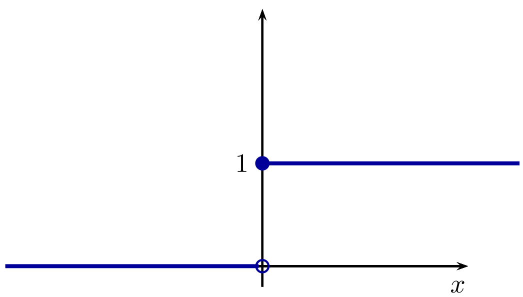
Proof. Assume that there exists a $\delta > 0$ such that $|f(x) - f(0)| < \frac{1}{2}$ for all $x \in \mathbb{R}$ with $|x - 0| < \frac{1}{2}$. This contradicts $|f(-\frac{1}{2}) - f(0)| = 1 > \frac{1}{2}$. □



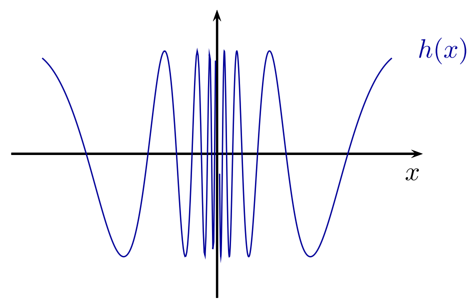
Example 5.2 (v): $f(x) = |x|$



Example 5.2 (vi): $f(x) = x^2$



Example 5.2 (vii): Heaviside function



$h(x) = \sin(x^{-1})$ (Exercise 6.2)

FIGURE 5.2: The functions in the first row are continuous, the functions in the second row are not.

(viii) The Dirichlet function

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

is nowhere continuous in $\mathbb{R}$.

Proof. Exercise 5.4. □

Definition 5.3. Let (X, d_X), (Y, d_Y) be metric spaces. A function $f : X \supseteq \mathcal{D} \to Y$ is called *Lipschitz continuous with Lipschitz constant* L if

$$x, y \in \mathcal{D} \quad \Longrightarrow \quad d_Y(f(x), f(y)) \leq L d_X(x, y).$$

Lipschitz continuity is stronger than continuity.

Theorem 5.4. *Every Lipschitz continuous functions is continuous.*

Proof. Let (X, d_X), (Y, d_Y) be metric spaces and $f : X \supseteq \mathcal{D} \to Y$ Lipschitz continuous with Lipschitz constant $L > 0$. Let $x_0 \in \mathcal{D}$ and $\varepsilon > 0$. Then for all $x \in \mathcal{D}$ with $d_X(x, x_0) < \frac{\varepsilon}{L}$ it follows that

$$d_Y(f(x), f(x_0)) \leq L d_X(x, y) < \varepsilon. \qquad \square$$

The next theorem gives a criterion for continuity of a function in a point in terms of sequences.

Theorem 5.5. *Let (X, d_X), (Y, d_Y) be metric spaces, $\mathcal{D} \subseteq X$, $f : X \supseteq \mathcal{D} \to Y$, $x_0 \in \mathcal{D}$. Then f is continuous in x_0 if and only if for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}$ which converges to x_0 the sequence $(f(x_n))_{n\in\mathbb{N}} \subseteq Y$ converges to $f(x_0)$.*

Proof. "$\Longrightarrow$" Let f be continuous in $x_0 \in \mathcal{D}$ and $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}$ with $x_n \to x_0$, $n \to \infty$. Let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $f(B_\delta(x_0) \cap \mathcal{D}) \subseteq B_\varepsilon(f(x_0))$. Since $x_n \to x_0$, $n \to \infty$, there exists an $N \in \mathbb{N}$ such that $x_n \in B_\delta(x_0)$, $n \geq N$. Therefore $d_Y(f(x_n), f(x_0)) < \varepsilon$, $n \geq N$ which implies that $f(x_n) \to f(x_0)$, $n \to \infty$.
"$\Longleftarrow$" Assume that f is not continuous in x_0. Then there exists an $\varepsilon > 0$ such that

$$\forall \delta > 0 \ \exists x \in \mathcal{D} : \ d_X(x, x_0) < \delta \ \wedge \ d_Y(f(x), f(x_0)) \geq \varepsilon.$$

In particular, we find a sequence $(x_n)_{n\in\mathbb{N}}$ such that

$$\forall n \in \mathbb{N} \ \exists x_n \in \mathcal{D} : \ d_X(x_n, x_0) < \frac{1}{n} \ \wedge \ d_Y(f(x_n), f(x_0)) \geq \varepsilon.$$

Hence $x_n \to x_0$ but $f(x_n) \not\to f(x_0)$, in contradiction to the assumption. $\square$

The previous theorem states that continuous functions and limits commute:

$$\lim_{x\to x_0} f(x) = f(\lim_{x\to x_0} x).$$

Next we show that continuity is compatible with algebraic operations.

Definition 5.6. Let X be a set, Y a vector space over a field $\mathbb{F}$, $\mathcal{D}_f, \mathcal{D}_g \subseteq X$, and $f : \mathcal{D}_f \to Y$, $g : \mathcal{D}_g \to Y$, $\lambda \in \mathbb{F}$. Let $\mathcal{D}_{\lambda f+g} := \mathcal{D}_f \cap \mathcal{D}_g$. Using the algebraic structure on Y we define and sum of two functions and the product with a scalar by

$$\lambda f + g : \mathcal{D}_{\lambda f+g} \to Y, \qquad (\lambda f + g)(x) := \lambda f(x) + g(x).$$

If Y is a field we set $\mathcal{D}_{fg} := \mathcal{D}_f \cap \mathcal{D}_g$, $\mathcal{D}_{\frac{f}{g}} := \{x \in X \, : \, x \in \mathcal{D}_f \cap \mathcal{D}_g, g(x) \neq 0\}$ and

$$fg : \mathcal{D}_{fg} \to Y, \qquad (fg)(x) := f(x)g(x),$$
$$\frac{f}{g} : \mathcal{D}_{\frac{f}{g}} \to Y, \qquad \frac{f}{g}(x) := \frac{f(x)}{g(x)}.$$

Theorem 5.7. *Let (X, d_X) be a metric space, $(Y, \|\cdot\|)$ a normed space over a field $\mathbb{F}$ and $f : X \supseteq \mathcal{D}_f \to Y$, $g : X \supseteq \mathcal{D}_g \to Y$ functions and $\lambda \in \mathbb{F}$. Let $x_0 \in \mathcal{D}_f \cap \mathcal{D}_g$ such that f and g are continuous in x_0. Then*

(i) *$\lambda f + g$ is continuous in x_0.*

If Y is a field, then

(ii) *fg is continuous in x_0.*

(iii) *$\frac{f}{g}$ is continuous in x_0 if $g(x_0) \neq 0$.*

Proof. Exercise 5.1. $\square$

Corollary 5.8. *Let (X, d_X) be a metric space and $(Y, \|\cdot\|)$ a normed space. Then the set*

$$C(X,Y) := \{f : X \to Y \text{ continuous }\}$$

is a linear subspace of the vector space $Y^X = \{f : X \to Y\}$.
If $Y = \mathbb{R}$ or $Y = \mathbb{C}$, then often the notation $C(X)$ is used instead of $C(X,\mathbb{R})$ or $C(X,\mathbb{C})$.

Example. Since the functions $f : \mathbb{R} \to \mathbb{R},\ f(x) = 1$, and $\mathrm{id} : \mathbb{R} \to \mathbb{R}$ are continuous, Theorem 5.7 implies that all polynomials

$$P : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \sum_{n=0}^{m} a_n x^n$$

are continuous in $\mathbb{R}$.

Example. Let $\mathbb{F}$ be a field, $m, n \in \mathbb{N}$. For the *multi-index* $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ we define $|\alpha| = \sum_{k=1}^{n} \alpha_k$ and $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for $x = (x_1, \dots, x_n) \in \mathbb{F}^n$.
A *polynomial of degree m with coefficients in* $\mathbb{F}$ is a function

$$P : \mathbb{F}^n \to \mathbb{F}, \qquad P(x) = \sum_{|\alpha| \le m} c_\alpha x^\alpha,$$

such that there exists at least one $\alpha \in \mathbb{N}^n$ such that $|\alpha| = m$ and $c_\alpha \neq 0$.
A function $R : \mathcal{D}_R \subseteq \mathbb{F}^n \to \mathbb{F}$ is called a *rational function* if there exist polynomials $P, Q : \mathbb{F}^n \to \mathbb{F}$ such that

$$R = \frac{P}{Q}, \qquad \mathcal{D}_R = \{x \in \mathbb{F}^n : Q(x) \neq 0\}.$$

If $\mathbb{F}$ is equipped with a norm, in particular, when $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, then all polynomials and all rational functions on $\mathbb{F}^n$ are continuous by Theorem 5.7 and the fact that the maps

$$\mathbb{F} \times \mathbb{F} \to \mathbb{F}, \quad (x,y) \mapsto x + y, \qquad \mathbb{F} \times \mathbb{F} \to \mathbb{F}, \quad (x,y) \mapsto xy$$

are continuous with the norm on $\mathbb{F} \times \mathbb{F}$ defined by $\|(x,y)\| = \sqrt{\|x\|^2 + \|y\|^2}$.

Theorem 5.9. *Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces and $f : X \supseteq \mathcal{D}_f \to Y$, $g : Y \supseteq \mathcal{D}_g \to Z$ functions such that $\mathrm{R}(f) \subseteq \mathcal{D}_g$. Let $x_0 \in \mathcal{D}_f$. If f is continuous in x_0 and g is continuous in $f(x_0)$ then $g \circ f$ is continuous in x_0.*

Proof. We will use the criterion of Theorem 5.5 to prove the continuity of $g \circ f$ in x_0. Let $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}_f$ such that $x_n \to x_0$ for $n \to \infty$. Then $f(x_n) \to f(x_0)$ because f is continuous in x_0. Since g is continuous in $f(x_0)$ it follows that

$$\lim_{n\to\infty} (g \circ f)(x_n) = \lim_{n\to\infty} g(f(x_n)) = g(\lim_{n\to\infty} f(x_n)) = g(f(x_0)) = (g \circ f)(x_0).$$

Therefore, by Theorem 5.5, $g \circ f$ is continuous in x_0. □

Remark. The continuity of $g \circ f$ does imply neither the continuity of f nor of g.

- f Heaviside function, $g : \mathbb{R} \to \mathbb{R}, g \equiv 0$. Then $g \circ f$ and g are continuous, but f is not continuous in 0.
- $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$, g Heaviside function. Then $g \circ f$ and f are continuous, but g is not continuous in 0.
  If, however, $\mathrm{R}(f) = \mathcal{D}_g$, then continuity of $g \circ f$ and f implies that g is continuous.

- If we choose $f = g =$ Heaviside function, or $f = g =$ Dirichlet function, then neither f nor g is continuous but their composition is.

Definition 5.10. Let (X, d_X) be a metric space and $M \subseteq X$. A point $x_0 \in X$ is called a *limit point* (or *cluster point*) of M if there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq M \setminus \{x_0\}$ such that $\lim_{n\to\infty} x_n = x_0$. In other words:

$$\forall \varepsilon > 0 \ \exists x_\varepsilon \in M : \ 0 < d_X(x, x_\varepsilon) < \varepsilon.$$

Remark.
- A limit point of M does not necessarily belong to M, for example, 1 is a cluster point of the interval $(0,1) \subseteq \mathbb{R}$ but $0 \notin (0,1)$.
- Not every point $x \in M$ is a limit point of M. For example, if $|M| < \infty$ then M contains non limit point.
- Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space X and $M := \{x_n \ n \in \mathbb{N}\}$. Then each limit point of M is a cluster value of x_n, but the converse is not true. For example, consider the sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ defined by $x_n = 1$, $n \in \mathbb{N}$. Then 1 is a cluster value of the sequence, but it is not a limit point of the corresponding set M.

Definition 5.11. Let (X, d_X), (Y, d_Y) metric spaces, $f : X \supseteq \mathcal{D}_f \to Y$ a function and x_0 a limit point of $\mathcal{D}_f$. A point $a \in Y$ is called *limit of f in x_0* if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in \mathcal{D}_f : \ \Big(0 < d_X(x, x_0) < \delta \implies d_Y(f(x), a) < \varepsilon\Big).$$

Theorem 5.12 shows that the limit is uniquely determined and we write

$$\lim_{x\to x_0} f(x) = a.$$

Remark. The existence of the limit of f in x_0 does not imply that f is defined in x_0.

The next theorem gives a criterion for the existence of the limit of a function in terms of sequences.

Theorem 5.12. *Let (X, d_X), (Y, d_Y) metric spaces, $f : X \supseteq \mathcal{D} \to Y$ a function and x_0 a limit point of $\mathcal{D}$. Then $\lim_{x\to x_0} f(x) = a$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \setminus \{x_0\}$ which converges to x_0 the sequence $(f(x_n))_{n\in\mathbb{N}} \subseteq Y$ converges to a.*

Proof. "$\Longrightarrow$" Assume that $\lim_{x\to x_0} f(x) = a$ exists. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \setminus \{x_0\}$ such that $x_n \to x_0$ for $n \to \infty$ and let $\varepsilon > 0$. By assumption there exists a $\delta > 0$ such that

$$x \in \mathcal{D} \ \wedge \ 0 < d_X(x, x_0) < \delta \quad \Longrightarrow \quad d_Y(f(x), a) < \varepsilon.$$

Since $(x_n)_{n\in\mathbb{N}}$ converges to x_0, there exists an $N \in \mathbb{N}$ such that $0 < d_X(x_n, x_0) < \delta$, $n \geq \mathbb{N}$, hence $d_Y(f(x_n), a) < \varepsilon$, $n \geq N$. Therefore $f(x_n)$ converges to a.

"$\Longleftarrow$" Assume $\lim_{n\to\infty} f(x_n) = a$ for each sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \setminus \{x_0\}$ with $\lim_{n\to\infty} x_n = x_0$. Assume $\lim_{x\to x_0} f(x) \neq a$. Then there exists an $\varepsilon > 0$ such that for every $n \in \mathbb{N}$ there exists an $x_n \in \mathcal{D}$ such that

$$0 < d_X(x_n, x_0) < \frac{1}{n} \ \text{ and } \ d_Y(f(x_n), a) \geq \varepsilon.$$

Since by construction the sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \setminus \{x_0\}$ converges to x_0, this is a contradiction to the assumption. □

Theorem 5.13. *Let (X, d_X), (Y, d_Y) metric spaces, $f : X \supseteq \mathcal{D} \to Y$ a function and x_0 a limit point of $\mathcal{D}$. Then f is continuous in x_0 if and only if the limit of f in x_0 exists and $\lim_{x\to x_0} f(x) = f(x_0)$.*

Proof. This follows immediately from Theorem 5.5 and Theorem 5.12. □

Note that the existence of the limit in x_0 is not sufficient for continuity of f in x_0. For example, for the function

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \begin{cases} 0, & x \neq 0, \\ 1, & x = 0 \end{cases}$$

the limit of f in x_0 exists and is equal to 0 but f is not continuous in x_0.

Theorem 5.14. *Let (X, d_X), (Y, d_Y) metric spaces, $f : X \supseteq \mathcal{D} \to Y$ a function and $x_0 \notin \mathcal{D}$ a limit point of $\mathcal{D}$. If the limit of f in x_0 exists, then f has a unique continuous extension $\widehat{f} : \mathcal{D} \cup \{x_0\} \to Y$.*

Proof. By Theorem 5.5 the function

$$\widehat{f} : \mathcal{D} \cup \{x_0\} \to Y, \qquad \widehat{f}(x) = \begin{cases} f(x), & x \in \mathcal{D} \\ \lim_{x\to x_0} f(x), & x = x_0 \end{cases}$$

is a continuous extension of f. For any continuous extension $\widetilde{f}$ of f to $\mathcal{D} \cup \{x_0\}$ it follows that $\widehat{f}(x) = f(x) = \widetilde{f}(x)$ for all $x \in \mathcal{D}$ and by continuity of $\widehat{f}$ and $\widetilde{f}$

$$\widehat{f}(x_0) = \lim_{x\to x_0} \widehat{f}(x) = \lim_{x\to x_0} \widetilde{f}(x) = \widetilde{f}(x_0),$$

therefore the continuous extension of f is unique. □

Theorem 5.15 (Cauchy criterion). *Let (X, d_X), (Y, d_Y) be metric spaces, Y a complete metric space, $f : X \supseteq \mathcal{D} \to Y$ a function and x_0 a limit point of $\mathcal{D}$. Then f has a limit in x_0 if and only if*

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x, y \in D_f : \\ \Big(0 < d_X(x, x_0) < \delta\ \wedge\ 0 < d_X(y, x_0) < \delta \implies d_Y\big(f(x), f(y)\big) < \varepsilon\Big). \tag{5.2}$$

Proof. Exercise 5.2.

□

If X is $\mathbb{R}$ or any other ordered field with a norm, also *one-sided limits* are defined.

Definition 5.16. Let (Y, d_Y) be a metric space, $(a, b) \subseteq \mathbb{R}$ an interval and $f : (a, b) \to Y$ a function. For $a \leq x_0 < b$ we define

$$f(x_0+) := \lim_{x\searrow x_0} f(x) = y \in Y$$

if $\lim_{n\to\infty} f(x_n) = y$ for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq (x_0, b)$ such that $x_n \to x_0$ for $n \to \infty$. Analogously

$$f(x_0-) := \lim_{x\nearrow x_0} f(x) = y \in Y$$

for $a < x_0 \leq b$ if $\lim_{n\to\infty} f(x_n) = y$ for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq (a, x_0)$ such that $x_n \to x_0$ for $n \to \infty$.
The function f is called

$$\begin{cases} \textit{right continuous at } x_0 \text{ if } f(x_0+) = f(x_0), \\ \textit{left continuous at } x_0 \text{ if } f(x_0-) = f(x_0). \end{cases}$$

Proposition 5.17. *Let (Y, d_Y) be a metric space, $(a, b) \subseteq \mathbb{R}$ and $x_0 \in (a, b)$. For a function $f : (a, b) \to Y$ the following is equivalent:*

(i) *f is continuous in x_0.*

(ii) *f is left and right continuous in x_0, that is, $f(x_0+) = f(x_0-) = f(x_0)$.*

Example. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) := [x] := \max\{k \in \mathbb{Z} : k \le x\}$. The function f is continuous in $\mathbb{R} \setminus \mathbb{Z}$, it is right continuous in $\mathbb{Z}$ but not left continuous in $\mathbb{Z}$.

Proof. Let $x_0 \in \mathbb{Z}$. Then

$$\lim_{x \searrow x_0} f(x) = x_0 = f(x_0), \qquad \lim_{x \nearrow x_0} f(x) = x_0 - 1 = f(x_0) - 1 \ne f(x_0). \qquad \square$$

Definition 5.18. Let $\mathcal{D} \subseteq \mathbb{R}$. A function $f : \mathcal{D} \to \mathbb{R}$ is called

(i) *monotonically increasing* $\iff$ $f(x) \ge f(y)$ if $x \ge y$,

(ii) *strictly monotonically increasing* $\iff$ $f(x) > f(y)$ if $x > y$,

(iii) (*strictly*) *monotonically decreasing*, if $-f$ is (strictly) monotonically increasing.

(iv) (*strictly*) *monotonic*, if it is either (strictly) monotonically increasing or (strictly) monotonically decreasing.

Examples. The functions $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$, $g : \mathbb{R} \to \mathbb{R}$, $g(x) = [x]$, $h : \mathbb{R}_+ \to \mathbb{R}$, $h(x) = \sqrt{x}$, are monotonically increasing.

Definition 5.19. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \supseteq \mathcal{D} \to Y$ is called *bounded* if $\mathrm{R}(f)$ is bounded in Y.

Theorem 5.20. *Let $(a, b) \subseteq \mathbb{R}$ and $f : (a, b) \to \mathbb{R}$ monotonic. Then f has one-sided limits in every $x_0 \in \mathcal{D}$.*

Proof. Without restriction we assume that f is monotonically increasing. Let $x_0 > a$ and let $s := \sup\{f(x) : a < x < x_0\}$. We will show that $\lim_{x \nearrow x_0} f(x) = s$. To this end, let $\varepsilon > 0$. Since s is the supremum of $\{f(x) : a < x < x_0\}$ there exists an $x_\varepsilon \in (a, x_0)$ such that $s - \varepsilon < f(x_\varepsilon) \le s$. Since f is monotonically increasing it follows that $s - \varepsilon < f(x) \le s$ for all $x \in (x_\varepsilon, x_0)$. This shows that f the left limit in x_0 exists. Analogously it is shown that the right limit in x_0 exists for $a \le x_0 < b$. $\square$

Theorem 5.21. *If $f : (a, b) \to \mathbb{R}$ is monotonic then it has at most countably many discontinuities.*

Proof. Without restriction we assume that f is monotonically increasing. Let $x_0 \in (a, b)$ be a discontinuity of f. Since f is monotonic, we have that the one-sided limits of f in x_0 exist and, again by the monotonicity of f, that $f(x_0-) < f(x_0+)$. By Proposition 3.19 there exists an $q_0 \in \mathbb{Q}$ such that $f(x_0-) < q_0 < f(x_0+)$. Since $\mathbb{Q}$ is countable and for each $q \in \mathbb{Q}$ there is at most one $x \in (a, b)$ such that $f(x-) < q < f(x+)$, f can have at most countably many discontinuities. $\square$

Definition 5.22. Let (Y, d_Y) be a metric space and $\mathcal{D} \subseteq \mathbb{R}$ an unbounded set. If $\mathcal{D}$ is unbounded from above, then a function $f : \mathcal{D} \to Y$ as the *limit* $a := \lim_{x \to \infty} f(x)$ *at* ∞ if

$$\forall \varepsilon > 0 \; \exists R \in \mathbb{R} : \; \forall x \in \mathcal{D} \; \Big(x \ge R \implies d_Y(f(x), a) < \varepsilon \Big).$$

If $\mathcal{D}$ is unbounded from below, then the *limit of f at* $-\infty$ is defined analogously.

Of course, f does not need to have limits a $\pm\infty$, as for example the function $f : \mathbb{R} \to \mathbb{R},\ f(x) = (x - [x])(1 - (x - [x]))$ shows.

Definition 5.23. Let (X, d_X) be a metric space, $\mathcal{D} \subseteq X$ and x_0 a limit point of $\mathcal{D}$. Then $f : \mathcal{D} \to \mathbb{R}$ *has the limit* ∞ *at* x_0, denoted by $\lim_{x\to x_0} f(x) = \infty$, if

$$\forall R \in \mathbb{R}\ \exists \delta > 0 :\ \forall x \in \mathcal{D}\ \Big(d_X(x, x_0) < \delta \implies f(x) > R\Big).$$

$-\infty$ is the limit of f at x_0, denoted by $\lim_{x\to x_0} f(x) = -\infty$, if ∞ is the limit of $-f$ at x_0.

5.2 Properties of continuous functions

In this section some important properties are discussed, for example the intermediate value theorem which basically says that real intervals have no holes.

Theorem 5.24 (Intermediate value theorem). *Let $a < b \in \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ a continuous function. Without restriction we assume $f(a) \leq f(b)$. Then for each $\gamma \in [f(a), f(b)]$ there exists an $c \in [a, b]$ such that $f(c) = \gamma$.*

In other words: The image of an interval under a continuous real function is convex in $\mathbb{R}$, that is, it is an interval.

Proof. Let $c := \sup\{x \in [a, b] : f(x) \leq \gamma\}$. In particular, $f(x) > \gamma$ for all $x \in (c, b]$. We will show that $f(c) = \gamma$. Assume that $f(c) < \gamma$. Then $\varepsilon := \gamma - f(c) > 0$. Since f is continuous in c, there exists a $\delta > 0$ such that $|f(x) - f(c)| < \varepsilon$ for all $x \in [a, b]$ such that $|x - c| < \delta$. Without restriction we can choose δ small enough that $c + \delta/2 \in [a, b]$. Then $f(c + \frac{\delta}{2}) \leq f(c) + |f(x) - f(c)| < f(c) + \varepsilon < \gamma$. This contradicts the definition of c.
Analogously, if $f(c) > \gamma$, then $\varepsilon := f(c) - \gamma > 0$. Since f is continuous in c, there exists a $\delta > 0$ such that $|f(x) - f(c)| < \varepsilon$ for all $x \in [a, b]$ such that $|x - c| < \delta$. Then $f(x) \geq f(c) - |f(x) - f(c)| \geq f(c) - \varepsilon > \gamma$ for all $x \in (c - \delta, c]$ which also contradicts the definition of c. □

The intermediate value theorem implies that the image of a continuous function defined on an interval is again an interval (see also Theorem 8.41).

Theorem 5.25. *Every polynomial in $\mathbb{R}$ with odd degree has at least one zero.*

Proof. Let $P(x) = \sum_{m=0}^{n} a_m x^m$ such that $a_n \neq 0$. Then x_0 is a zero of P if and only if x_0 is a zero of the polynomial $f = \frac{1}{a_n} P$. For $x \neq 0$ we have that

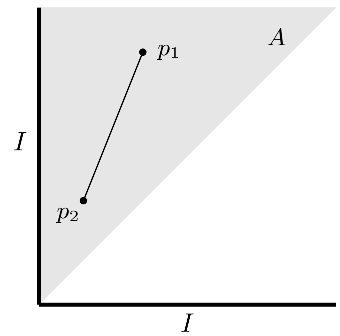
$$f(x) = x^n g(x) \quad \text{with} \quad g(x) = 1 + \frac{a_{n-1}}{a_n} x^{-1} + \cdots + \frac{a_0}{a_n} x^{-n}.$$

Since $g(x) \to 1$ for $x \to \infty$ and $x \to -\infty$ and $\lim_{x\to\infty} x^n = \infty$, $\lim_{x\to-\infty} x^n = -\infty$, there exist $x_\pm \in \mathbb{R}$ such that $f(x_-) < 0$ and $f(x_+) > 0$.
Since the polynomial f is continuous, Theorem 5.24 implies that there exists an $x_0 \in (x_-, x_+)$ such that $f(x_0) = 0$. □

Theorem 5.26. *Let $I \subseteq \mathbb{R}$ a interval and $f : I \to \mathbb{R}$ continuous. Then*

$$f \text{ injective} \iff f \text{ strictly monotonic.}$$

Proof. Exercise 5.5 "$\Longleftarrow$" clear.



"$\Longrightarrow$" Let $f : I \to \mathbb{R}$ continuous and injective. We define the set

$$A := \{(x, y) \in I \times I : x < y\}.$$

Note that A is convex, that is

$$p_1, p_2 \in A \quad \Longrightarrow \quad \{p_1 + t(p_2 - p_1) : 0 \le t \le 1\} \subseteq A.$$

By assumption, the function

$$\varphi : A \to \mathbb{R}, \quad \varphi(x, y) = f(x) - f(y)$$

has no zeros. Assume that f is not monotonic. Then there exist $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ in A such that $\varphi(p_1) < 0$ and $\varphi(p_2) > 0$. Since the function

$$\psi : [0, 1] \to \mathbb{R}, \quad \psi(t) = \varphi(p_1 + t(p_2 - p_1))$$

is continuous and $\psi(0) < 0$ and $\psi(1) > 0$, then intermediate value theorem (Theorem 5.24) implies that there exists an $t_0 \in (0, 1)$ such that

$$0 = \psi(t_0) = \varphi(p_1 + t_0(p_1 - p_2)),$$

in contradiction to the assumption that $\varphi \neq 0$ on A. □

Theorem 5.27. *Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ a strictly monotonic function. Then f is invertible in the sence that there exists a (unique) function $f^{-1} : R(f) \to I$ such that $f \circ f^{-1} = \mathrm{id}_{R_f}$ and $f^{-1} \circ f = \mathrm{id}_I$. The function f^{-1} is strictly monotonic and continuous.*

Proof. Existence, uniqueness and monotonicity of f^{-1} are clear. It remains to be shown that f^{-1} is continuous. To this end, let $p \in I$ such that p is not boundary point and let $\varepsilon > 0$. Since I is an interval, we can assume without restriction that ε is so small that $(p - \varepsilon,\, p + \varepsilon) \subseteq I$. Monotonicity of f implies that there exists a $\delta > 0$ such that

$$f(p - \varepsilon) < f(p) - \delta < f(p) < f(p) + \delta < f(p + \varepsilon).$$

By monotonicity of f^{-1} we obtain for all $y \in \mathcal{D}(f^{-1}) = R(f)$:

$$|y - f(p)| < \delta \quad \Longrightarrow \quad p - \varepsilon < f^{-1}(y) < p + \varepsilon.$$

The proof for p being a boundary point of I is analogous. □

Definition 5.28. Let (X, d_X) be a metric space. A subset $K \subseteq X$ is called *compact* if and only if every sequence $(x_n)_{n\in\mathbb{N}} \subseteq K$ contains a subsequence which converges in K.

Proposition 5.29. *An interval $I \subseteq \mathbb{R}$ is compact if and only if there exist $a \le b \in \mathbb{R}$ such that $I = [a, b]$.*

Proof. "$\Longleftarrow$" Let $a \le b \in \mathbb{R}$ and $I = [a, b]$. By the theorem of Bolzano-Weierstraß every sequence $(x_n)_{n\in\mathbb{N}}$ in I contains a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$. Since $a \le x_{n_k} \le b$ for all $k \in \mathbb{N}$, it follows that $a \le \lim_{k\to\infty} x_{n_k} \le b$.
"$\Longrightarrow$" Assume for example that $I = (a, b]$. Then the sequence $(a + \frac{b-a}{2n})_{n\in\mathbb{N}}$ does not converge in I. □

Theorem 5.30. *Let (X, d_X) be a metric space, $K \subseteq X$ compact and $f : K \to \mathbb{R}$ a continuous function. Then f attains its infimum and supremum, that is,*

$$\exists\, p, q \in K : \quad \forall x \in K \quad f(p) \le f(x) \le f(q).$$

Proof. We show that f attains its supremum. Let $s := \sup\{f(x) : x \in K\}$. Then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $f(x_n) \to s$. Since K is compact, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ and a $q \in K$ such that $x_{n_k} \to q$ for $k \to \infty$. It follows that

$$f(q) = f(\lim_{k\to\infty} x_{n_k}) = \lim_{k\to\infty} f(x_{n_k}) = s,$$

in particular $s < \infty$.
Applying the above to the function $-f$ it follows that f attains its infimum. □

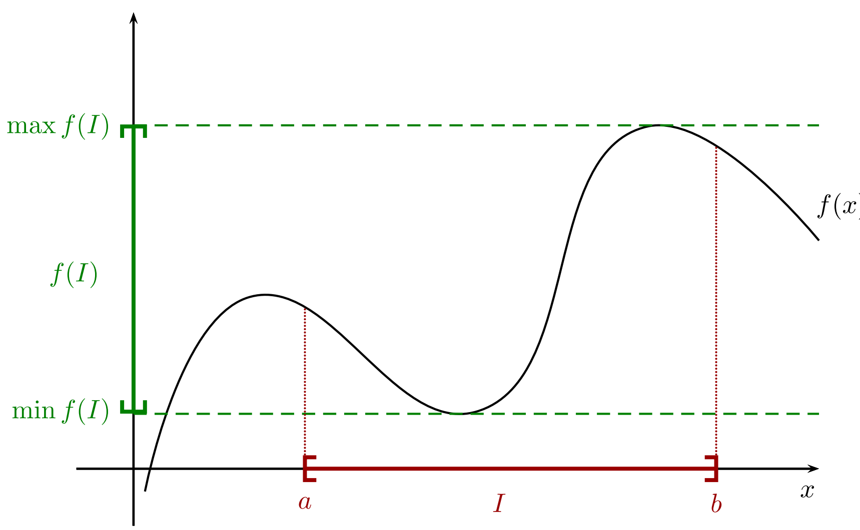


FIGURE 5.3: Intermediate value theorem (Theorem 5.24) and theorem of the minumum and maximum (Theorem 5.30): The continuous function f attains a minimum and a maximum on the closed interval $I = [a, b]$. The image of the interval I is again an interval: $[\min f(I), \max f(I)]$.

Corollary 5.31. *Let K be compact, $f : K \to \mathbb{R}$ such that $f(x) > 0$ for all $x \in K$. Then* $\inf\{f(x) : x \in K\} > 0$.

Definition 5.32. Let (X, d_X) and (Y, d_Y) be metric spaces and $\mathcal{D} \subseteq X$. A function $f : \mathcal{D} \to Y$ is called *uniformly continuous* if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in \mathcal{D} : \big(d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon\big).$$

Obviously, every uniformly continuous function is continuous.

If f is continuous in $\mathcal{D}$ then the δ in the definition of continuity may depend both on ε and the point x_0 at which the function is considered. If f is uniformly continuous, then the same δ is good enough for all x_0 in $\mathcal{D}$.

Examples.

- Every Lipschitz continuous function is uniformly continuous.
- $f : [0, 1] \to \mathbb{R},\ f(x) = \sqrt{x}$ is uniformly continuous but not Lipschitz continuous (see Exercise 5.6).
- $f : (0, 1] \to \mathbb{R},\ f(x) = \frac{1}{x}$ is not uniformly continuous.

Theorem 5.33. *A continuous function on a compactum is uniformly continuous.*

Proof. Let (Y, d_Y) be a metric space, K a compact subset of a metric space (X, d_X) and $f : K \to Y$ a continuous function. We will show that f is uniformly continuous. Assume that f is not uniformly continuous. Then there exists an $\varepsilon > 0$ such that

$$\forall\, n \in \mathbb{N}\ \exists\, x_n, y_n \in K : \quad d_X(x_n, y_n) < \frac{1}{n} \quad \wedge \quad d_Y(f(x_n), f(y_n)) \geq \varepsilon.$$

Since K is compact, the sequence $(x_n)_{n\in\mathbb{N}}$ contains a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to some $p \in K$. Since $d(x_{n_k}, y_{n_k}) < \frac{1}{n_k}$ for all $k \in \mathbb{N}$ it follows that also the subsequence $(y_{n_k})_{k\in\mathbb{N}}$ converges to p. The continuity of f implies

$$\lim_{k\to\infty} f(x_{n_k}) = f(p) = \lim_{k\to\infty} f(y_{n_k}),$$

in contradiction to the assumption $f(x_{n_k}) - f(x_{n_k}) \geq \varepsilon$ for all $k \in \mathbb{N}$. □

5.3 Sequences and series of functions

In this section we consider sequences and series of functions. We will consider two types of convergence of sequences of functions: pointwise convergence and uniform convergence.

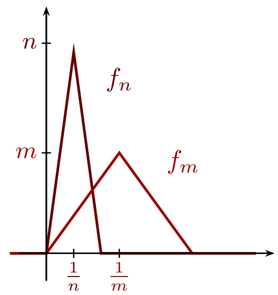
Definition 5.34. Let X be a set, (Y, d_Y) a metric space and $(f_n)_{n\in\mathbb{N}} \subseteq Y^X$ a sequence of functions $f_n : X \to Y$. The sequence *converges pointwise* to the function $f : X \to Y$ if $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in X$, i. e.,

$$\forall \varepsilon > 0 \quad \forall x \in X \quad \exists N \in \mathbb{N} \quad \forall n \geq N : \quad d(f_n(x), f(x)) < \varepsilon.$$

The following example shows that a pointwise convergent function is not necessarily as "close" to the limit function as might be expected.

Example 5.35. Let

$$f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \begin{cases} n^2 x, & 0 \le x \le \frac{1}{n}, \\ 2n - n^2 x, & \frac{1}{n} < x \le \frac{2}{n}, \\ 0, & x > \frac{2}{n} \end{cases}$$



Obviously, $f_n(x) \xrightarrow{n\to\infty} 0$ for all $x \in \mathbb{R}$, hence $f_n \to 0$ pointwise (see 5.7).

The following example shows that the limit function of a pointwise convergent sequence of continuous functions is not necessarily continuous.

Example 5.36. Let $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \frac{nx}{1+|nx|}$.
Obviously, every f_n is continuous in $\mathbb{R}$ and for all $x \in \mathbb{R}$ we have $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) \to \operatorname{sign}(x)$ for $n \to \infty$, hence the pointwise limit function is not continuous.

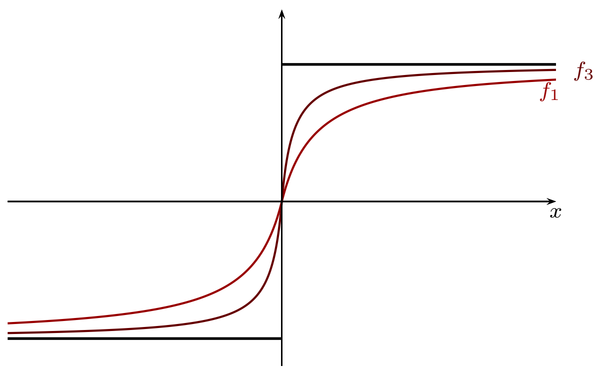


FIGURE 5.4: The pointwise limit of the sequence $(f_n)_{n\in\mathbb{N}}$ with the continuous functions $f_n(x) = nx(1 + |nx|)^{-1}$ is the non-continuous function sign$(\cdot)$, see Example 5.36.

We need a stronger notion of convergence that guarantees that the limit of continuous functions is again continuous.

Definition 5.37. Let X be a set and (Y, d_Y) be a metric space. A sequence $(f_n)_{n\in\mathbb{N}}$ of functions $f_n : X \to Y$ is called *uniformly convergent* to a function $f : X \to Y$ if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in X \quad \forall n \ge N : \quad d(f_n(x), f(x)) < \varepsilon.$$

In contrast to the definition of pointwise convergence, the N depends only on ε, not on x.

Definition 5.38. For a set X and a normed space $(Y, \|\cdot\|)$ we set

$$B(X, Y) = \{f : X \to Y \text{ bounded}\}.$$

The *supremum norm* of a function $f \in B(X, Y)$ is

$$\|f\|_\infty := \sup\{\|f(x)\| : x \in X\}.$$

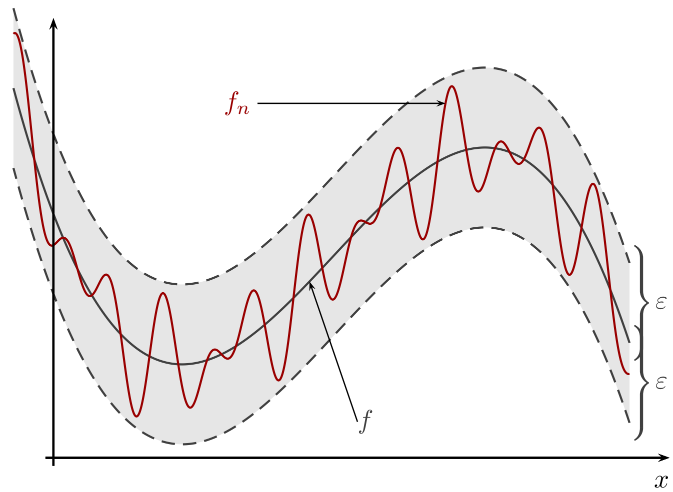


FIGURE 5.5: Uniform convergence (Definition 5.37): For every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that the graphs of all f_n with $n \geq N$ lie in an ε-tube about the graph of the limit function f.

If $f : X \to Y$ is unbounded, we set $\|f\|_\infty := \infty$.

Remark.

- $f_n \to f$ uniformly $\implies$ $f_n \to f$ pointwise.
- $f_n \to f$ pointwise $\iff$ $f_n(x) \to f(x)$, $x \in X$.
- $f_n \to f$ uniformly $\iff$ $\|f_n - f\|_\infty \to 0$.

Theorem 5.39. (i) *$(B(X,Y), \|\cdot\|_\infty)$ is a normed space over $\mathbb{K}$ when Y is a normed space over $\mathbb{K}$.*

(ii) *If Y is a complete normed space, then $(B(X,Y), \|\cdot\|_\infty)$ a complete normed space, i. e. a Banach space.*

Proof. (i) Clearly, $\|f\|_\infty \in \mathbb{R}_+^0$ and $\|\lambda f\|_\infty = |\lambda| \, \|f\|_\infty$ for all $f \in B(X,Y)$ and $\lambda \in \mathbb{K}$. Let $f, g \in B(X,Y)$ and $x \in X$. Then

$$\|(f+g)(x)\| = \|f(x) + g(x)\| \leq \|f(x)\| + \|g(x)\|.$$

Taking the supremum over all $x \in X$ yields the triangle inequality in $B(X,Y)$: $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

(ii) Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $B(X,Y)$. We have to show that it converges to some $f \in B(X,Y)$. Let $\varepsilon > 0$. By assumption, there exists an $N \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty < \frac{\varepsilon}{2}$ for all $m, n \geq N$. In particular, for each $x \in X$, the sequence $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in Y, hence convergent because Y is a Banach space. Therefore, the function

$$f : X \to Y, \qquad f(x) := \lim_{n\to\infty} f_n(x)$$

is well defined. We will show that $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f. For $m, n \geq N$ and $x \in X$ we have

$$\begin{aligned}\|f_n(x) - f(x)\| &\le \|f_n(x) - f_m(x)\| + \|f_m(x) - f(x)\| \\ &\le \|f_n - f_m\|_\infty + \|f_m(x) - f(x)\| < \frac{\varepsilon}{2} + \|f_m(x) - f(x)\|.\end{aligned}$$

Taking the limit $m \to \infty$ yields that $\|f_m(x) - f(x)\| \le \frac{\varepsilon}{2}$. Therefore for $n \ge N$

$$\forall x \in X \quad \|f_n(x) - f(x)\| \le \varepsilon. \tag{5.3}$$

Taking the supremum over all $x \in X$ finally yields $\|f_n - f\|_\infty \le \frac{\varepsilon}{2} < \varepsilon$ for all $n \ge N$. In particular, f is bounded because

$$|f(x)| \le |f_N(x)| + |f(x) - f_N(x)| \le \|f_N\|_\infty + \frac{\varepsilon}{2} < \infty,$$

and $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f by (5.3). □

Theorem 5.40. *Let (X, d_X) be a metric space, $(Y, \|\cdot\|)$ a normed space and $f_n : X \to Y$ continuous and $f : X \to Y$. If $f_n \xrightarrow[\text{unif.}]{n\to\infty} f$, then f is continuous.*

In other words: The uniform limit of continuous functions is continuous.

Proof. Let $x_0 \in X$ and $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $\|f_N - f\|_\infty < \frac{\varepsilon}{3}$. Since f_N is continuous, there exists a $\delta > 0$ such that $\|f_N(x) - f_N(x_0)\| < \frac{\varepsilon}{3}$ for all $x \in X$ with $d_X(x, x_0) < \delta$. Hence we obtain for all $x \in X$ with $d_X(x, x_0) < \delta$

$$\begin{aligned}\|f(x) - f(x_0)\| &\le \|f(x) - f_N(x)\| + \|f_N(x) - f_N(x_0)\| + \|f_N(x_0) - f(x_0)\| \\ &\le \|f - f_N\|_\infty + \|f_N(x) - f_N(x_0)\| + \|f_N - f\|_\infty < \varepsilon.\end{aligned}$$ □

The above theorem shows that for a uniformly convergent sequence of continuous functions the limits commute:

$$\lim_{n\to\infty} \lim_{x\to x_0} f_n(x) = \lim_{x\to x_0} \lim_{n\to\infty} f_n(x).$$

If the sequence $(f_n)_{n\in\mathbb{N}}$ converges only pointwise, then, in general, the limits cannot be commuted, as Example 5.35 shows.

Theorem 5.39 and Theorem 5.40 show that the set of all bounded continuous functions on a metric space X together with the supremum norm are a Banach space. Since every continuous function on a compact metric space is bounded, we obtain

Theorem 5.41. *Let (X, d_X) be a compact metric space and $(Y, \|\cdot\|)$ a normed space. Then $(C(X, Y), \|\cdot\|_\infty)$ is a Banach space.*

Since series are special case of sequences, we have the notion on pointwise and uniform convergence also for series of functions:

$$\begin{aligned}\sum_{n=1}^\infty f_n \text{ converges pointwise} &\iff \forall\, x \in X \quad \sum_{n=1}^\infty f_n(x) \text{ converges in } Y, \\ \sum_{n=1}^\infty f_n \text{ converges uniformly} &\iff \text{the sequence of the partial sums} \\ &\qquad \Big(\sum_{k=1}^n f_k\Big)_{n\in\mathbb{N}} \text{ converges uniformly.}\end{aligned}$$

Since $(B(X, Y), \|\cdot\|_\infty)$ is a Banach space, we obtain the following criterion for convergence of a series of functions.

Theorem 5.42 (Weierstraß criterion). *Let X be a set, $(Y, \|\cdot\|)$ be a complete normed space and $(f_n)_{n\in\mathbb{N}} \subseteq B(X,Y)$. If $\sum_{n=1}^{\infty} \|f_n\|_\infty < \infty$, then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly to a function $f \in B(X,Y)$ and for each $x \in X$ the series $\sum_{n=1}^{\infty} f_n(x)$ converges absolutely in Y.*

Proof. Let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $\sum_{k=m}^{n} \|f_k\|_\infty < \varepsilon$ for all $m, n \geq N$. The triangle inequality yields

$$\Big\| \sum_{k=m}^{n} f_n \Big\|_\infty \leq \sum_{k=m}^{n} \|f_n\|_\infty < \varepsilon, \qquad m, n \geq N.$$

The Cauchy criterion (in the complete normed space $(B(X,Y), \|\cdot\|_\infty)$ shows that the series of functions converges absolutely in $(B(X,Y), \|\cdot\|_\infty)$ which is equivalent to uniform convergence. Since for every $x \in X$

$$\sum_{k=1}^{n} \|f_k(x)\| \leq \sum_{k=1}^{n} \|f_k\|_\infty < \infty,$$

also the assertion on pointwise absolute convergence is proved. □

5.4 Power series

Definition 5.43. Let $a \in \mathbb{C}$ and $(c_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$. Then

$$\sum_{n=0}^{\infty} c_n (z-a)^n \tag{5.4}$$

is called a *power series centred in a* (or a *power series in* $(z-a)$) with coefficients in c_n.
The *radius of convergence* of the power series (5.4) is

$$R := \sup\{t \in \mathbb{R} : (c_n t^n)_{n\in\mathbb{N}} \text{ is bounded}\}.$$

Depending on the coefficients, the series (5.4) converges for all $z \in \mathbb{C}$, for no $z \in \mathbb{C}$, or for z a subset of $\mathbb{C}$.

Theorem 5.44. *Let R be the radius of convergence of the power series (5.4).*

(i) *For $z \in \mathbb{C}$ such that $|z-a| > R$, the series (5.4) diverges.*

(ii) *For $z \in \mathbb{C}$ such that $|z-a| < R$, the series (5.4) is absolutely convergent.*
For $0 < r < R$ On $\overline{B_r(a)}$ the series converges uniformly to the continuous function

$$\overline{B_r(a)} \to \mathbb{C}, \qquad z \mapsto \sum_{n=0}^{\infty} c_n (z-a)^n.$$

Proof. (i) Since by assumption $(c_n |z-a|^n)_{n\in\mathbb{N}}$ is not bounded, the series diverges (Theorem 4.48).
(ii) Let $r \in \mathbb{R}$ such that $r < R$. Then there exists a t such that $r < t < R$. By definition of R there exists an M such that $|c_n t^n| \leq M$ for all $n \in \mathbb{N}$. For each $z \in \mathbb{C}$ with $|z-a| \leq r < t$ we obtain

$$|c_n(z-a)^n| \leq |c_n r^n| \leq |c_n t^n| \Big(\frac{r}{t}\Big)^n \leq M \Big(\frac{r}{t}\Big)^n.$$

Therefore, $\|c_n(\cdot - a)\|_\infty \leq A\big(\frac{r}{t}\big)^n$ where $\|\cdot\|_\infty$ is the supremum norm of bounded functions on $\overline{B_r(a)}$. By the Weierstraß criterion (Theorem 5.42) the series of polynomials (5.4) converges uniformly on

$\overline{B_r(a)}$ and for fixed z the series converges absolutely. Since all polynomials are continuous, the function

$$\overline{B_r(a)} \to \mathbb{C}, \quad z \mapsto \sum_{n=0}^{\infty} c_n (z-a)^n$$

is continuous (Theorem 5.40). □

Theorem 5.45. *Let R be the radius of convergence of the power series (5.4). Then*

(i) $R = \big(\limsup_n \sqrt[n]{|c_n|}\big)^{-1}$,

(ii) $R = \lim_{n\to\infty} \dfrac{|c_n|}{|c_{n+1}|}$ *if the limit exists*

Proof. (i) Let $\widetilde{R} := \big(\limsup_n \sqrt[n]{|c_n|}\big)^{-1}$. $\widetilde{R} = R$ follows immediately from the root test and the characterisation of R in Theorem 5.44:

$$\limsup_n \sqrt[n]{|c_n(z-a)^n|} = |z-a| \limsup_n \sqrt[n]{|c_n|} = |z-a|\widetilde{R} \quad \begin{cases} < 1 & \text{if } |z-a| < \widetilde{R}^{-1}, \\ > 1 & \text{if } |z-a| > \widetilde{R}^{-1} \end{cases}$$

with the convention that $\widetilde{R}^{-1} = 0$ if $\widetilde{R} = \infty 0$ and $\widetilde{R}^{-1} = \infty$ if $\widetilde{R} = 0$.
(ii) follows analogously with the ratio test. □

The following theorem follows immediately from Theorem 4.48 and Theorem 4.69 (Cauchy product).

Theorem 5.46. *Let*

$$\sum_{n=0}^{\infty} b_n (z-a)^n \qquad \textit{and} \qquad \sum_{n=0}^{\infty} c_n (z-a)^n$$

complex power series in $(z-a)$ with radii of convergence R_b and R_c respectively. Then for all $z \in \mathbb{C}$ with $|z-a| < \min\{R_c, R_d\}$

$$\Big(\sum_{n=0}^{\infty} b_n (z-a)^n\Big) + \Big(\sum_{n=0}^{\infty} c_n (z-a)^n\Big) = \sum_{n=0}^{\infty} (b_n + c_n)(z-a)^n,$$

$$\Big(\sum_{n=0}^{\infty} b_n (z-a)^n\Big) \cdot \Big(\sum_{n=0}^{\infty} c_n (z-a)^n\Big) = \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} c_k d_{n-k}\Big)(z-a)^n.$$

Let R be the radius of convergence of (5.4). We know that for $|z-a| < R$ the series is absolutely convergent and that for $|z-a| > R$ it is divergent. For $|z-a| = R$ the series can diverge or converge.

Examples 5.47. Even when the function represented by a power series is continuous in the limit points of the interval of convergence, the series does not need to converge.

(i) $\sum_{n=0}^{\infty} z^n$. The radius of convergence is $R = 1$. The series diverges for $z = \pm 1$.
Note that $\sum_{n=1} z^n = \frac{1}{1+z}$ for $|z| < 1$ and $\frac{1}{1-(-1)} = \frac{1}{2}$.

(ii) $\sum_{n=0}^{\infty} z^{2n}$. The radius of convergence is $R = 1$. The series diverges for $z = \pm 1$ and $\sum_{n=1} z^{2n} = \frac{1}{1-z^2}$ for $|z| < 1$ and $\frac{1}{1+(\pm 1)^2} = \frac{1}{2}$.

(iii) $\displaystyle\sum_{n=0}^{\infty} \frac{z^n}{n}$. The radius of convergence is $R = 1$. The series diverges for $z = 1$ and is conditionally convergent for $z = -1$.

From Theorem 5.44 we already know that a power series defines a continuous function on the the open ball of convergence. Next we will show that f is continuous in that points z with $|z - a| = R$ for which the power series converges.
To this end we use

Remark (Summation by parts). Let $(a_n)_{n\in\mathbb{N}_0}$ and $(b_n)_{n\in\mathbb{N}_0}$ be sequences in $\mathbb{C}$ and define

$$A_{-1} := 0, \qquad A_n := \sum_{k=1}^{n} a_k, \quad n \in \mathbb{N}.$$

Then obviously we have $a_n = A_n - A_{n=1} =: \Delta A_n$. The following rules can be verified straighforwardly:

(i) Product rule: $\underbrace{\Delta(AB)_k}_{:=A_kB_k - A_{k-1}B_{k-1}} = (\Delta A_k)B_k + A_{k-1}\Delta B_k$,

(ii) $\displaystyle\sum_{k=0}^{n} \Delta A_k \cdot B_k = A_nB_n - \sum_{k=0}^{n} A_{k-1}\Delta B_k$,

(iii) $\displaystyle\sum_{k=0}^{n} a_k \cdot B_k = A_nB_n + \sum_{k=0}^{n-1} A_k\Delta B_{k+1}$,

(iv) Then for $0 \le m \le n$ it follows that $\displaystyle\sum_{k=m}^{n} a_kb_k = \sum_{k=m}^{n-1} A_k(b_k - b_{k+1}) + A_nb_n - A_{m-1}b_m$.

Theorem 5.48 (Abel's theorem). *Let $(c_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and assume that the power series*

$$\sum_{n=0}^{\infty} c_n(x-a)^n \tag{5.5}$$

converges in $I := [a - R,\, a + R]$. Then the series converges uniformely in I and its limit is a continuous function.

Proof. It suffices to prove the uniform convergence because all partial sums are polynomials, hence continuous, and by Theorem 5.40 the uniform limit of continuous functions is continuous. Obviously, it suffices to show uniform convergence on $[a,\, a+R]$ and $[a-R,\, a]$. Let us apply the transformation of the variable x

$$\xi := \frac{x-a}{R}$$

Then, obviously, the series in (5.5) converges uniformly on $[a, a+R]$ if and only if

$$\sum_{n=0}^{\infty} c_n\xi^n$$

converges uniformly on $[0, 1]$. Let us show uniform convergence on $[0, 1]$. To this end fix $\varepsilon > 0$ and choose $m \in \mathbb{N}$ such that for all $n \ge m$

$$\left|\sum_{k=m}^{n} c_k\right| < \varepsilon.$$

Such m exists because the series $\sum_{k=0}^{n} c_k = \sum_{k=0}^{n} 1^k c_k$ converges by assumption. Now define

$$a_k := \begin{cases} 0, & k < m, \\ c_k, & k \geq m. \end{cases}$$

Then

$$\sum_{k=0}^{\infty} c_k \xi^k - \sum_{k=0}^{m-1} c_k \xi^k \sum_{k=0}^{\infty} a_k \xi^k.$$

Summation by parts applied to $A_k = \sum_{n=1}^{k} a_k$, $B_k = \xi^k$ for $\xi \in [0, 1]$ yields

$$\begin{aligned} \sum_{k=0}^{n} \underbrace{a_k}_{=\Delta A_k} \underbrace{\xi^k}_{B_k} &= A_n B_n - \sum_{k=0}^{n-1} A_k (\xi^k - \xi^{k+1}) = A_n \xi^n - \sum_{k=0}^{n-1} A_k (\xi^k - \xi^{k+1}) \\ &= A_n \xi^n - (1 - \xi) \sum_{k=0}^{n-1} A_k \xi^k. \end{aligned}$$

Using that $|A_k| < \varepsilon$ by assumption we obtain for $0 \leq \xi < 1$

$$\left| \sum_{k=0}^{n} a_k \xi^k \right| \leq |A_n| \xi^n + (1 - \xi) \sum_{k=0}^{n-1} |A_k| \xi^k \leq \varepsilon \xi^n + (1 - \xi) \frac{\varepsilon}{1 - \xi} = 2\varepsilon.$$

Obviously, the inequality is also true in the case $\xi = 1$. In summary, we showed that the series converges uniformly on $[0, 1]$, hence the series in (5.5) converges uniformly on $[a, a + R]$. To show that it converges uniformly on $[a - R, a]$, we apply the substitution

$$\xi := -\frac{x - a}{R}. \qquad \square$$

The exponential function is defined as a power series.

Definition 5.49. The *exponential function* is defined by

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

and the sine and cosine functions are defined by

$$\sin : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \sum_{n=0}^{\infty} \frac{(-)^n z^{2n+1}}{(2n+1)!}, \qquad \cos : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \sum_{n=0}^{\infty} \frac{(-)^n z^{2n}}{(2n)!}.$$

These functions are well-defined by Theorem 5.45 and continuous in $\mathbb{C}$ by Theorem 5.44.

Theorem 5.50 (Properties of $\exp$**).** *For the function* $\exp : \mathbb{C} \to \mathbb{C}$*,* $\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$ *and the Euler's number* e *defined in Theorem 4.71 gilt:*

(i) $\exp(\bar{z}) = \overline{\exp(z)}, \quad z \in \mathbb{C}$,

(ii) $\exp(z + w) = \exp(z) \exp(w), \quad z, w \in \mathbb{C}$,

(iii) $\exp(n) = \mathrm{e}^n, \quad n \in \mathbb{Z}$,

(iv) $\exp(z) \neq 0, \quad z \in \mathbb{C}$,

(v) $|\exp(\mathrm{i}x)| = 1, \quad x \in \mathbb{R}$.

Proof. Exercise 5.11. □

Theorem 5.51 (Euler's formula).

$$\exp(\mathrm{i}z) = \cos(z) + \mathrm{i}\sin(z), \qquad z \in \mathbb{C}, \tag{5.6}$$

consequently

$$\cos(z) = \frac{1}{2}\big(\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)\big), \quad \sin(z) = \frac{1}{2\mathrm{i}}\big(\exp(\mathrm{i}z) - \exp(-\mathrm{i}z)\big). \tag{5.7}$$

In particular, it follows for all $x \in \mathbb{R}$ that $\exp(x) = \big(\exp(x/2)\big)^2 \geq 0$. Directly from the defintion of exp we obtain that it is monotonically increasing in $[0, \infty)$. Using the fact that $\exp(-x) = \big(\exp(x)\big)^{-1}$ and that exp is positive, it follos that exp is monotonically increasing in $\mathbb{R}$.

Proof. Let $z \in \mathbb{C}$. Since exp, sin and cos are absolutely convergent on $\mathbb{C}$, we have

$$\begin{aligned}
\exp(\mathrm{i}z) &= \lim_{n\to\infty} \sum_{k=0}^{2n} \frac{(\mathrm{i}z)^k}{k!} = \lim_{n\to\infty} \Big(\sum_{k=0}^{n} \frac{(\mathrm{i}z)^{2k}}{(2k)!} + \sum_{k=0}^{n-1} \frac{(\mathrm{i}z)^{2k+1}}{(2k+1)!} \Big) \\
&= \lim_{n\to\infty} \Big(\sum_{k=0}^{n} \frac{(-1)^k z^{2k}}{(2k)!} + \sum_{k=0}^{n-1} \frac{(-1)^k z^{2k+1}}{(2k+1)!} \Big) \\
&= \lim_{n\to\infty} \Big(\sum_{k=0}^{n} \frac{(-1)^k z^{2k}}{(2k)!} \Big) + \mathrm{i} \lim_{n\to\infty} \Big(\sum_{k=0}^{n-1} \frac{(-1)^k z^{2k+1}}{(2k+1)!} \Big) = \cos(z) + \mathrm{i}\sin(z),
\end{aligned}$$

Formulae (5.7) follow because $\sin(z) = \sin(-z)$ and $\cos(-z) = \cos(z)$ which follows directly from the definition. □

Definition 5.52. The function $\mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x)$ is continuous and by Theorem 5.50 monotonically increasing with range $\mathrm{R}(\exp) = \mathbb{R}_+$, hence by Theorem 5.26 it is invertible and the inverse is continuous. The inverse function is called the (*natural*) *logarithm* denoted by

$$\ln : (0, \infty) \to \mathbb{R}.$$

Remark. Sometimes the logarithm is denoted by log instead of ln.

Uniqueness of the power series representation

For the proof of the uniqueness of the power series representation of a function we need the following technical lemma.

Lemma 5.53. *Let $\sum_{n=0}^{\infty} c_n(z-a)^n$ be a comples power series with radius of convergence R. Then for every $m \in \mathbb{N}_0$ and every $r \in (0, R)$ there exists an $M > 0$ such that for all $z \in \mathbb{C}$ with $|z-a| \leq r$:*

$$\Big| \sum_{n=m}^{\infty} c_n(z-a)^n \Big| \leq M|z-a|^m.$$

Proof. Let $m \in \mathbb{N}$. Then the series $\sum_{n=m}^{\infty} c_n(z-a)^n$ and $\sum_{n=m}^{\infty} c_n(z-a)^{n+m}$ have the same radius of convergence. For $|z-a| \le r$ Lemma 4.60 implies

$$\Big|\sum_{n=m}^{\infty} c_n(z-a)^n\Big| \le |z-a|^m \sum_{n=m}^{\infty} |c_n|\underbrace{|z-a|}_{\le r}{}^{n-m} \le |z-a|^m \sum_{n=m}^{\infty} |c_n| r^{n-m}$$
$$= |z-a|^m \underbrace{\sum_{n=0}^{\infty} |c_{n+m}| r^n}_{=:M<\infty,\ \text{since } r<R} . \qquad \square$$

Theorem 5.54. *Let $\sum_{n=0}^{\infty} b_n(z-a)^n$ and $\sum_{n=0}^{\infty} c_n(z-a)^n$ be complex power series with radii of convergence R_b and R_c respectively. If*

$$\sum_{n=0}^{\infty} b_n(z-a)^n = \sum_{n=0}^{\infty} c_n(z-a)^n, \quad |z-a| \le r,$$

for some $0 < r \le \min\{R_b, R_c\}$, then $a_n = b_n,\ n \in \mathbb{N}_0$.

Proof. Without restriction we assume $a = 0$. By Theorem 5.44 it suffices to show that $\sum_{n=0}^{\infty} b_n z^n = 0$ for all $|z| \le r$ implies $b_n = 0,\ n \in \mathbb{N}_0$. Let $N := \min\{n \in \mathbb{N} : c_n \ne 0\}$. Then, by the proof of Theorem 5.44, there exists an $M > 0$ such that for all $z \in \mathbb{C}$ with $|z| \le r$

$$b_N z^N = \Big|\underbrace{\sum_{n=0}^{\infty} b_n z^n}_{=0} - b_N z^N\Big| = \Big|\sum_{n=N+1}^{\infty} b_n z^n\Big| \le z^{N+1} M,$$

in particular $|z| \ge \frac{b_N}{M}$ for all $0 < |z| \le r$. Since $|z|$ can be chosen arbitrarily small, this implies $b_N = 0$. $\square$

Another proof for the uniqueness of the power series representations follows from the Taylor expansion.

Chapter 6

Integration and Differentiation in $\mathbb{R}$

6.1 Differentiable functions

Continuity of a function f in a point x_0 implies that the function values $f(x)$ do not deviate too much from $f(x_0)$ if x is close to x_0.
In this section we investigate the local behaviour of functions further. We will consider mainly functions $f : \mathcal{D} \supseteq \mathbb{R} \to \mathbb{R}$. A function is called differentiable in a point x_0 if it can be approximated by an affine function. More generally, a function is n times differentiable if it can be approximated locally by a polynomial of degree n. This is the main assertion of Taylor's theorem.

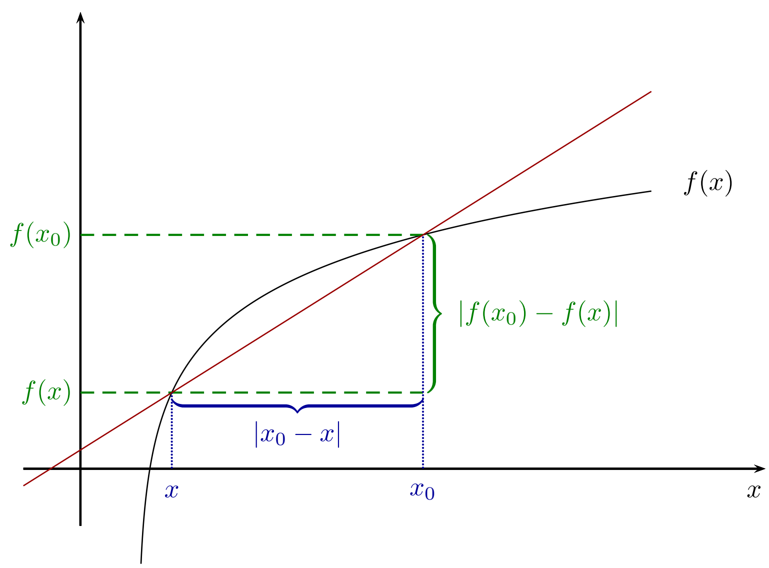


FIGURE 6.1: Geometric interpretation of the difference quotient in the case $\mathbb{F} = Y = \mathbb{R}$: The difference quotient $\dfrac{f(x) - f(x_0)}{x - x_0}$ is the slope of the secant of the graph of f through the points $(x_0, f(x_0))$ and $(x, f(x))$. For $x \to x_0$ the secant becomes the tangent of the graph of f in the point $(x_0, f(x_0))$; $f'(x_0)$ is the slope of the tangent.

Definition 6.1. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $(Y, \|\cdot\|)$ a normed space over $\mathbb{F}$ and $x_0 \in \mathcal{D}$ a limit point of the set $\mathcal{D} \subseteq \mathbb{F}$. A function $f : \mathcal{D} \to Y$ is called *differentiable in* x_0 if there exists a function $\Phi : \mathcal{D} \to Y$ continuous in x_0 such that

$$f(x) - f(x_0) = \Phi(x)(x - x_0), \qquad x \in \mathcal{D}. \tag{6.1}$$

Then $\Phi(x_0) =: f'(x_0)$ is called the *derivative of* f *at* x_0. The function is called *differentiable* if every point of $\mathcal{D}$ is a limit point of $\mathcal{D}$ and f is differentiable in every point $x_0 \in \mathcal{D}$. In this case, the function

$$f' : \mathcal{D} \to Y, \qquad x \mapsto f'(x)$$

is called the *derivative* of f.

Note that the function Φ depends on f and x_0.

Theorem 6.2. *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $(Y, \|\cdot\|)$ a normed space over $\mathbb{F}$. Let $x_0 \in \mathcal{D} \subseteq \mathbb{F}$ such that x_0 is a limit point of $\mathcal{D}$ and let $f : \mathcal{D} \to Y$. Then the following is equivalent:*

(i) *f is differentiable in x_0.*

(ii) *There exists an $a \in Y$ and a function $\varphi : \mathcal{D} \to Y$ which is continuous in x_0 with $\varphi(x_0) = 0$ and*

$$f(x) = f(x_0) + a(x - x_0) + \varphi(x)(x - x_0), \qquad x \in \mathcal{D}. \tag{6.2}$$

(iii) *The following limit exists:*

$$b := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{6.3}$$

If f is differentiable in x_0, then $f'(x_0) = a = b$.

Proof. "(i) $\implies$ (ii)" Let $a := f'(x_0)$ and $\varphi : \mathcal{D} \to Y$, $\varphi(x) = \Phi(x) - f'(x_0)$. Then φ is continuous in x_0 and $\varphi(x_0) = \Phi(x_0) - f'(x_0) = 0$ by definition of $f'(x_0)$ and obviously φ satisfies (6.2).
"(ii) $\implies$ (iii)" By assumption

$$a = \lim_{x \to x_0} \varphi(x) + a = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = b.$$

"(iii) $\implies$ (i)" Since the limit in (6.3) exists, the function

$$\Phi : \mathcal{D} \to Y, \quad \Phi(x) := \begin{cases} \frac{f(x) - f(x_0)}{x - x_0}, & \text{if } x \neq x_0, \\ b, & \text{if } x = x_0 \end{cases}$$

is continuous in x_0. Obviously it satisfies (6.1) and $f'(x_0) = \Phi(x_0) = b$. □

The characterisation of differentiability in Definition 6.1 is useful for proofs and can be extended to functions f between normed spaces. The characterisation (ii) of Theorem 6.2 gives a geometric interpretation of the derivative (see Remark 6.4) and (iii) is useful to calculate derivatives of functions.

Corollary 6.3. *If f is differentiable in x_0 then f is continuous in x_0.*

Proof. This follows immediately from (6.1) because

$$\lim_{x \to x_0} (f(x) - f(x_0)) = \big(\lim_{x \to x_0} \Phi(x)\big)\big(\lim_{x \to x_0} (x - x_0)\big) = 0.$$ □

Note that the converse is not true, for example the absolute value function on $\mathbb{R}$ is continuous in 0 but not differentiable. There exist functions that are continuous on $\mathbb{R}$ but nowhere differentiable, for example the Weierstraß function $f(x) = \sum_{n=0}^{\infty} \frac{\cos(15^k \pi x)}{2^k}$, $x \in \mathbb{R}$.

Notation. Other notations for $f'(x_0)$ and f' are $\frac{\mathrm{d}}{\mathrm{d}x} f(x_0)$, $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$, $\mathrm{D}f(x_0)$ and $\frac{\mathrm{d}}{\mathrm{d}x} f$, $\frac{\mathrm{d}f}{\mathrm{d}x}$, $\mathrm{D}f$, respectively.

Remark 6.4. Theorem 6.2 shows that f is differentiable in x_0 if and only if it can be approximated by a linear function at x_0, that is, there exists a linear function

$$L : \mathbb{F} \to Y, \qquad L(x) = f(x_0) + a(x - x_0)$$

such that $f(x) - L(x)$ tends to 0 faster than $x - x_0$ for $x \to x_0$. The constant a is then $f'(x_0)$.

Remark 6.5. The space of all linear functions from $\mathbb{F}$ to Y is denoted by $L(\mathbb{F}, Y)$. Note that every $a \in Y$ induces the linear map $\mathbb{F} \to Y,\ x \mapsto ax$.
Let $f : \mathbb{F} \supseteq \mathcal{D} \to Y$ be differentiable. The *differential* $\mathrm{d}f$ of f is the map

$$\mathrm{d}f : \mathcal{D} \to L(\mathbb{F}, \mathbb{F}), \qquad x \mapsto \mathrm{d}_x f : \mathbb{F} \to \mathbb{F},\ h \mapsto f'(x)h.$$

Since the differential $\mathrm{d}x$ of the function $\mathbb{F} \to \mathbb{F},\ x \mapsto x$ is the identity, it follows that $\mathrm{d}f = f' \mathrm{d}x$.

Examples 6.6.

- $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^n$ for $n \in \mathbb{N}_0$ is differentiable in $\mathbb{R}$ with $f'(x) = 0$ if $n = 0$ and

$$f'(x) = nx^{n-1}, \qquad n \geq 1.$$

Proof. For $n = 0$ the assertion is clear. Now let $n \geq 1$ and fix $x_0 \in \mathbb{R}$. For $x \in \mathbb{R} \setminus \{0, x_0\}$ it follows from the formula for the geometric sum (4.5) that

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^n - x_0^n}{x - x_0} = \frac{x^n}{x} \frac{1 - (\frac{x_0}{x})^n}{1 - \frac{x_0}{x}} = x^{n-1} \sum_{j=0}^{n-1} \Big(\frac{x_0}{x}\Big)^j.$$

For $x \to x_0$ this tends to nx_0^{n-1}. □

- $f : [0, \infty) \to \mathbb{R},\ f(x) = \sqrt{x}$ is differentiable in $(0, \infty)$ with

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

It is not differentiable in 0 (see Exercise 6.1).

- $f : \mathbb{R} \to \mathbb{R},\ f(x) = |x|$ is not differentiable in 0.

Proof. $\displaystyle\lim_{x \searrow 0} \frac{|x| - 0}{x - 0} = 1 \neq -1 = \lim_{x \nearrow 0} \frac{|x| - 0}{x - 0}$. □

Example 6.7. The exponential function $\mathbb{C} \to \mathbb{C}, z \mapsto \exp(z)$ is differentiable with derivative $\exp' = \exp$.

Proof. First we show that exp is differentiable in $z = 0$. For $z \neq 0$

$$\frac{\exp(z) - \exp(0)}{z - 0} = \frac{\exp(z) - 1}{z} = \frac{1}{z}\Big(\sum_{k=0}^{\infty} \frac{z^k}{k!} - 1\Big) = \frac{1}{z}\sum_{k=1}^{\infty} \frac{z^k}{k!} = \sum_{k=0}^{\infty} \frac{z^k}{(k+1)!}.$$

It is easy to see that radius of convergence of the last series is ∞, therefore it is uniformly convergent which implies

$$\exp'(0) = \lim_{z\to 0} \sum_{k=0}^{\infty} \frac{z^k}{(k+1)!} = \sum_{k=0}^{\infty} \lim_{z\to 0} \frac{z^k}{(k+1)!} = 1.$$

Now let $z_0 \in \mathbb{C}$ arbitrary. It follows that

$$\begin{aligned} \exp'(z_0) = \frac{\exp(z) - \exp(z_0)}{z - z_0} &= \exp(z_0)\frac{\exp(z - z_0) - 1}{z - z_0} \\ &\xrightarrow{z\to z_0} \exp(z_0)\exp'(0) = \exp(z_0). \end{aligned}$$

□

Theorem 6.8. *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $(Y, \|\cdot\|_Y)$ a normed space over $\mathbb{F}$. Let $x_0 \in \mathcal{D} \subseteq \mathbb{F}$ such that x_0 is a limit point of $\mathcal{D}$ and assume that $f, g : \mathcal{D} \to Y$ are differentiable in x_0. Then*

(i) *For all $\alpha \in \mathbb{F}$ the linear combination $\alpha f + g$ is differentiable in x_0 with*

$$(\alpha f + g)'(x_0) = \alpha f'(x_0) + g'(x_0). \tag{6.4}$$

(ii) *If $Y = \mathbb{F}$ then the product fg is differentiable in x_0 with*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0). \tag{6.5}$$

(iii) *If $g(x_0) \neq 0$ then the function $\frac{f}{g}$ is differentiable in x_0 with*

$$\Big(\frac{f}{g}\Big)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}. \tag{6.6}$$

Proof. Let Φ_f and Φ_g as in (6.2), that is, Φ_f and Φ_g are continuous in x_0 and

$$\begin{aligned} f(x) - f(x_0) &= \Phi_f(x)(x - x_0), \qquad x \in \mathcal{D}, \\ g(x) - g(x_0) &= \Phi_g(x)(x - x_0), \qquad x \in \mathcal{D}. \end{aligned}$$

(i) follows from

$$\begin{aligned} (\alpha f + g)(x) - (\alpha f + g)(x_0) &= \alpha(f(x) - f(x_0)) + g(x) - g(x_0) \\ &= \big[\underbrace{\alpha\Phi_f(x) + \Phi_g(x)}_{:=\Phi_{\alpha f+g}(x)}\big](x - x_0). \end{aligned}$$

Since $\Phi_{\alpha f+g}$ is continuous in x_0 and tends to $\alpha f'(x_0) + g'(x_0)$ for $x \to x_0$, the function $\alpha f + g$ is differentiable in x_0 by (6.2) and (6.4) holds.

(ii) follows similarly:

$$\begin{aligned} (fg)(x) - (fg)(x_0) &= f(x)g(x) - f(x_0)g(x_0) \\ &= \big(f(x) - f(x_0)\big)g(x) + f(x_0)\big(g(x) - g(x_0)\big) \\ &= \big[\underbrace{\Phi_f(x)g(x) + f(x_0)\Phi_g(x)}_{:=\Phi_{fg}}\big](x - x_0). \end{aligned}$$

Since Φ_{fg} is continuous at x_0 and it tends to $f'(x_0)g(x_0)+f(x_0)g'(x_0)$ for $x\to x_0$, the function fg is differentiable in x_0 and (6.5) holds.

(iii) by (ii) it suffices to show that $\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}$. This follows from

$$\frac{1}{g(x)} - \frac{1}{g(x_0)} = \frac{g(x_0)-g(x)}{g(x)g(x_0)} = \frac{\Phi_g(x)}{g(x)g(x_0)}\,(x-x_0)$$

because $\frac{\Phi_g(x)}{g(x)g(x_0)} \to \frac{g'(x_0)}{g(x_0)^2}$ for $x\to x_0$. □

Corollary 6.9. *Polynomials and rational functions are differentiable.*

Theorem 6.10 (Chain rule). *Let $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$ and $f:\mathbb{F}\supseteq\mathcal{D}_f\to\mathbb{F}$, $g:\mathbb{F}\supseteq\mathcal{D}_g\to Y$ functions such that $f(\mathcal{D}_f)\subseteq\mathcal{D}_g$. Let $x_0\in\mathcal{D}_f$ be a limit point of $\mathcal{D}$ and $f(x_0)$ be a limit point of $\mathcal{D}_g$. If f is differentiable in x_0 and g is differentiable in $f(x_0)$ then $g\circ f$ is differentiable in x_0 with derivative*

$$(g\circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. By assumption on f and g there exist functions $\Phi_f:\mathcal{D}_f\to\mathbb{F}$ continuous in x_0 and $\Phi_g:\mathcal{D}_g\to\mathbb{R}$ continuous in $f(x_0)$ such that

$$\begin{aligned} f(x)-f(x_0) &= \Phi_f(x)(x-x_0), \qquad & \Phi_f(x_0) &= f'(x_0),\\ g(x)-g(x_0) &= \Phi_g(x)(x-x_0), \qquad & \Phi_g(x_0) &= g'(x_0). \end{aligned}$$

Therefore

$$\begin{aligned} (g\circ f)(x)-(g\circ f)(x_0) &= g(f(x))-g(f(x_0)) = \Phi_g(f(x))\big(f(x)-f(x_0)\big)\\ &= \underbrace{\Phi_g(\Phi_f(x))\Phi_f(x)}_{:=\Phi_{g\circ f(x)}}(x-x_0). \end{aligned}$$

Since $\Phi_{g\circ f}$ is continuous in x_0 and tends to $g'(f(x_0))f'(x_0)$ for $x\to x_0$, the assertion is proved. □

Examples.

- $f:\mathbb{R}_+\to\mathbb{R}$, $f(x)=\sqrt{x^3+42x+7}$.
  The function f is a composition of differentiable functions, therefore it is differentiable. Using chain rule we obtain

$$f'(x) = \frac{1}{2}\frac{1}{\sqrt{x^3+42x+7}}\big(3x^2+42\big).$$

- $f:\mathbb{R}_+\to\mathbb{R}$, $f(x)=\sqrt{x}^3+\sqrt{42x}+7$.
  As a composition of differentiable functions, f is differentiable. Chain rule yields

$$f'(x) = 3\sqrt{x}^2\cdot\frac{1}{2}\frac{1}{\sqrt{x}} + \frac{1}{2}\frac{1}{\sqrt{42x}}42 = \frac{3}{2}\sqrt{x} + \frac{\sqrt{42}}{2\sqrt{x}}.$$

For functions defined on intervals in $\mathbb{R}$ we can define one-sided differentiability.

Definition 6.11. Let $(Y,\|\cdot\|)$ be a normed space over $\mathbb{R}$ and $\mathcal{D}\subseteq\mathbb{R}$, $x_0\in\mathcal{D}$ such that x_0 is a limit point of $\mathcal{D}\cap[x_0,\infty)$. Then f is called *differentiable from the right* if there exists a function $\Phi:\mathcal{D}\cap[x_0,\infty)\to\mathbb{R}$, continuous in x_0 such that $f(x)-f(x_0)=\Phi(x)(x-x_0)$ for all $x\in\mathcal{D}\cap[x_0,\infty)$. In this case, $f'_+(x_0):=\Phi(x_0)$ is called the *derivative from the right of f in x_0*. f is called *differentiable from the right* if it is so in every point $x\in\mathcal{D}$. The *derivative from the left* is defined similarly.

Definition 6.12. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $(Y, \|\cdot\|)$ a normed space over $\mathbb{F}$ and $x_0 \in \mathcal{D} \subseteq \mathbb{F}$ such that x_0 is a limit point of $\mathcal{D}$. We define $f^{[0]}(x_0) = f(x_0)$. If f is differentiable in x_0 we set $f^{[1]}(x_0) = f'(x_0)$. Inductively, higher order derivatives are defined: Assume that $f^{[0]}, f^{[1]}, \dots, f^{[n-2]}$ are differentiable in $\mathcal{D}$ and that $f^{[n-1]}$ is differentiable in x_0, then f is called *n-times differentiable in x_0* and

$$f^{[n]}(x_0) := \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x_0) := \big(f^{[n-1]}\big)'(x_0)$$

is the *nth derivative of f at x_0*. The function f is called *n-times differentiable* if it is n-times differentiable in every $x \in \mathcal{D}$. In this case, the function

$$f^{[n]} : \mathcal{D} \to Y, \quad x \mapsto f^{[n]}(x)$$

is the *nth derivative of f*.

If the nth derivative of f is continuous, then f is called *n-times continuously differentiable*.

The following vector spaces of functions are defined:

$$C^n(D) := C^n(D, Y) := \{f : D \to Y \,:\, f \text{ is } n\text{-times continuously differentiable}\},$$
$$C^\infty(D) := \bigcap_{n=0}^{\infty} C^n(D, Y).$$

Remark. Obviously $C^\infty \subseteq C^{n+1}(D) \subseteq C^n(D, Y) \subseteq C^0(D) = C(D), \quad n \in \mathbb{N}$.

Differentiation in Banach spaces

In this section the definition of differentiabilty is generalized to functions f between (subsets of) normed spaces. All normed spaces in this subsection are assumed to be real or complex vector spaces.

Definition. Let X and Y be normed spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A map $T : X \to Y$ is called *linear* if for all $x, y \in X$ and $\lambda \in \mathbb{F}$

$$T(x + \lambda y) = T(x) + \lambda T(y).$$

The linear map $T : X \to Y$ is called *bounded* if and only if

$$\|T\| := \sup\{\|Tx\| : x \in X,\ \|x\| = 1\} < \infty.$$

In this case, $\|T\|$ is called the *norm* of T. The set of all bounded linear maps from X to Y is denoted by $L(X, Y)$. It is easy to check that $(L(X, Y), \|\cdot\|)$ is a normed space over $\mathbb{F}$.

Remark. Let $T \in L(X, Y)$.

(i) $\|Tx\| \leq \|T\|\,\|x\|$ for all $x \in X$.

Proof. If $x \neq 0$, then $\|Tx\| = \|T\frac{x}{\|x\|}\|\,\|x\| \leq \|T\|\,\|x\|$. The assertion is clear if $x = 0$. □

(ii) If Y is a Banach space, then $L(X, Y)$ is a Banach space.

Proof. Note that for a linear map $T \in L(X, Y)$ its restriction to the unit ball B_X in X is bounded and that $\|T\| = \|T|_{B_X}\|_\infty$ (i. e., the norm of T as a linear map is equal to the supremum norm of the restriction of T to B_X). Let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L(X, Y)$. Then the sequence of the restrictions to B_X are a Cauchy sequence in $B(B_X, Y)$ (the set of all bounded functions from

B_X to Y with the supremum norm). By Theorem 5.39 there exists an $\widetilde{T} \in B(B_X, Y)$ such that the restrictions of T_n converge uniformly to $\widetilde{T}$. $\widetilde{T}$ can be extended to a linear function T on X by setting $T0 = 0$ and $Tx = \|x\|\widetilde{T}\frac{x}{\|x\|}$. It is not hard to check that T is well-defined, linear, bounded and that $\|T_n - T\| \to 0$ for $n \to \infty$. □

(iii) If $\dim X < \infty$ then every linear function $T : X \to Y$ is bounded because the unit ball B_X in X is compact (by the Heine-Borel theorem, Theorem 8.33).

Now let us assume that X is a vector space over $\mathbb{F}$ with an inner product $\langle \cdot \,, \cdot \rangle$. Then X becomes a normed space if we set $\|x\| = \langle x \,, x \rangle$ for all $x \in X$.

Definition 6.1'. Let X be a Banach space over $\mathbb{F}$ and Y be normed space over $\mathbb{F}$, $\mathcal{D} \subseteq X$ and $x_0 \in \mathcal{D}$ a limit point of $\mathcal{D}$. A function $f : \mathcal{D} \to Y$ is called *differentiable in* x_0 if there exists a function $\Phi : \mathcal{D} \to Y$ continuous in x_0 such that

$$f(x) - f(x_0) = \Phi(x)(x - x_0). \tag{6.1'}$$

Then $\Phi(x_0) =: f'(x_0)$ is called the *Fréchet derivative of* f *at* x_0. The function is called *differentiable* if every point of $\mathcal{D}$ is a limit point and f is differentiable in every point $x_0 \in \mathcal{D}$. In this case, the function

$$f' : \mathcal{D} \to L(X, Y), \qquad x \mapsto f'(x)$$

is called the *Fréchet derivative* of f.

Note that the function Φ depends on f and x_0 and that $f'(x_0) \in L(X, Y)$.

Theorem 6.2'. Let X and Y be normed spaces. Let $x_0 \in \mathcal{D} \subseteq X$ such that x_0 is a limit point of $\mathcal{D}$ and let $f : \mathcal{D} \to Y$. Then the following is equivalent:

(i) f is differentiable in x_0.

(ii) There exists an $A \in L(X, Y)$ and a function $\varphi : \mathcal{D} \to Y$ which is continuous in x_0 with $\lim\limits_{x \to x_0} \frac{\varphi(x)}{\|x - x_0\|} = 0$ and

$$f(x) = f(x_0) + A(x - x_0) + \varphi(x), \qquad x \in \mathcal{D}. \tag{6.2'}$$

(6.2')

(iii) There exists a $B \in L(X, Y)$ such that

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - B(x - x_0)\|}{\|x - x_0\|} = 0. \tag{6.3'}$$

If f is differentiable in x_0, then $f'(x_0) = A = B$.

Proof. "(i) $\implies$ (ii)" Let $\varphi : \mathcal{D} \to Y$, $\varphi(x) = (\Phi(x) - \Phi(x_0))(x - x_0)$ and $A := f'(x_0) = \Phi(x_0)$. Then φ is continuous in x_0 and

$$\lim_{x \to x_0} \frac{\|\varphi(x)\|}{\|x - x_0\|} = \lim_{x \to x_0} \|\Phi(x) - \Phi(x_0)\| = 0$$

because Φ is continuous in x_0. Moreover, by definition of φ,

$$f(x) - f(x_0) = \Phi(x)(x - x_0) = \Phi(x_0)(x - x_0) + \varphi(x),$$

so φ satisfies (6.2').

"(ii) $\Longrightarrow$ (i)" Let $x \in \mathcal{D}$. By the Hahn-Banach theorem[1] there exists a linear functional $\psi_x : X \to \mathbb{F}$ such that $\psi_x(\frac{x-x_0}{\|x-x_0\|}) = 1$ and $\|\psi_x\| = 1$. Let $\Phi : \mathcal{D} \to L(X,Y)$ be defined by

$$\Phi(x) : X \to Y, \quad \Phi(x)v = \begin{cases} Av + \|x-x_0\|^{-1}\psi_x(v)\,\varphi(x), & x \neq x_0, \\ Av, & x = x_0. \end{cases}$$

Φ is continuous in x_0 because

$$\begin{aligned} \lim_{x\to x_0} \|\Phi(x) - \Phi(x_0)\| &= \lim_{x\to x_0} \sup\{\big\| \|x-x_0\|^{-1}\psi_x(v)\,\varphi(x)\big\| : v \in X,\, \|v\| = 1\} \\ &\leq \lim_{x\to x_0} \sup\{\|x-x_0\|^{-1}\|\psi_x\|\,\|v\|\,\|\varphi(x)\| : v \in X,\, \|v\| = 1\} \\ &= \lim_{x\to x_0} \|x-x_0\|^{-1}\|\varphi(x)\| = 0 \end{aligned}$$

by assumption on φ. Obviously, Φ satisfies (6.1'). Hence f is differentiable in x_0 and $f'(x_0) = \Phi(x_0)$.

"(ii) $\Longleftrightarrow$ (iii)" The equivalence is obvious with $A = B$. □

If X is an inner product space with scalar product $\langle\cdot\,,\cdot\rangle$ such that $\|x\|^2 = \langle x\,,x\rangle$, $x \in X$, then the function Φ in the proof "(ii) $\Longrightarrow$ (i)" is given by

$$\Phi(x) : X \to Y, \quad \Phi(x)v = \begin{cases} Av + \|x-x_0\|^{-2}\langle x-x_0\,,v\rangle\,\varphi(x), & x \neq x_0, \\ Av, & x = x_0. \end{cases}$$

Corollary. *The derivative $f'(x_0)$ is uniquely determined.*

Proof. Assume that there exist $A, B \in L(X,Y)$ and $\varphi_A, \varphi_B : \mathcal{D} \to Y$ such that

$$f(x) - f(x_0) = A(x-x_0) + \varphi_A(x-x_0) = B(x-x_0) + \varphi_B(x-x_0)$$

and $\lim\limits_{x\to x_0} \frac{\|\varphi_A(x)\|}{\|x-x_0\|} = \lim\limits_{x\to x_0} \frac{\|\varphi_B(x)\|}{\|x-x_0\|} = 0$. it follows that

$$0 = \lim_{x\to x_0} \frac{\|\varphi_A(x) - \varphi_B(x)\|}{\|x-x_0\|} = \lim_{x\to x_0} \frac{\|(A-B)(x-x_0)\|}{\|x-x_0\|} = \lim_{x\to x_0} \Big\|(A-B)\frac{(x-x_0)}{\|x-x_0\|}\Big\|.$$

Let $v \in X$ with $\|v\| = 1$. For every $\lambda \in \mathbb{F}$ there exists an $x_\lambda \in X$ such that $\lambda v = x_0 - x_\lambda$. Obviously, $\frac{\|x-x_0\|}{|\lambda|} = 1$ for all $\lambda \neq 0$ and $x_\lambda \to x_0$ for $\lambda \to 0$. Therefore

$$\|(A-B)v\| = \Big\|(A-B)\frac{x_0 - x_\lambda}{\lambda}\Big\| = \lim_{\lambda\to 0}\Big\|(A-B)\frac{x_0 - x_\lambda}{\lambda}\Big\| = 0.$$

This implies that $\|A-B\| = \sup\{\|(A-B)v\| : v \in X,\, \|v\| = 1\} = 0$, therefore $A = B$. □

As for functions defined on a subset of $\mathbb{F}$ we have the following corollary.

Corollary 6.3'. If f is differentiable in x_0 then f is continuous in x_0.

Proof. This follows immediately from (6.2') because

$$\lim_{x\to x_0} \|f(x) - f(x_0)\| = \lim_{x\to x_0} \|\varphi(x) - \varphi(x_0)\| = 0.$$ □

Obviously, product and chain rule hold also for functions between Banach spaces (see Theorem 6.8 and Theorem 6.10).

[1]Let X be a normed space over $\mathbb{F}$, $U \subseteq X$ a subspace of X and $u' : U \to \mathbb{F}$ a bounded linear map. Then there exists a bounded linear extension $u : X \to \mathbb{F}$ of u' such that $\|u\| = \|u'\|$.

6.2 Local behaviour of differentiable functions

In this section we prove theorems about the local behaviour of real valued functions with domain in $\mathbb{R}$. In particular, criteria for maxima and minima of functions in terms of the derivative are given. For the proof, the mean value theorem is used. We start with a special case of the mean value theorem.

Theorem 6.13 (Rolle's theorem). *Let $a < b \in \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ be a continuous function such that f is differentiable in (a, b). If $f(a) = f(b)$, then there exists a $p \in (a, b)$ such that $f'(p) = 0$.*

Proof. If f is constant, the assertion is clear. Now assume that f is not constant. Without restriction we assume that $f(x) < 0$ for at least one $x \in (a, b)$. Then 0 is not the minimum of f. By Theorem 5.30 f attains its minimum, hence there exists a $p \in (a, b)$ such that $f(p) = \min\{f(x) : x \in \mathcal{D}\}$. Since f is differentiable in p there exists a $\Phi : [a, b] \to \mathbb{R}$ that is continuous in p and that satisfies

$$f(x) - f(p) = \Phi(x)(x - p), \qquad \Phi(p) = f'(p).$$

Since $f(x) - f(p) \geq 0$ for all $x \in \mathcal{D}$ by definition of p, it follows that

$$\Phi(x) = \frac{f(x) - f(p)}{x - p} \begin{cases} < 0, & \text{for } x > p, \\ > 0, & \text{for } x < p. \end{cases}$$

This implies that $f'(p) = \Phi(p) = 0$ because the continuity of Φ in x_0 yields

$$0 \leq \lim_{x \searrow p} \Phi(x) = \Phi(p) = \lim_{x \nearrow p} \Phi(x) \leq 0.$$ □

Theorem 6.14 (Mean value theorem). *Let $a < b \in \mathbb{R}$, $f : [a, b] \to \mathbb{R}$ continuous and differentiable in (a, b). Then there exists a $p \in (a, b)$ such that*

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

Proof. The function

$$h : [a, b] \to \mathbb{R}, \qquad h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a)$$

is continuous in $[a, b]$, differentiable in (a, b) and $h(a) = h(b) = 0$. By Rolle's theorem (Theorem 6.13) there exists a $p \in (a, b)$ such that

$$0 = h'(p) = f'(p) - \frac{f(b) - f(a)}{b - a}.$$ □

Theorem 6.15. *Let $f : (a, b) \to \mathbb{R}$ differentiable. Then*

(i) $f' = 0 \iff f$ *is constant.*

(ii) $f' \geq 0 \iff$ *is monotonically increasing.*
$f' \leq 0 \iff$ *is monotonically decreasing.*

(iii) $f' > 0 \implies f$ *is strictly monotonically increasing.*
$f' < 0 \implies f$ *is strictly monotonically decreasing.*

Note that in (iii) the converse implication is not true: $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$, is strictly increasing but $f'(0) = 0$.

Proof. (i) "$\Longleftarrow$" is clear. To show "$\Longrightarrow$" fix an arbitrary $c \in (a,b)$. By the mean value theorem, for every $q \in (a,b) \setminus \{c\}$ there exists an $p_q \in (a,b)$ such that

$$f(c) - f(q) = f'(p_q)(c-q).$$

Since $f' = 0$ it follows that $f(q) = f(c)$ for all $q \in (a,b)$.

(ii) We prove only the first equivalence.

"$\Longleftarrow$" Let $a < x < y < b$. Then there exists a $p \in (x,y)$ such that $f(y) - f(x) = f'(p)(y-x)$. Since $f'(p) \geq 0$ and $y - x > 0$ it follows that $f(y) \geq f(x)$.

"$\Longrightarrow$" Since for $a < x < y < b$ every difference quotient $\frac{f(y)-f(x)}{y-x}$ is nonnegative, the same is true for $f'(x) = \lim_{y \searrow x} \frac{f(y)-f(x)}{y-x}$.

(iii) is proved as the analogous statement in (ii).

The assertions about (strictly) decreasing functions are proved similarly. □

Remark. Let $f : (a,b) \to \mathbb{R}$ differentiable and assume that $f'(x_0) > 0$ for some $x_0 \in (a,b)$. Then it follows that there exists an $\delta > 0$ such that $f(r) < f(x_0) < f(s)$ for all $r, s \in B_\delta(x_0)$ with $r < x_0 < s$ because

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) > 0.$$

Hence there exists a $\delta > 0$ such that $\frac{f(x)-f(x_0)}{x-x_0} > 0$ if $|x - x_0| < \delta$ and the assertion follows. Note however, that $f(r) < f(x_0) < f(s)$ for $r < x_0 < s$ in a neighbourhood U of x_0 does *not* imply that f is locally increasing at x_0. A counterexample is

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x + 2x^2(1 + (\sin x^{-1})^2).$$

The function f is everywhere differentiable with

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = 1 > 0$$

and for $x \in \mathbb{R} \setminus \{0\}$

$$\begin{aligned} f'(x) &= 1 + 4x(1 + (\sin x^{-1})^2) + 4 \sin x^{-1} \cos x^{-1} \\ &= 1 + 4x(1 + (\sin x^{-1})^2) + 2 \sin(2x)^{-1} \end{aligned}$$

Since the second term tends to zero for $x \to 0$ and the last term oscillates between -2 and 2, there is not interval J around 0 such that the restriction of f to J is either strictly positive or strictly negative. Therefore, by Theorem 6.15, f is not strictly monotonic at 0. Note, however, that $f(x) < f(0) < f(y)$ for $x < 0 < y$ in a neighbourhood of 0. (See also Exercise 6.9.)

Definition 6.16. Let (X, d_X) be a metric space, $p \in \mathcal{D} \subseteq X$ and $f : \mathcal{D} \to \mathbb{R}$. Then $f(p)$ is a *local maximum* of f if

$$\exists\, \delta > 0 \quad \forall x \in \mathcal{D} \cap B_\delta(p) \quad f(x) \leq f(p). \tag{6.7}$$

$f(p)$ is a *global maximum* of f if

$$\forall x \in \mathcal{D} \quad f(x) \leq f(p). \tag{6.8}$$

If in (6.7) or (6.8) strict inequality holds for $x \neq p$, then the maximum is called *isolated*. The value $f(p)$ is *local* or *global minimum* of f if it is a local or global maximum of $-f$. $f(p)$ is called a

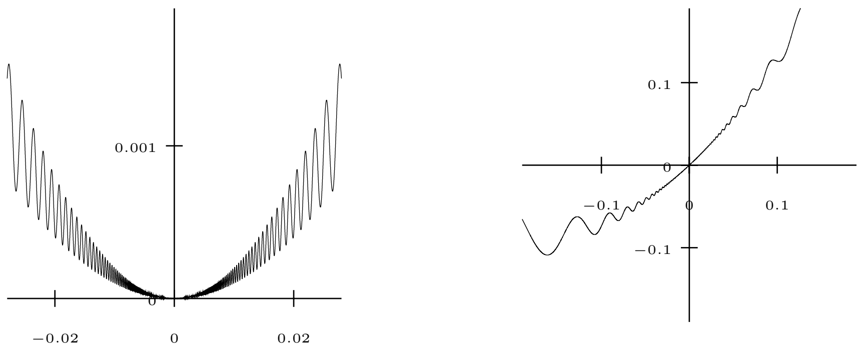


FIGURE 6.2: The function $f(x) = x^2(1 + (\sin x^{-1})^2)$ in the left picture has a global isolated minimum at 0 but there is no right neighbourhood J of f such that $f|_J$ is monotonically increasing.
The function $g(x) = x + 2x^2(1 + (\sin x^{-1})^2)$ in the right picture has derivative $g'(0) = 1$ but it is not monotonic locally at 0.

local extremum if it is a local minimum or maximum, it is called a *global extremum* if it is a global minimum or maximum.

If $X = \mathbb{R}$, then we say that f is *locally increasing at* $p \in \mathcal{D}$ if there exists an $\delta > 0$ such that the restriction of f to $B_\delta(p) \cap \mathcal{D}$ is increasing. The notions *strictly locally increasing* and (*strictly*) *decreasing* are defined analogously.

If a function is arbitrarily often differentiable and not all derivatives in a point p vanish, then it is locally at p either monotonic or it has an isolated local extremum as Theorem 6.18 shows.

Lemma 6.17. *Let $(a, b) \subseteq \mathbb{R}$ and $p \in (a, b)$. Let $f : (a, b) \to \mathbb{R}$ differentiable and assume that $f'(p) = 0$.*

(i) *If there exists a $\delta > 0$ such that*

$$f'(x)(x - p) > 0, \quad x \in (p - \delta, p + \delta) \setminus \{p\},$$

then f has an isolated local minimum at p. In particular this is the case when f' is strictly increasing locally at p.

(ii) *If f' has an isolated local minimum at p, then f is strictly increasing locally at p.*

Proof. (i) By assumption $f'(x) > 0$ for $x \in (p, p + \delta)$ and $f'(x) < 0$ for $x \in (p - \delta, p)$. Therefore f is strictly increasing in $(p, p + \delta)$ and strictly decreasing in $(p - \delta, p)$ which implies $f(x) > f(p)$ for all $x \in (p - \delta, p + \delta) \setminus \{p\}$.

(ii) By assumption, there exists a $\delta > 0$ such that $f'(x) > 0$ for all $x \in (p - \delta, p + \delta) \setminus \{p\}$, hence f is strictly increasing in $(p - \delta, p + \delta)$. □

Theorem 6.18. *Let $(a, b) \subseteq \mathbb{R}$, $p \in (a, b)$ and $n \in \mathbb{N}$, $n \geq 2$. If $f : (a, b) \to \mathbb{R}$ is $(n - 1)$-times differentiable and n-times differentiable in p and*

$$f^{[n]}(p) \neq 0, \qquad f^{[k]}(p) = 0, \quad k = 0, \ldots, n - 1,$$

then exactly one of the following statements holds:

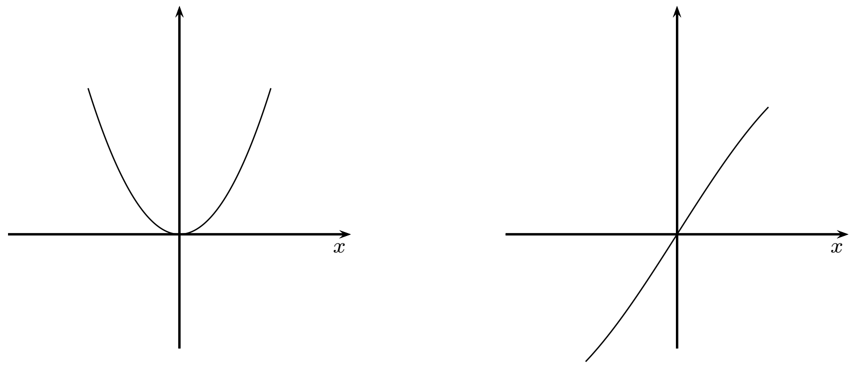


FIGURE 6.3: If f looks like the function on the left, then its derivative looks like the function on the right and vice versa.

$$
\begin{array}{lll}
n \text{ even}, & f^{[n]}(p) > 0 & \implies f \text{ has an isolated local minimum at } p \\
n \text{ even} & f^{[n]}(p) < 0 & \implies f \text{ has an isolated local maximum at } p \\
n \text{ odd} & f^{[n]}(p) > 0 & \implies f \text{ is strictly increasing locally at } p \\
n \text{ odd} & f^{[n]}(p) < 0 & \implies f \text{ is strictly decreasing locally at } p.
\end{array}
$$

Proof. We show only the case when $f^{[n]}(p) > 0$. By assumption, $f^{[n-2]''}(p) > 0$, therefore $f^{[n-2]'}$ is strictly increasing locally at p. Lemma 6.17 (i) implies that $f^{[n-2]}(p)$ has an isolated local minimum at p. By Lemma 6.17 (ii) it follows that $f^{[n-3]}$ is strictly increasing locally at p. Inductively we obtain: $f^{[n-2k]}$ has a an isolated local minimum at p and $f^{[n-2k-1]}$ is strictly increasing locally at p. Depending on whether n is even or odd, $f = f^{[0]}$ has an isolated local minimum at p or it is strictly increasing locally at p. □

The theorem implies that locally at p the function f behaves like the function

$$x \mapsto f^{[n]}(p)(x-p)^n.$$

This will be discussed in more detail in the section about Taylor expansion in Chapter 7.
When all derivatives of f at a point p vanish f does not necessarily behave as described in the theorem above. An example is the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2 \sin(x^{-1}) \text{ for } x \neq 0, \ f(0) = 0.$$

Corollary 6.19. *Let $f : (a, b) \to \mathbb{R}$ and $p \in (a, b)$ such that f is differentiable in p and $f'(p) = 0$.*

(i) $f''(p) > 0 \implies f$ *has an isolated local minimum in p.*

(ii) $f''(p) < 0 \implies f$ *has an isolated local maximum in p.*

Proof. The proof is analogously to the proof of Rolle's theorem. Without restriction we assume that f has a local minimum at p. By assumption the function $\mathcal{D} \to \mathbb{R}$, $x \mapsto \frac{f(x)-f(p)}{x-p}$ is continuous in p with value $f'(p)$. Therefore the claim follows from

$$0 \leq \lim_{x \searrow p} \frac{f(x) - f(p)}{x - p} = f'(p) = \lim_{x \nearrow p} \frac{f(x) - f(p)}{x - p} \leq 0. \qquad \square$$

Definition 6.20. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $(Y, \|\cdot\|)$ a normed space over $\mathbb{F}$ and let $f : \mathbb{F} \supset \mathcal{D} \to Y$ be differentiable in a point $p \in \mathcal{D}$. The p is called a *critical point* of f if $f'(p) = 0$.

All candidates for local extrema of a function $f : [a, b] \to \mathbb{R}$ are:

- the critical points of f,
- points where f is not differentiable,
- the end points of the interval where f is defined.

Theorem 6.21 (Inverse function theorem). *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $x_0 \in \mathcal{D} \subseteq \mathbb{F}$ such that x_0 is a limit point of $\mathcal{D}$. Assume that $f : \mathcal{D} \to \mathbb{F}$ is injective and differentiable in x_0. Moreover, assume that f^{-1} is continuous in $y_0 := f(x_0)$. Then*

$$f^{-1} \text{ differentiable in } y_0 \qquad \iff \qquad f'(x_0) \neq 0.$$

In this case

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}. \tag{6.9}$$

Proof. "$\Longrightarrow$" If f^{-1} is differentiable in y_0, then chain rule yields

$$1 = \frac{\mathrm{d}}{\mathrm{d}x}(f^{-1} \circ f)(x_0) = (f^{-1})'(f(x_0))\, f'(x_0) = (f^{-1})'(y_0)\, f'(f^{-1}(y_0)).$$

In particular, $f'(x_0) \neq 0$ and formula (6.9) holds.
"$\Longleftarrow$" First we show that $y_0 = f(x_0)$ is a limit point of $\mathcal{D}_{f^{-1}} = \mathrm{R}(f)$. Since x_0 is a limit point of $\mathcal{D}$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \setminus \{x_0\}$ that converges to x_0. The injectivity and continuity of f imply that $(f(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}_{f^{-1}} \setminus \{y_0\}$ and that it converges to y_0. Let Φ as in the definition of continuity of f, i. e., Φ is continuous in x_0 and

$$f(x) - f(x_0) = \Phi(x)(x - x_0), \qquad \Phi(x_0) = f'(x_0) \neq 0. \tag{6.10}$$

Since f is injective, $\Phi(x) \neq 0$ for all $x \in \mathcal{D}$ and we obtain from (6.10) (with $f(x) = y$)

$$\underbrace{f^{-1}(y) - f^{-1}(y_0)}_{=x-x_0} = \frac{1}{\Phi(f^{-1}(y))}\,(y - y_0).$$

Since f^{-1} is continuous in y_0 and Φ is continuous in $x_0 = f^{-1}(y_0)$, the assertion is proved. □

Example 6.22. The derivative of $\ln : (0, \infty) \to \mathbb{R}$ defined in Definition 5.52 is

$$\ln'(x) = \frac{1}{x}, \qquad x > 0.$$

Proof. Since the logarithm is the inverse of the real exponential function and $\exp'(x) \neq 0$ for all $x \in \mathbb{R}$, the theorem of the inverse function(Theorem 6.21) yields

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$ □

Example 6.23 (Inverse functions of trigonometric functions). By Exercise 5.1 the functions sin and cos are differentiable on $\mathbb{R}$ and the tangent $\tan := \frac{\sin}{\cos}$ is differentiable on $\mathbb{R}\setminus\{(k+\frac{1}{2})\pi : k \in \mathbb{Z}\}$ with derivatives

$$\sin' = \cos, \qquad \cos' = -\sin, \qquad \tan' = \frac{1}{\cos^2} = 1 + \tan^2.$$

Hence the restrictions

$$\sin : [-\pi/2, \pi/2] \to \mathbb{R}, \qquad \cos : [0, \pi] \to \mathbb{R}, \qquad \tan : [-\pi/2, \pi/2] \to \mathbb{R}$$

are strictly monotonic (see definition of π in Exercise 5.11) and therefore invertible with inverse functions

$$\arcsin : [-1, 1] \to \mathbb{R}, \qquad \arccos : [-1, 1] \to \mathbb{R}, \qquad \arctan : \mathbb{R} \to \mathbb{R}.$$

Their derivatives are

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}, \qquad \arccos'(x) = -\frac{1}{\sqrt{1-x^2}}, \qquad x \in (-1, 1),$$
$$\arctan'(x) = \frac{1}{1+x^2}, \qquad x \in \mathbb{R}.$$

Note that arcsin and arccos are not differentiable in ± 1.

Theorem 6.24 (Generalised mean value theorem). *Let $f, g : [a, b] \to \mathbb{R}$ continuous and differentiable in (a, b). Then there exists a $p \in (a, b)$ such that*

$$\big(f(b) - f(a)\big) g'(p) = \big(g(b) - g(a)\big) f'(p).$$

If $g'(x) \neq 0$, $x \in \mathcal{D}$, then $g(a) \neq g(b)$ and

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

Proof. Let

$$h(x) = \big[f(x) - f(a)\big]\big[g(b) - g(a)\big] - \big[g(x) - g(a)\big]\big[f(b) - f(a)\big].$$

Then h is differentiable in (a, b) and $h(a) = h(b) = 0$. Therefore, by Rolle's theorem, there exists an $p \in (a, b)$ such that

$$0 = h'(p) = f'(p)\big[g(b) - g(a)\big] - g'(p)\big[f(b) - f(a)\big].$$

Note that $g(a) \neq g(b)$ because otherwise, by Rolle's theorem, there would exist a $p \in (a, b)$ such that $g'(p) = 0$. □

Theorem 6.14 follows from Theorem 6.24 for the special case $g = \mathrm{id}$.

Theorem 6.25 (l'Hospital's rules). *Let $-\infty \leq a < b \leq \infty$ and $f, g : (a, b) \to \mathbb{R}$ differentiable functions such that $g'(x) \neq 0$ for all $x \in (a, b)$. Assume that one of the conditions holds:*

(i) *$f(x) \to 0$, $g(x) \to 0$ for $x \searrow a$,*

(ii) *$g(x) \to \infty$ for $x \searrow a$,*

then the existence of $\displaystyle\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ implies the existence of $\displaystyle\lim_{x \searrow a} \frac{f(x)}{g(x)}$ and

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \lim_{x \searrow a} \frac{f(x)}{g(x)}.$$

Analogous statements hold for $x \nearrow b$.

Proof. (i) If $a \neq -\infty$ then f and g can be extended continuously to $[a, b)$ by setting $f(a) = g(a) = 0$. Since $g' \neq 0$, g is either strictly increasing or decreasing, hence $g(x) \neq 0$ for all $x \in (a, b)$ (Darboux's theorem, see Exercise 6.6). The generalised mean mean value theorem (Theorem 6.14) implies that for every $x > 0$ there exists an $p_x \in (a, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(p_x)}{g'(p_x)}.$$

For an arbitrary sequence $x_n \searrow a$, $n \to \infty$ in (a, b) it follows that $p_{x_n} \to a$, hence the statement is proved.
Now let $a = -\infty$. Without restriction we assume $b < 0$. Then the functions $(b^{-1}, 0) \to \mathbb{R}$, $t \mapsto f(t^{-1})$, $t \mapsto g(t^{-1})$. satisfy assumption (i) for $x \nearrow b$ (with $b = 0$). Therefore

$$\lim_{x\to-\infty} \frac{f(x)}{g(x)} = \lim_{t\nearrow 0} \frac{f(t^{-1})}{g(t^{-1})} = \lim_{t\nearrow 0} \frac{\frac{\mathrm{d}}{\mathrm{d}t} f(t^{-1})}{\frac{\mathrm{d}}{\mathrm{d}t} g(t^{-1})} = \lim_{t\nearrow 0} \frac{f'(t^{-1}) \frac{\mathrm{d}}{\mathrm{d}t} t^{-1}}{g'(t^{-1}) \frac{\mathrm{d}}{\mathrm{d}t} t^{-1}} = \lim_{x\to-\infty} \frac{f'(x)}{g'(x)}.$$

(ii) Again, we first consider the case $a \neq -\infty$. Without loss of generality we can assume $g > 0$ and $g' < 0$ in (a, b), the latter again by Darboux's theorem (Exercise 6.6). Let $C = \lim_{x\searrow a} \frac{f'(x)}{g'(x)}$ and fix $\varepsilon > 0$. Then there exists $\delta' > 0$ such that $a + \delta' \leq b$ and

$$C - \varepsilon < \frac{f'(x)}{g'(x)} < C + \varepsilon, \qquad x \in (a, a + \delta').$$

The generalised mean value theorem (Theorem 6.14) implies for $a < x < p < a + \delta'$

$$C - \varepsilon < \frac{f(x) - f(p)}{g(x) - g(p)} < C + \varepsilon.$$

A little bit of algebra shows

$$C - \varepsilon + \frac{\big(f(p) - g(p)\big)\big(C - \varepsilon\big)}{g(x)} \;<\; \frac{f(x)}{g(x)} \;<\; C + \varepsilon + \frac{\big(f(p) - g(p)\big)\big(C - \varepsilon\big)}{g(x)}.$$

Since $g(x) \to \infty$ for $x \searrow a$, there exists a $\delta_0 > 0$ such that

$$C - 2\varepsilon \;<\; \frac{f(x)}{g(x)} \;<\; C + 2\varepsilon, \qquad x \in (a, a + \delta).$$

The case $a = -\infty$ can be treated similarly. □

Similarly it can be shown that

$$\lim_{x\searrow\alpha} \frac{f(x)}{g(x)} = \infty$$

if $f, g : (\alpha,\, \beta) \to \mathbb{R}$ are differentiable functions with $g'(x) \neq 0$ in (α, β) and $\lim\limits_{x\searrow\alpha} g(x) = \lim\limits_{x\searrow\alpha} \dfrac{f'(x)}{g'(x)} = \infty$ (see Exercise 6.9).

Inequalities

Definition 6.26. Let $I \subseteq \mathbb{R}$ a nonempty real interval. A function $f : I \to \mathbb{R}$ is called *convex* if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y), \qquad x, y \in (a, b),\ \lambda \in [0, 1].$$

A function f is called *concave* if $-f$ is convex.

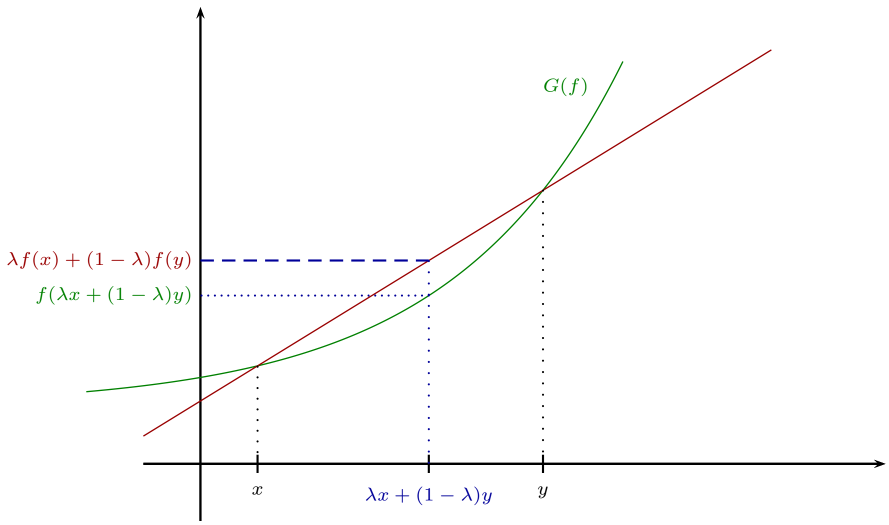


FIGURE 6.4: Convex function: $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$ for all $x<y$ in the domain of f:

Note that $\lambda x+(1-\lambda)y \in [x,y]$ for $\lambda \in [0,1]$.

Theorem 6.27. *Let $I \subseteq \mathbb{R}$ be a nonempty open interval and $f : I \to \mathbb{R}$ twice differentiable. Then*

$$f \text{ convex} \qquad \Longleftrightarrow \qquad f'' \geq 0.$$

Proof. "$\Longleftarrow$" Note that f' is monotonically increasing on I because $f'' \geq 0$. Let $x, y \in I$, without restriction $x<y$. Then for all $\lambda \in (0,1)$ it follows that

$$p := \lambda x+(1-\lambda)y \in (x,y).$$

By the mean value theorem (Theorem 6.14) there exist $p_x \in (x,p)$ and $p_y \in (p,y)$ such that

$$\frac{f(p)-f(x)}{p-x} = f'(p_x) \leq f'(p_y) = \frac{f(y)-f(p)}{y-p}. \tag{6.11}$$

Inequality (6.11) yields

$$(y-x)f(p) \leq f(x)(y-p)+f(y)(p-x) = \lambda f(x)(y-x)+(1-\lambda)f(y)(y-x)$$

since

$$\begin{aligned} y-p &= y-\lambda x-(1-\lambda)y = \lambda(y-x), \\ p-x &= \lambda x+(1-\lambda)y-x = (1-\lambda)(y-x). \end{aligned}$$

"$\Longrightarrow$" Now assume that f is convex. We will show that f' is monotonically increasing. Let $x, y \in I$, without restriction $x<y$. For $\lambda \in (0,1)$ let p be defined as above. Since f is convex it follows that

$$0 \leq \lambda f(x)+(1-\lambda)f(y)-f(p).$$

Multiplication by $y-x$ gives

$$\begin{aligned} 0 &\leq \underbrace{\lambda(y-x)}_{=y-p} f(x) + \underbrace{(1-\lambda)(y-x)}_{=p-x} f(y) - \underbrace{(y-x)}_{y-p+p-x} f(p) \\ &= (y-p)[f(x)-f(p)] + (p-x)[f(y)-f(p)]. \end{aligned}$$

Hence, for all $p \in (x, y)$:

$$\frac{f(p) - f(x)}{p - x} \le \frac{f(y) - f(p)}{y - p}.$$

Since f is differentiable in I it is in particular continuous in I and it follows that

$$\begin{aligned} f'(x) &= \lim_{p \searrow x} \frac{f(p) - f(x)}{p - x} \le \lim_{p \searrow x} \frac{f(y) - f(p)}{y - p} = \frac{f(y) - f(x)}{y - x} \\ &= \lim_{p \nearrow y} \frac{f(p) - f(x)}{p - x} \le \lim_{p \nearrow y} \frac{f(y) - f(p)}{y - p} = f'(y). \end{aligned}$$

□

Examples. $\exp : \mathbb{R} \to \mathbb{R}$ is convex since $\exp'' = \exp > 0$.
$\ln : \mathbb{R}_+ \to \mathbb{R}$ is concave since $\ln''(x) = -x^{-2} < 0$, $x \in \mathbb{R}_+$.

Theorem 6.28 (Young's inequality). *Let $p, q \in (1, \infty)$ such that*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then for all $a, b \ge 0$:

$$ab \le \frac{1}{p}\, a^p + \frac{1}{q}\, b^q. \tag{6.12}$$

Proof. If $ab = 0$, then inequality (6.12) is clear. Now assume $ab > 0$. Since the logarithm is concave and $\frac{1}{p} + \frac{1}{q} = 1$ is follows that

$$\ln\Big(\frac{1}{p}\, a^p + \frac{1}{q}\, b^q\Big) \ge \frac{1}{p} \ln(a^p) + \frac{1}{q} \ln(b^q) = \ln(a) + \ln(b) = \ln(ab).$$

Since $\exp : \mathbb{R} \to \mathbb{R}$ is monotonically increasing, application of exp on both sides of the above inequality proves (6.12). □

Theorem 6.29 (Hölder's inequality). *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $p, q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. For $x = (x_j)_{j=1}^n$ let*

$$\|x\|_p := \Big(\sum_{j=1}^n |x_j|^p\Big)^{\frac{1}{p}}. \tag{6.13}$$

Then for all $x = (x_j)_{j=1}^n$, $y = (y_j)_{j=1}^n \in \mathbb{F}^n$ the following inequality holds:

$$\sum_{j=1}^n |x_j y_j| \;\le\; \|x\|_p \cdot \|y\|_q.$$

Proof. If $x = 0$ or $y = 0$ then the inequality (6.13) clearly holds.
Now assume $x, y \ne 0$. The Young inequality (6.12) with

$$a = \frac{|x_j|}{\|x\|_p}, \quad b = \frac{|y_j|}{\|y\|_q}$$

yields

$$\frac{|x_j|\,|y_j|}{\|x\|_p\,\|y\|_q} \;\le\; \frac{1}{p}\,\frac{|x_j|^p}{\|x\|_p^p} + \frac{1}{q}\,\frac{|y_j|^q}{\|y\|_q^q}\,.$$

Taking the sum over $j = 1, \dots, n$ gives

$$\frac{1}{\|x\|_p \, \|y\|_q} \sum_{j=1}^n |x_j y_j| \;\le\; \underbrace{\frac{1}{p} \frac{1}{\|x\|_p^p} \underbrace{\sum_{j=1}^n |x_j|^p}_{=\|x\|_p^p}}_{=1} + \underbrace{\frac{1}{q} \frac{1}{\|y\|_q^q} \underbrace{\sum_{j=1}^n |y_j|^q}_{=\|y\|_q^q}}_{=1} \;=\; \frac{1}{p} + \frac{1}{q} \;=\; 1. \qquad \square$$

In the special case $p = q = 2$ we obtain the Cauchy-Schwarz inequality.

Corollary 6.30 (Cauchy-Schwarz inequality). *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $n \in \mathbb{N}$. For $x = (x_j)_{j=1}^n$, $y = (y_j)_{j=1}^n \in \mathbb{F}^n$ let*

$$\langle x \,, y \rangle := \sum_{j=1}^n x_j \overline{y_j}.$$

be the Euclidean inner product *on $\mathbb{F}$. Then*

$$|\langle x \,, y \rangle| \le \|x\|_2 \, \|y\|_2.$$

Theorem 6.31 (Minkowski inequality). *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $p \in (1, \infty)$. For all $x, y \in \mathbb{F}^n$ it follows that*

$$\|x + y\|_p \;\le\; \|x\|_p + \|y\|_p. \tag{6.14}$$

Proof. If $x + y = 0$ then (6.14) clearly holds.
Now assume $x + y \neq 0$. Let $q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. The Hölder inequality (6.13) yield

$$\begin{aligned}
\|x + y\|_p^p \;&=\; \sum_{j=1}^n |x_j + y_j| \cdot |x_j + y_j|^{p-1} \\
&\le\; \sum_{j=1}^n |x_j| \, |\underbrace{x_j + y_j}_{:=\widetilde{y}_j}|^{p-1} + \sum_{j=1}^n |y_j| \, |x_j + y_j|^{p-1} \\
&\le\; \|x\|_p \underbrace{\Big(\sum_{j=1}^n |x_j + y_j|^{\overbrace{(p-1)q}^{=p}} \Big)^{\frac{1}{q}}}_{\|\widetilde{y}\|_q} + \|y\|_p \Big(\sum_{j=1}^n |x_j + y_j|^{\overbrace{(p-1)q}^{p}} \Big)^{\frac{1}{q}} \\
&=\; \big(\|x\|_p + \|y\|_p \big) \, \|x + y\|_p^{\frac{p}{q}}.
\end{aligned}$$

Since $p - \frac{p}{q} = p\big(1 - \frac{1}{q}\big) = 1$ division by $\|x + y\|_p^{\frac{p}{q}}$ proves (6.14). $\square$

Note that the Minkowski inequality is the triangle inequality for $\| \cdot \|_p$:

Corollary 6.32. $(\mathbb{F}6n, \| \cdot \|_p)$ *is normed space for $p \in (1, \infty)$.*

6.3 The Riemann-Stieltjes integral in $\mathbb{R}$

A motivation for integration is to determine the area under the graph of a nonnegative function defined on an interval $(a, b) \subseteq \mathbb{R}$.

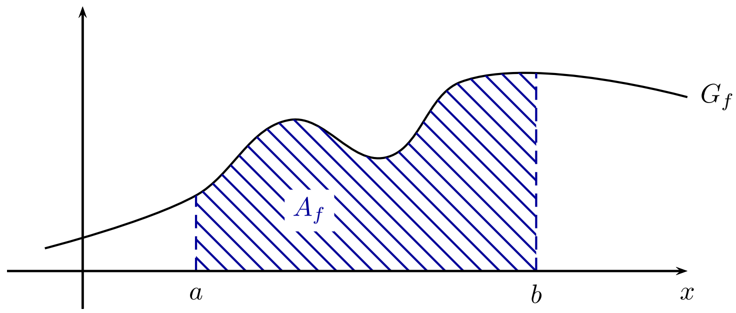


FIGURE 6.5: Geometric interpretation of the Riemann integral.

In the special case that f is piecewise constant, the area is $A_f = \sum_{j=1}^n f_j(c_j - c_{j-1})$ if $f(x) = c_j$ for $x \in (c_j - c_{j-1})$. In the general case, the integral will be defined as the limit of integrals of piecewise functions that approximate f in a suitable sense.

In this section we always assume that $-\infty < a < b < \infty$.

Definition 6.33. A *partition of* $[a, b]$ is a finite set of points $P := \{x_0, \dots, x_n\}$ such that

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

If P, P' are partitions of $[a, b]$ and $P \subseteq P'$, then P' is called a *refinement* of P.

Obviously, if P, Q are partitions of $[a, b]$, then $P \cup Q$ is a common refinement of both P and Q.

In the following we will always assume that $\alpha : [a, b] \to \mathbb{R}$ is an increasing function. In particular, α is bounded because

$$-\infty < \alpha(a) \le \alpha(x) \le \alpha(b) < \infty, \qquad x \in [a, b].$$

Definition 6.34. Let $[a, b] \subseteq \mathbb{R}$, $f : [a, b] \to \mathbb{R}$ a bounded function. Given a partition $P = \{x_0, x_1, \dots, x_n\} \subseteq [a, b]$ we define for $j = 1, \dots, n$:

$$\begin{aligned} \Delta\alpha_j &:= \alpha(x_j) - \alpha(x_{j-1}), \\ m_j &:= \inf\{f(x) : x \in [x_{j-1}, x_j]\}, \\ M_j &:= \sup\{f(x) : x \in [x_{j-1}, x_j]\}, \end{aligned}$$

we define the sums

$$s(f, \alpha, P) := \sum_{j=1}^n m_j \Delta\alpha_j, \qquad S(f, \alpha, P) := \sum_{j=1}^n M_j \Delta\alpha_j$$

and the numbers

$$\begin{aligned} \int_{*\,a}^{b} f \,\mathrm{d}\alpha &:= \sup\{s(f, \alpha, P) : P \text{ partition of } [a, b]\}, \\ \int_{a}^{*\,b} f \,\mathrm{d}\alpha &:= \inf\{S(f, \alpha, P) : P \text{ partition of } [a, b]\}. \end{aligned}$$

Note that

$$\sum_{j=1}^{n} \Delta\alpha_j = \alpha(b) - \alpha(a).$$

Remark 6.35. Let $m, M \in \mathbb{R}$ such that $m \le f \le M$. Then for a fixed partition P of $[a,b]$:

$$m(\alpha(b) - \alpha(a)) \le s(f,\alpha,P) \le S(f,\alpha,P) \le M(\alpha(b) - \alpha(a)),$$

hence

$$\int_{*\,a}^{b} f \,\mathrm{d}\alpha \ge m(\alpha(b) - \alpha(a)) > -\infty,$$
$$\int_{a}^{*\,b} f \,\mathrm{d}\alpha \le M(\alpha(b) - \alpha(a)) < \infty.$$

Lemma 6.36. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function.*

(i) *Let P, P' be partitions of $[a,b]$ such that $P \subseteq P'$. Then*

$$s(f,\alpha,P) \le s(f,\alpha,P') \le S(f,\alpha,P') \le S(f,\alpha,P).$$

(ii) $\displaystyle\int_{*\,a}^{b} f \,\mathrm{d}\alpha \le \int_{a}^{*\,b} f \,\mathrm{d}\alpha.$

Proof. (i) The middle estimate follows from Remark 6.35. Let us show the first estimate. The last estimate is proved analogously.
Let $P = \{x_0, x_1, \dots, x_n\}$. If $P = P'$ then the estimate is clear. Now assume $P \ne P'$. It suffices to show the estimate in the case when $P \setminus P' = \{y\}$, for the case $P \setminus P' = \{y_1, \dots, y_n\}$ follows then by induction. Let $k \in \{1, \dots, n\}$ such that $x_{k-1} < y < x_k$. Then

$$m_k^- := \inf\{f(x) \,:\, x \in [x_{k-1}, y]\} \ge m_k,$$
$$m_k^+ := \inf\{f(x) \,:\, x \in [y, x_k]\} \ge m_k$$

and it follows that

$$\begin{aligned}
s(f,\alpha,P') - s(f,\alpha,P)& \\
&= m_k^-(\alpha(y) - \alpha(x_{k-1})) + m_k^+(\alpha(x_k) - \alpha(y)) - m_k(\alpha(x_k) - \alpha(x_{k-1})) \\
&= \underbrace{(m_k^- - m_k)}_{\ge 0}\underbrace{(\alpha(y) - \alpha(x_{k-1}))}_{\ge 0} + \underbrace{(m_k^+ - m_k)}_{\ge 0}\underbrace{(\alpha(x_k) - \alpha(y))}_{\ge 0} \ge 0.
\end{aligned}$$

(ii) For partitions P_1, P_2 it follows by (i) that

$$s(f,\alpha,P_1) \le s(f,\alpha,P_1 \cup P_2) \le S(f,\alpha,P_1 \cup P_2) \le S(f,\alpha,P_2).$$

Taking the supremum over all partitions P_1 on the left hand side and the infimum over all partitions P_2 on the right side proves the assertion. □

Definition 6.37. A bounded function $f : [a,b] \to \mathbb{R}$ is called *Riemann-Stieltjes integrable* (or simply *integrable*) with respect to α if

$$\int_{*\,a}^{b} f \,\mathrm{d}\alpha = \int_{a}^{*\,b} f \,\mathrm{d}\alpha.$$

In this case

$$\int_a^b f\,\mathrm{d}\alpha := \int_{*\,a}^{b} f\,\mathrm{d}\alpha = \int_a^{*\,b} f\,\mathrm{d}\alpha$$

is called the *Riemann-Stieltjes integral* of f.

Remark 6.38. In the case when $\alpha = \mathrm{id}$, the integral is called the *Riemann-integral.* For positive functions f the integral of f is the area between the graph of f and the x-axis. The following notation is used:

$$\int_a^b f\,\mathrm{d}\alpha =: \int_a^b f(x)\,\mathrm{d}\alpha(x),$$
$$\int_a^b f\,\mathrm{d}\alpha =: \int_a^b f\,\mathrm{d}x =: \int_a^b f(x)\,\mathrm{d}x \qquad \text{if } \alpha = \mathrm{id}\,.$$

Notation.

$$\mathcal{R}(\alpha) := \{f : [a,b] \to \mathbb{R} \,:\, f \text{ is Riemann-Stieltjes integrable with respect to } \alpha\},$$
$$\mathcal{R} := \mathcal{R}([a,b]) := \{f : [a,b] \to \mathbb{R} \,:\, f \text{ is Riemann integrable}\},$$

Remark 6.39. If α is constant, then obviously every bounded function f is Riemann-Stieltjes integrable with respect to α and $\int_a^b f\,\mathrm{d}\alpha = 0$.

Theorem 6.40 (Riemann criterion). *Let $f : [a,b] \to \mathbb{R}$ a bounded function. Then $f \in \mathcal{R}(\alpha)$ if and only if*

$$\forall\,\varepsilon > 0 \quad \exists\,P_\varepsilon \text{ partition of } [a,b] : \quad S(f,\alpha,P_\varepsilon) - s(f,\alpha,P_\varepsilon) < \varepsilon.$$

Proof. "$\Longleftarrow$" Let $\varepsilon > 0$ and P_ε as above. By Lemma 6.36 (ii) it follows that

$$0 \le \underbrace{\int_a^{*\,b} f\,\mathrm{d}\alpha}_{\le S(f,\alpha,P_\varepsilon)} - \underbrace{\int_{*\,a}^{b} f\,\mathrm{d}\alpha}_{\ge s(f,\alpha,P_\varepsilon)} \le S(f,\alpha,P) - s(f,\alpha,P) < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the assertion is proved.
"$\Longrightarrow$" Assume that f is Riemann-Stieltjes integrable with respect to α and let $\varepsilon > 0$. By Definition 6.34 there exist partitions P_1, P_2 of $[a,b]$ such that

$$\int_a^b f\,\mathrm{d}\alpha - s(f,\alpha,P_1) < \frac{\varepsilon}{2}, \qquad S(f,\alpha,P_2) - \int_a^b f\,\mathrm{d}\alpha < \frac{\varepsilon}{2}.$$

Addition of the inequalities gives

$$\varepsilon > S(f,\alpha,P_2) - s(f,\alpha,P_1) \ge S(f,\alpha,P_1 \cup P_2) - s(f,\alpha,P_1 \cup P_2) \ge 0. \qquad \square$$

Theorem 6.41. *Every continuous function $f : [a,b] \to \mathbb{R}$ is Riemann-Stieltjes integrable.*

Proof. We use the Riemann criterion to show the integrability of f. Let $\varepsilon > 0$. In the case when α is constant, the assertion follows immediately from Remark 6.39. Now assume that α is not constant. In particular, it follows that $\alpha(a) \neq \alpha(b)$. Since f is continuous on the compact set $[a,b]$, it is uniformly continuous (Theorem 5.33), so there exists an $\delta > 0$ such that

$$\forall x, y \in [a,b] \qquad |x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{\alpha(b) - \alpha(a)}.$$

Now choose n large enough such that

$$\frac{b-a}{n} \leq \delta$$

and define the partition $P = \{x_0, x_1, \dots, x_n\}$ by

$$x_j = a + j\,\frac{b-a}{n}, \qquad j = 0, \dots, n.$$

Then

$$S(f,\alpha,P) - s(f,\alpha,P) = \sum_{j=1}^{n} (M_j - m_j)\Delta\alpha_j < \sum_{j=1}^{n} \frac{\varepsilon}{\alpha(b)-\alpha(a)}\,\frac{\alpha(b)-\alpha(a)}{n} = \varepsilon.$$

By the Riemann criterion (Theorem 6.40) f is integrable. □

Theorem 6.42. *If $f : [a,b] \to \mathbb{R}$ is monotonic and α is increasing and continuous, then $f \in \mathcal{R}(\alpha)$.*

Proof. Let $\varepsilon > 0$ and $n \in \mathbb{N}$ such that

$$n^{-1}(\alpha(b) - \alpha(a))|f(b) - f(a)| < \varepsilon.$$

Since α is continuous there exists a partition $P = \{x_0, x_1, \dots, x_n\}$ of $[a,b]$ such that

$$\Delta\alpha_j = \frac{\alpha(b)-\alpha(a)}{n}, \qquad j = 1, \dots, n.$$

Without restriction we assume that f is increasing. Then

$$f(x_{j-1}) = m_j \leq M_j = f(x_j), \qquad j = 1, \dots, n,$$

and therefore

$$\begin{aligned} S(f,\alpha,P) - s(f,\alpha,P) &= \sum_{j=1}^{n} (M_j - m_j)\Delta\alpha_j \leq \sum_{j=1}^{n} \big(f(x_j) - f(x_{j-1})\big)\frac{\alpha(b)-\alpha(a)}{n} \\ &= n^{-1}(f(b) - f(a))(\alpha(b) - \alpha(a)) < \varepsilon \end{aligned}$$

by the choice of n. Therefore f is integrable by the Riemann criterion (Theorem 6.40). □

Theorem 6.43 (Properties of the Riemann-Stieltjes integral).

(i) *Let $f : [a,b] \to \mathbb{R}$ and let $c \in (a,b)$. Set $f_1 := f|_{[a,c]}$, $f_2 := f|_{[c,b]}$ and $\alpha_1 := \alpha|_{[a,c]}$, $\alpha_2 := \alpha|_{[c,b]}$. Then $f \in \mathcal{R}(\alpha)$ if and only if $f_1 \in \mathcal{R}(\alpha_1)$ on $[a,c]$ and $f_2 \in \mathcal{R}(\alpha_2)$ on $[c,b]$. In this case*

$$\int_a^b f \,\mathrm{d}\alpha = \int_a^c f \,\mathrm{d}\alpha + \int_c^b f \,\mathrm{d}\alpha.$$

Now let $f, g \in \mathcal{R}(\alpha)$ and $\gamma \in \mathbb{R}$.

(ii) *$f + \gamma g \in \mathcal{R}(\alpha)$ and*

$$\int_a^b f + \gamma g \,\mathrm{d}\alpha = \int_a^b f \,\mathrm{d}\alpha + \gamma \int_a^b g \,\mathrm{d}\alpha.$$

(iii) *If* $f \leq g$ *then*

$$\int_a^b f \, \mathrm{d}\alpha \leq \int_a^b g \, \mathrm{d}\alpha.$$

(iv) *If* $f \in \mathcal{R}(\alpha_1)$ *and* $f \in \mathcal{R}(\alpha_2)$ *on* $[a,b]$ *then* $f \in \mathcal{R}(\alpha_1 + \gamma\alpha_2)$ *on* $[a,b]$ *then*

$$\int_a^b f \, \mathrm{d}(\alpha_1 + \gamma\alpha_2) = \int_a^b f \, \mathrm{d}\alpha_1 + \gamma \int_a^b f \, \mathrm{d}\alpha_2.$$

Proof. Exercise. □

Theorem 6.44. *If* $f : [a,b] \to \mathbb{R}$ *is bounded and has only finitely many discontinuities and* α *is continuous at every point where* f *is discontinuous, then* f *is integrable with respect to* α.

Proof. By theorem 6.43 (i) we can write the interval $[a,b]$ as union of smaller intervals each of which contains only one discontinuity of f. We may even assume that the discontinuity of f is at the boundary of the interval. Without restriction we will assume that f is continuous in $(a,b]$ and that α is continuous at a. Let $\varepsilon > 0$ and $M > \sup\{|f(x)| : x \in [a,b]\}$. Since α is continuous in a, there exists $0 < \delta < (b-a)/2$ such that $|\alpha(a) - \alpha(t)| < \frac{\varepsilon}{4M}$ for all $t \in [a, a+2\delta]$. Since f is continuous in $I = [a+\delta, b]$, it is integrable there. So we can choose a partition P of I such that

$$S(f|_I, \alpha|_I, P) - s(f|_I, \alpha|_I, P) < \frac{\varepsilon}{2}.$$

Then $Q := P \cup \{a\}$ is a partition of $[a,b]$ and

$$\begin{aligned} S(f,\alpha,Q) - s(f,\alpha,Q) &= \Big(\sup\{f(t) : t \in [a,a+\delta]\} - \inf\{f(t) : t \in [a,a+\delta]\}\Big)(\alpha(a+\delta) - \alpha(a)) \\ &\quad + S(f|_I, \alpha|_I, P) - s(f|_I, \alpha|_I, P) \\ &< 2M\frac{\varepsilon}{4M} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Hence f is integrable on $[a,b]$ by the Riemann criterion (Theorem 6.40). □

Theorem 6.42 and Theorem 6.44 and show that every function $f : [a,b] \to \mathbb{R}$ that is either monotonic or has only finitely many discontinuities is Riemann integrable.

Theorem 6.45. *Let* $f \in \mathcal{R}(\alpha)$ *and* $m, M \in \mathbb{R}$ *such that* $\mathrm{R}(f) \subseteq [m, M]$. *If* $\Phi : [m,M] \to \mathbb{R}$ *is continuous, Then* $h = \Phi \circ f \in \mathcal{R}(\alpha)$.

Proof. Let $\varepsilon > 0$. Since Φ is uniformly continuous on $[m,M]$ there exists a $\delta \in (0,\varepsilon)$ such that

$$|x - y| < \delta \implies |\Phi(x) - \Phi(y)| < \varepsilon, \qquad x, y \in [a,b] \tag{6.15}$$

Let M_j, m_j for f as in Definition 6.34 and m_j', M_j' the analogon for h.
Since by assumption $f \in \mathcal{R}(\alpha)$, there exists a partition P of $[a,b]$ such that

$$S(f,\alpha,P) - s(f,\alpha,P) \leq \delta^2. \tag{6.16}$$

Let $A := \{j : M_j - m_j < \delta\}$, $B := \{j : M_j - m_j \geq \delta\}$. Then

$$\begin{aligned} j \in A &\implies M_j' - m_j' \leq \varepsilon \qquad \text{by (6.15)} \\ j \in B &\implies M_j' - m_j' \leq 2\|\Phi\|_\infty. \end{aligned}$$

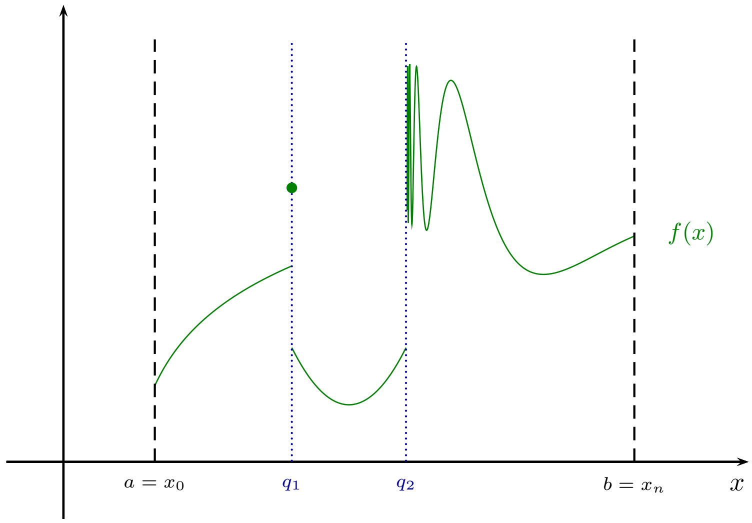


FIGURE 6.6: The function f has only finitely many discontinuities. Use $[a, b] = [a, p_1] \cup \cdots \cup [q_n, b]$ such that f restricted to these subintervals has sigularities only at the boundaries of these subintervals.

From (6.16) it follows that

$$\sum_{j \in B} \Delta\alpha_j \le \sum_{j \in B} \underbrace{\frac{M_j - m_j}{\delta}}_{\le 1} \Delta\alpha_j \le \frac{1}{\delta} \big(S(f, \alpha, P) - s(f, \alpha, P) \big) \overset{(6.16)}{<} \delta.$$

Therefore $h \in \mathbb{R}(\alpha)$ by the Riemann criterion (Theorem 6.40) because

$$\begin{aligned} S(h, \alpha, P) - s(h, \alpha, P) &= \sum_{j \in A} (M_j' - m_j') \Delta\alpha_j + \sum_{j \in B} (M_j' - m_j') \Delta\alpha_j \\ &\le \varepsilon\big(\alpha(b) - \alpha(a)\big) + 2\|\Phi\|_\infty \delta < \varepsilon\big(\alpha(b) - \alpha(a) + 2\|\Phi\|_\infty\big) \end{aligned}$$

and $\varepsilon > 0$ was arbitrary. □

Theorem 6.46. *Let $f \in \mathcal{R}(\alpha)$. Then also $|f| \in \mathcal{R}(\alpha)$ and*

$$\Big| \int_a^b f \, \mathrm{d}\alpha \Big| \le \int_a^b |f| \, \mathrm{d}\alpha.$$

Proof. Since $| \cdot | : \mathbb{R} \to \mathbb{R}$ is continuous, $|f| \in \mathcal{R}(\alpha)$ by Theorem 6.45. Chose $c \in \{\pm 1\}$ such that

$$c \int_a^b f \, \mathrm{d}\alpha \ge 0.$$

By Theorem 6.43 (ii) and (iii) it follows that

$$\Big| \int_a^b f \, \mathrm{d}\alpha \Big| = c \int_a^b f \, \mathrm{d}\alpha = \int_a^b \underbrace{cf}_{\le |f|} \, \mathrm{d}\alpha \le \int_a^b |f| \, \mathrm{d}\alpha.$$ □

Theorem 6.47. *Let $f, g \in \mathcal{R}(\alpha)$. Then also f^2 and $fg \in \mathcal{R}(\alpha)$.*

Proof. Since $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ is continuous, Theorem 6.45 implies that $f^2 \in \mathcal{R}(\alpha)$. In order to see that $fg \in \mathcal{R}(\alpha)$ note that $fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$. □

Theorem 6.48. *Let* $\alpha : [a,b] \to \mathbb{R}$ *be increasing and differentiable with* $\alpha' \in \mathcal{R}$. *If* $f : [a,b] \to \mathbb{R}$ *is bounded, then* $f \in \mathcal{R}(\alpha)$ *if and only if* $f\alpha' \in \mathcal{R}$. *In this case*

$$\int_a^b f \,\mathrm{d}\alpha = \int_a^b f\alpha' \,\mathrm{d}x. \tag{6.17}$$

Proof. Let $\varepsilon > 0$. Since $\alpha' \in \mathcal{R}$, there exists a partition $P_\varepsilon = \{x_0, x_1, \dots, x_n\}$ of $[a,b]$ such that

$$S(\alpha', P_\varepsilon) - s(\alpha', P_\varepsilon) < \varepsilon.$$

By the mean value theorem, for all $j = 1, \dots, n$ there exists $t_j \in [x_{j-1}, x_j]$ such that $\Delta\alpha_j = \alpha'(t_j)\Delta x_j$. For arbitrary $s_j \in [x_{j-1}, x_j]$ we have

$$\begin{aligned}\Big|\sum_{j=1}^n f(s_j)\underbrace{\Delta\alpha_j}_{=\alpha'(t_j)\Delta x_j} - \sum_{j=1}^n f(s_j)\alpha'(s_j)\Delta x_j\Big| &= \Big|\sum_{j=1}^n f(s_j)\big[\alpha'(t_j) - \alpha'(s_j)\big]\Delta x_j\Big| \\ &\le \|f\|_\infty \Big|\sum_{j=1}^n \big[\alpha'(t_j) - \alpha'(s_j)\big]\Delta x_j\Big| \le \|f\|_\infty \Big(S(\alpha', P_\varepsilon) - s(\alpha', P_\varepsilon)\Big) < \varepsilon\|f\|_\infty.\end{aligned}$$

Since the s_j are chosen arbitrarily in $[x_{j-1}, x_j]$, we can chose them such that $0 \le S(f, \alpha, P_\varepsilon) - \sum_{j=1}^n f(s_j)\Delta\alpha_j < \varepsilon$. Then the above inequality implies

$$\begin{aligned} S(f, \alpha, P_\varepsilon) &< \varepsilon + \sum_{j=1}^n f(s_j)\Delta\alpha_j < \varepsilon + \varepsilon\|f\|_\infty + \sum_{j=1}^n f(s_j)\alpha'(s_j)\Delta x_j \\ &\le \varepsilon(1 + \|f\|_\infty) + S(f\alpha', P_\varepsilon). \end{aligned} \tag{6.18}$$

Analogously, if we chose the s_j such that $0 \le S(f\alpha', P_\varepsilon) - \sum_{j=1}^n f(s_j)\alpha'(s_j)\Delta x_j < \varepsilon$, then the above inequality implies

$$S(f\alpha', P_\varepsilon) < \varepsilon(1 + \|f\|_\infty) + S(f, \alpha, P_\varepsilon). \tag{6.19}$$

Inequalities (6.18) and (6.19) imply

$$|S(f\alpha', P_\varepsilon) - S(f, \alpha, P_\varepsilon)| < \varepsilon(1 + \|f\|_\infty). \tag{6.20}$$

Analogously

$$|s(f\alpha', P_\varepsilon) - s(f, \alpha, P_\varepsilon)| < \varepsilon(1 + \|f\|_\infty) \tag{6.21}$$

can be shown. From the inequalities (6.20) and (6.21) if follows that $f \in \mathcal{R}(\alpha)$ if and only if $f\alpha' \in \mathcal{R}$ and in this case, formula (6.17) holds. □

Theorem 6.49 (Change of variables). *Let* $[a,b]$ *and* $[A,B]$ *nonempty intervals in* $\mathbb{R}$ *and* $\varphi : [A,B] \to [a,b]$ *a monotonically increasing bijection. Suppose that* $\alpha : [a,b] \to \mathbb{R}$ *is monotonically increasing and that* $f \in \mathcal{R}(\alpha)$. *Let*

$$\beta := \alpha \circ \varphi : [A,B] \to \mathbb{R}, \qquad g := f \circ \varphi : [A,B] \to \mathbb{R}.$$

Then $g \in \mathcal{R}(\beta)$ *and*

$$\int_A^B g \,\mathrm{d}\beta = \int_a^b f \,\mathrm{d}\alpha.$$

Proof. Since φ is increasing, also β is an increasing function on $[A,B]$. The bijection φ induces a bijection between the partitions of $[A,B]$ and the partitions of $[a,b]$:

$$\varphi^* : \{P : \text{partition of } [A,B]\} \to \{P : \text{partition of } [a,b]\},$$
$$\{x_0, x_1, \dots, x_n\} \mapsto \{\varphi(x_0), \varphi(x_1), \dots, \varphi(x_n)\}.$$

Since $S(g,\beta,P) = S(f\circ\varphi, \alpha\circ\varphi, P) = S(g,\beta,\varphi^*P)$ and analogously $s(g,\alpha,P) = s(g,\beta,\varphi^*P)$ for every partition P of $[A,B]$ it follows that

$$\int_a^{*\,b} f\,\mathrm{d}\alpha = \int_a^{*\,b} g\,\mathrm{d}\beta, \qquad \int_{*\,a}^{b} f\,\mathrm{d}\alpha = \int_{*\,a}^{b} g\,\mathrm{d}\beta. \qquad \square$$

Corollary 6.50. *In the special case when* $\alpha = \mathrm{id}$ *and* $\beta = \varphi$ *is differentiable such that* $\varphi' \in \mathcal{R}$*, we obtain the* transformation formula

$$\int_A^B (f\circ\varphi)(y)\varphi'(y)\,\mathrm{d}y = \int_{\varphi(A)}^{\varphi(B)} f(x)\,\mathrm{d}x.$$

6.4 Riemann integration and differentiation

Theorem 6.51 (Intermediate value theorem of integration). *Let* $f : [a,b] \to \mathbb{R}$ *continuous and* $g : [a,b] \to \mathbb{R}$ *Riemann integrable with* $g \ge 0$*. Then there exists a* $p \in [a,b]$ *such that*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(p)\int_a^b g(x)\,\mathrm{d}x$$

Proof. Since f is continuous on the compact interval $[a,b]$, there exist $m, M \in \mathbb{R}$ such that $\mathrm{R}(f) = [m,M]$ (Theorem 5.24 and Theorem 5.30). It follows that $mg \le fg \le Mg$ because $g \ge 0$. By Theorem 6.43 we obtain

$$m\int_a^b g(x)\,\mathrm{d}x \le \int_a^b f(x)g(x)\,\mathrm{d}x \le M\int_a^b g(x)\,\mathrm{d}x.$$

Hence there exists an $\mu \in [m,M]$ such that

$$\mu\int_a^b g(x)\,\mathrm{d}x = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

By the intermediate value theorem (Theorem 5.24) there exists a $p \in [a,b]$ such that $f(p) = \mu$. $\square$

Notation 6.52. If $f : [a,b] \to \mathbb{R}$ is Riemann integrable, we set

$$\int_a^a f(x)\,\mathrm{d}x := 0, \qquad \int_b^a f(x)\,\mathrm{d}x := -\int_a^b f(x)\,\mathrm{d}x.$$

Theorem 6.53. *Let* $f : [a,b] \to \mathbb{R}$ *Riemann integrable. Let*

$$F_a(x) := \int_a^x f(t)\,\mathrm{d}t, \qquad t \in [a,b].$$

Then F_a *is continuous in* $[a,b]$*. If* f *is continuous in* $x_0 \in [a,b]$*, then* F_a *is differentiable in* x_0 *and*

$$F_a'(x_0) = f(x_0).$$

Proof. Since f is integrable, it is bounded in $[a,b]$. Let $M \geq |f|$ and $x < y \in [a,b]$. Then the continuity of F_a follows from

$$|F_a(x) - F_a(y)| = \left| \int_x^y f(t) \, \mathrm{d}t \right| \leq \int_x^y |f(t)| \, \mathrm{d}t \leq M|y-x|.$$

Now assume that f is continuous in $x_0 \in [a,b]$ and let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that

$$|x_0 - x| < \delta \implies |f(x_0) - f(x)| < \varepsilon, \qquad x \in [a,b].$$

Theorem 6.46 implies for $x \in [x_0, x_0 + \delta] \cap [a,b]$:

$$\begin{aligned} \left| \frac{F_a(x) - F_a(x_0)}{x - x_0} - f(x_0) \right| &= \frac{1}{x - x_0} \left| \int_{x_0}^x f(x) - f(x_0) \, \mathrm{d}x \right| \\ &\leq \frac{1}{x - x_0} \int_{x_0}^x \underbrace{|f(x) - f(x_0)|}_{<\varepsilon} \, \mathrm{d}x < \varepsilon. \end{aligned}$$

Analogously for $x \in [x_0 - \delta, x_0] \cap [a,b]$. □

The proof shows that F_a is even Lipschitz continuous.

Definition 6.54. Let $f : [a,b] \to \mathbb{R}$ Riemann integrable. A differentiable function $F : [a,b] \to \mathbb{R}$ is called an *antiderivative* of f if

$$F'(x) = f(x), \qquad x \in [a,b].$$

In this case we write

$$F(x) = \int f(x) \, \mathrm{d}x.$$

Proposition 6.55. *Let $f : [a,b] \to \mathbb{R}$ Riemann integrable and F an antiderivative of f. Then $G : [a,b] \to \mathbb{R}$ is an antiderivative of f if and only if $F - G = const.$*

Proof. Assume that $F - G \equiv c \in \mathbb{R}$. Then G is differentiable and $G' = F' = f$.
Now assume that G is an antiderivative of f. Then $(F-G)' = f - f = 0$. Therefore $F - G = const.$ by Theorem 6.15. □

Theorem 6.56 (Fundamental theorem of calculus). *Let $f : [a,b] \to \mathbb{R}$ continuous and $F : [a,b] \to \mathbb{R}$ an antiderivative of f. Then*

$$\int_a^b f(t) \, \mathrm{d}t = F(b) - F(a).$$

Proof. Let F_a be the antiderivative of f defined in Theorem 6.53. By Proposition 6.55 there exists a constant c such that $F = F_a - c$, hence

$$F(b) - F(a) = (F(b) - c) - (F(a) - c) = F_a(b) - F_a(a) = \int_a^b f(t) \, \mathrm{d}t.$$ □

Corollary 6.57. *If $F : [a,b] \to \mathbb{R}$ is continuously differentiable, then*

$$F(x) = F(a) + \int_a^x F'(t) \, \mathrm{d}t.$$

The fundamental theorem implies two methods to find the integral of a given function.

Theorem 6.58 (Substitution rule). *Let $f : [a, b] \to \mathbb{R}$ continuous, $\varphi : [A, B] \to [a, b]$ continuously differentiable. Then*

$$\int_A^B (f \circ \varphi)(y)\varphi'(y)\,\mathrm{d}y = \int_a^b f(x)\,\mathrm{d}x.$$

Proof. Let F be an antiderivative of f. Then $F \circ \varphi$ is an antiderivative of $(f \circ \varphi)\varphi'$ because by the chain rule

$$(F \circ \varphi)'(y) = F'(\varphi(y))\varphi'(y), = f(\varphi(y))\varphi'(y), \qquad y \in [A, B].$$

The Fundamental Theorem of Calculus implies

$$\begin{aligned}\int_A^B (f \circ \varphi)(y)\varphi'(y)\ \mathrm{d}y &= (F \circ \varphi)(B) - (F \circ \varphi)(A) = F(\varphi(B)) - F(\varphi(A)) \\ &= \int_{\varphi(A)}^{\varphi(B)} f(x)\ \mathrm{d}x.\end{aligned}$$ □

Corollary. *Let $f : [a, b] \to \mathbb{R}$ continuous and $c \in \mathbb{R}$. Then*

(i) $\displaystyle\int_{A+c}^{B+c} f(x)\ \mathrm{d}x = \int_A^B f(x + c)\ \mathrm{d}x \qquad$ *if* $[A + c, B + c] \subseteq [a, b]$,

(ii) $\displaystyle\int_{cA}^{cB} f(x)\ \mathrm{d}x = \int_A^B f(cx)\ \mathrm{d}x \qquad$ *if* $[cA, cB] \subseteq [a, b]$.

Theorem 6.59 (Integration by parts). *Let $f, g : [a, b] \to \mathbb{R}$ be continuously differentiable. Then*

$$\int_a^b f'(x)g(x)\ \mathrm{d}x + \int_a^b g'(x)f(x)\ \mathrm{d}x = [f(x)g(x)]_a^b\,.$$

Proof. The formula follows immediately from the Fundamental Theorem of Calculus because fg is an antiderivative of $f'g + fg'$. □

Improper integrals

Until now, we considered integrals of bounded functions on bounded and closed intervals. Next we want to extend the integral also to functions that are defined on open or halfopen intervals and possibly unbounded.

Definition 6.60. Let $D \subseteq \mathbb{R}$ be an interval. A function $f : D \to \mathbb{R}$ is called *locally Riemann integrable* if for every compact interval $[\alpha, \beta] \subseteq [a, b]$ the restriction $f|_{[\alpha,\beta]}$ is Riemann integrable. For a locally integrable function f its *improper integral* $\int_D f\,\mathrm{d}x$ of f is defined by

(i) if $D = (a, b]$: $\displaystyle\int_D f\,\mathrm{d}x = \lim_{t \searrow a} \int_t^b f(x)\,\mathrm{d}x \qquad$ if the limit exists,

(ii) if $D = [a, b)$ analogously,

(iii) if $D = (a, b)$: for arbitrary $c \in (a, b)$:

$$\int_D f\,\mathrm{d}x = \lim_{t \searrow a} \int_t^c f(x)\,\mathrm{d}x = \lim_{t \nearrow b} \int_c^t f(x)\,\mathrm{d}x \qquad \text{if both limits exist.}$$

Remark. (i) If f is Riemann integrable then its Riemann integral and its improper Riemann integral are equal. Therefore we use the notation $\int_a^b f(x)\,\mathrm{d}x$ also for improper integrals.

(ii) The properties of Theorem 6.43 hold also for improper integrals. In particular, the definition in (iii) does not depend on the chosen c.

Examples 6.61. (i) $$\int_1^\infty \frac{1}{x^s}\,\mathrm{d}x = \begin{cases} \frac{1}{s-1}, & s>1,\\ \text{diverges to } \infty, & s\le 1.\end{cases}$$

Proof. For $\beta > 1$ and $s \neq 1$ we have that

$$\int_1^\beta \frac{1}{x^s}\,\mathrm{d}x = \left[\frac{1}{-s+1}x^{-s+1}\right]_1^\beta = \frac{\beta^{-s+1}-1}{-s+1} \xrightarrow{\beta\to\infty} \begin{cases} \frac{1}{s-1}, & -s+1<0,\\ \infty, & -s+1>0.\end{cases}$$

For $s = 1$ we find $\displaystyle\int_1^\beta \frac{1}{x^s}\,\mathrm{d}x = [\ln x]_1^\beta = \ln\beta \to \infty$ for $\beta\to\infty$. □

(ii) $$\int_0^1 \frac{1}{x^s}\,\mathrm{d}x = \begin{cases} \frac{1}{1-s}, & s<1,\\ \text{diverges to } \infty, & s\ge 1.\end{cases}$$

Proof. Analogously as in (i). □

(iii) $$\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \pi.$$

Proof. Let $c\in\mathbb{R}$ arbitrary. For $a<c$ and $b>c$ we have that

$$\int_c^b \frac{1}{1+x^2}\,\mathrm{d}x = \arctan x\Big|_c^b = \arctan b - \arctan c \xrightarrow{b\to\infty} \frac{\pi}{2} - \arctan c,$$
$$\int_a^c \frac{1}{1+x^2}\,\mathrm{d}x = \arctan x\Big|_a^c = \arctan c - \arctan a \xrightarrow{a\to-\infty} \arctan c - \frac{\pi}{2}.$$

Therefore the improper integral exists and

$$\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2} - \arctan c + \arctan c - \frac{\pi}{2} = \pi.$$ □

Proposition 6.62. *Let $-\infty \le a < b \le \infty$ and $f:(a,b)\to\mathbb{R}$ such that the improper integral $\displaystyle\int_a^b |f(x)|\,\mathrm{d}x$ converges. Then also $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ converges.*

Proof. Let $c\in(a,b)$. We use apply the Cauchy criterion for convergence of a continuous function (see Theorem 5.15) to the continuous function $F(x) := \int_c^x f(t)\,\mathrm{d}t$. For arbitrary $\alpha<\beta\in(a,b)$ it follows that

$$\big|F(\beta)-F(\alpha)\big| = \Big|\int_\alpha^\beta f(t)\,\mathrm{d}t\Big| \le \int_\alpha^\beta |f(t)|\,\mathrm{d}t \to 0 \quad \text{if } \alpha,\beta\to a \text{ or } \alpha,\beta\to b.$$ □

Proposition 6.63 (Monotone convergence). *Let $f:[a,b)\to\mathbb{R}$ be a positive Riemann integrable function. Then $\int\limits_a^b f\,\mathrm{d}t$ converges if and only if f is bounded.*

Proof. For $x \in [a,b)$ let $F(x) = \int_a^x f \, \mathrm{d}t$ and set $s = \sup\{F(x) : x \in [a,b)\}$. If $s < \infty$, then for every $\varepsilon > 0$ there exists an $x_0 \in [a,b)$ such that $F(x_0) > s - \varepsilon$. Since F is monotonically increasing it follows that $F(x) \in (s-\varepsilon, s)$ for all $x \geq x_0$, hence $\lim_{x\to b} F(x) = s$. If $s = \infty$, then it follows analogously that $\lim_{x\to b} F(x) = \infty$. □

Theorem 6.64 (Integral test for convergence of series). *Let $f : [0,\infty) \to \mathbb{R}$ be a monotonically decreasing function. Then*

$$\sum_{n=1}^{\infty} f(n) \quad \textit{converges} \quad \iff \quad \int_1^\infty f(x)\,\mathrm{d}x \quad \textit{converges.}$$

Proof. Since f is decreasing, it follows that

$$f(k+1) \leq \int_k^{k+1} f(t)\ \mathrm{d}t \leq f(k), \qquad k \in \mathbb{N}.$$

Summation from 1 to n yields

$$\sum_{k=2}^{n+1} f(k) \leq \int_1^{n+1} f(t)\ \mathrm{d}t \leq \sum_{k=1}^{n} f(k), \qquad k \in \mathbb{N}.$$

Therefore the series converges if and only if the integral converges. □

Example 6.65. Let $s > 1$. Since by Theorem 6.64

$$\int_1^\infty x^{-s}\ \mathrm{d}x < \sum_{n=1}^{\infty} n^{-s} < 1 + \int_1^\infty x^{-s}\ \mathrm{d}x,$$

we have the chain of strict inequalities

$$\frac{1}{s-1} < \sum_{n=1}^{\infty} n^{-s} < \frac{s}{s-1}.$$

For $s > 1$ let $\zeta(s) = \sum_{n=1}^\infty n^{-s}$. ζ is called the *Riemann zeta function.*

For series we could proof that the convergence of $\sum_{j=1}^\infty a_n$ implies that $a_n \to \infty$ for $n \to \infty$. For improper integrals, however, this is no longer true as the following example show.

Example 6.66 (Fresnel integral).

$$\int_0^\infty \sin(t^2)\ \mathrm{d}t = \lim_{b\to\infty} \int_0^b \sin(t^2)\ \mathrm{d}t = \lim_{b\to\infty} \frac{1}{2} \int_0^b \frac{\sin(u)}{\sqrt{u}}\ \mathrm{d}u$$

where we used the substitution $u = t^2$. To see that the integral converges we write $b = \pi n + s$ with $n \in \mathbb{N}$ and $s \in [0,\pi)$. Then

$$\int_0^b \frac{\sin(u)}{\sqrt{u}}\ \mathrm{d}u = \sum_{m=1}^{n} \int_{\pi(m-1)}^{\pi m} \frac{\sin(u)}{\sqrt{u}}\ \mathrm{d}u + \int_{\pi m}^{\pi m + s} \frac{\sin(u)}{\sqrt{u}}\ \mathrm{d}u.$$

For $n \to \infty$, the sum sum converges by the Leibniz criterion for alternating series (Theorem 4.54) while the absolute value of the last integral is smaller than $\frac{\pi}{\sqrt{n}}$. There exist functions that are unbounded but whose integral is finite:

$$\int_0^\infty 2t \sin(t^4)\ \mathrm{d}t = \int_0^\infty \sin(u^2)\ \mathrm{d}u < \infty,$$

where we used the same substitution as above.

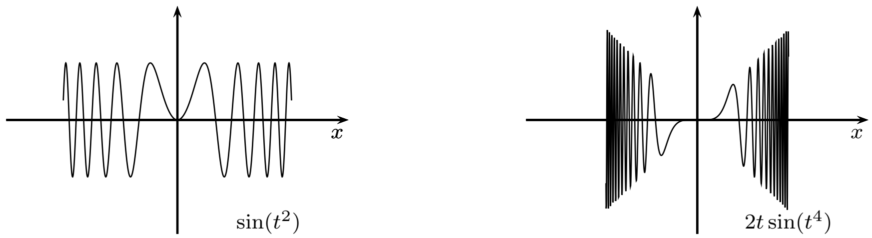


FIGURE 6.7: The function does not tend to zero, yet its integral is finite.

6.5 Differentiation and integration of sequences of functions

Theorem 6.67. *For all $n \in \mathbb{N}$ let $f_n : [a, b] \to \mathbb{R}$ be continuous and assume that $(f_n)_{n\in\mathbb{N}}$ converges uniformly. Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x. \tag{6.22}$$

Proof. Let f be the uniform limit of $(f_n)_{n\in\mathbb{N}}$. Then f is continuous by Theorem 5.40 and Riemann integrable by Theorem 6.41. Equation 6.22 follows from

$$\begin{aligned}\Big|\int_a^b f_n\ \mathrm{d}x - \int_a^b f\ \mathrm{d}x\Big| = \int_a^b \big|f_n - f\big|\ \mathrm{d}x \leq \int_a^b \big\|f_n - f\big\|_\infty\ \mathrm{d}x \\ \leq (b-a)\big\|f_n - f\big\|_\infty.\end{aligned}$$

□

Example 6.68. In Theorem 6.67 pointwise convergence of the f_n is not enough. For $n \in \mathbb{N}$ let

$$f_n(x) := \begin{cases} 2n^2x, & 0 \leq x \leq \frac{1}{2n}, \\ 2n - 2n^2x, & \frac{1}{2n} < x \leq \frac{1}{n}, \\ 0, & x \geq \frac{1}{2n}. \end{cases}$$

We saw in Example 5.35 and Exercise 5.7 that the sequence of functions converges pointwise to 0. Obviously

$$\int_0^\infty f(x)\,\mathrm{d}x = 0 \qquad \text{but} \qquad \int_0^\infty f_n(x)\,\mathrm{d}x = \frac{1}{2}, \quad n \in \mathbb{N}.$$

Remark. Theorem 6.67 implies that the integral is a continuous linear operator from the space of the continuous functions on $[a, b]$ to $\mathbb{R}$:

$$\int : \quad C([a,b],\mathbb{R}) \to \mathbb{R}, \qquad f \mapsto \int_a^b f(t)\ \mathrm{d}t.$$

Theorem 6.69. *For all $n \in \mathbb{N}$ let $f_n : [a, b] \to \mathbb{R}$ be continuously differentiable and assume that the sequence of the derivatives $(f_n')_{n\in\mathbb{N}}$ converges uniformly and there exists a $p \in [a, b]$ such that $(f_n(p))_{n\in\mathbb{N}}$ converges. Then the sequence $(f_n)_{n\in\mathbb{N}}$ converges pointwise to a continuously differentiable function $f : [a, b] \to \mathbb{R}$ and*

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(\lim_{n\to\infty} f_n)(x) = \lim_{n\to\infty} f_n'(x), \qquad x \in [a, b]. \tag{6.23}$$

Proof. Note that if the f_n converge uniformly, then the restrictions $f_n|_D$ to any subinterval $D \subseteq [a,b]$ also converges uniformly. For all $n \in \mathbb{N}$ and $x \in [a,b]$ we have that $f_n(x) = \int_p^x f_n'(t)\,\mathrm{d}t$, therefore we can define f as the pointwise limit of the sequence $(f_n)_{n\in\mathbb{N}}$ by

$$f(x) := \lim_{n\to\infty} f_n(p) + \lim_{n\to\infty} \int_p^x f_n(t)'\ \mathrm{d}t = \lim_{n\to\infty} f_n(p) + \int_p^x \lim_{n\to\infty} f_n(t)'\ \mathrm{d}t$$

where the last equality follows from Theorem 6.67. Therefore, f is a continuously differentiable function and satisfies (6.23). □

Note that in the preceding theorem all assumptions are necessary. For example, the sequence $(f_n)_{n\in\mathbb{N}}$ defined by

$$f_n(x) = \frac{\sin(nx)}{n}, \qquad x \in \mathbb{R},$$

converges uniformly to 0, but the sequence of its derivatives $f_n'(x) = \cos(nx)$ does not even converge pointwise.

Corollary 6.70. *Let f be defined by a power series $\sum_{n=0}^{\infty} c_n(x-a)^n$ with radius of convergence R. Then the* formal integral *and the* formal derivative *of f*

$$\sum_{n=1}^{\infty} nc_n(x-a)^{n-1} \qquad \text{and} \qquad \sum_{n=0}^{\infty} \frac{c_n}{n+1}(x-a)^{n+1}$$

have the same radius of convergence R and are power series representations of f' and $\int f\,\mathrm{d}x$, respectively, in $B_R(a)$.

Proof. The assertion about the radius of convergence follows easily from Theorem 5.45. All other assertions follow from Theorem 6.67 and Theorem 6.69. □

Example 6.71. The power series representation of $\ln(1+x)$ is

$$\sum_{n=1}^{\infty} \frac{(-)^{n+1}}{n}\, x^n, \qquad |x| < 1, \tag{6.24}$$

and

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \cdots \tag{6.25}$$

The first formula follows because for $|x| < 1$ we have

$$\ln(1+x) = \int_1^x \frac{1}{1+t}\ \mathrm{d}t = \int_1^x \sum_{n=0}^{\infty} (-t)^n\ \mathrm{d}t = \sum_{n=0}^{\infty} \int_1^x (-t)^n\ \mathrm{d}t = \sum_{n=0}^{\infty} \frac{(-)^n\, t^{n+1}}{n+1}\,.$$

Since the logarithm is continuous at $x+1 = 2$ and the series (6.24) converges also for $x = 1$, formula (6.25) follows from Abel's theorem (Theorem 5.48).

Chapter 7

Taylor series and approximation of functions

7.1 Taylor series

Assume that the function f has the power series representation

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n \tag{7.1}$$

with radius $R > 0$.
We already know that f can be differentiated arbitrarily often on $B_R(a)$ (Corollary 6.70). Note that

$$\frac{\mathrm{d}}{\mathrm{d}x^k}\, c_n(x-a)^n = \begin{cases} 0, & k > n, \\ \frac{n!}{(n-k)!}\, c_n (x-a)^{n-k}, & k \le n. \end{cases}$$

This implies that

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} f(a) = \frac{k!}{(k-k)!}\, c_k (x-a)^{k-k} = k!\, c_k.$$

If we insert the resulting formula for the coefficients c_k into the power series representation of f we obtain

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!}\, f^{[n]}(a)\,(x-a)^n. \tag{7.2}$$

Therefore the coefficients of the power series representation (7.1) of f in a are determined by the derivatives of f in a. In particular, the power series representation of f in $B_R(a)$ is unique.

The questions we address in this chapter are whether every function can be approximated by a polynomial and whether every C^∞ function has a power series representation.

Definition 7.1. Let $D \subseteq \mathbb{R}$ an interval, $f : D \to \mathbb{R}$ n-times differentiable at some $p \in D$. Then the polynomial

$$j_p^n f(t) := \sum_{k=0}^{n} \frac{f^{[k]}(p)}{k!}\, t^k \tag{7.3}$$

is called the *nth Taylor polynomial* (*or the n-jet*) of f at p. If $f \in C^\infty(D)$ (i. e., if f is arbitrarily often differentiable), then the power series

$$j_p^\infty f(t) := j_p f(t) := \sum_{k=0}^\infty \frac{f^{[k]}}{k!}\, t^k \tag{7.4}$$

is the *Taylor series* (or *jet*) of f at p. For $n \in \mathbb{N}$ and $x \in D$ we define the *remainder term*

$$R_n(x) := f(x) - j_p^n f(x-p), \tag{7.5}$$

Remark 7.2.

- $f(x) = j_p^n f(x-p) + R_n(x), \quad n \in \mathbb{N},\ x \in D,$
- $f^{[k]}(p) = \dfrac{\mathrm{d}^k}{\mathrm{d}t^k} j_p^n f(0), \quad 0 \le k \le n,$
- $R_n^{[k]}(p) = 0, \quad 0 \le k \le n,$

Formula (7.5) is only the definition of the remainder term. This representation of f is useful because $|R_n|$ can be expressed in terms of the $(n+1)$th derivative of f (if it exists). Hence, when $f^{(n+1)}$ can be estimated, then the nth Taylor polynomial is a good approximation of f.

Theorem 7.3 (Taylor's theorem). *Let $D \subseteq \mathbb{R}$ and interval and $f \in C^{[n+1]}(D)$ (i. e., f is $(n+1)$-times continuously differentiable in D), and let $p,\ x \in D$. Then*

$$f(x) = j_p^n f(x-p) + R_n(x)$$

with

$$R_n(x) = \frac{1}{n!}\int_p^x (x-t)^n f^{[n+1]}(t)\ \mathrm{d}t. \tag{7.6}$$

Proof. We prove formula (7.6) by induction. Since $j_p^n f$ is a polynomial of degree less or equal to n, it follows that $f^{[n+1]} = R_n^{[n+1]}$. Note that

$$R_n^{[n+1]}(p) = f^{[n+1]}(p) \quad \text{and} \quad R_n^{[k]}(p) = 0, \quad 0 \le k \le n.$$

Integration by parts yields

$$\int_p^x (x-t)^n R_n^{[n+1]}(t)\ \mathrm{d}t = \Big[(x-t)^n R_n^{[n]}(t)\Big]_{t=p}^x + n\int_p^x (x-t)^{n-1} R_n^{[n]}(t)\ \mathrm{d}t.$$

For $n \ge 1$ the term in brackets vanishes. Integrating the left hand side n-times by parts we obtain

$$\int_p^x (x-t)^n R_n^{[n+1]}(t)\ \mathrm{d}t = n!\int_p^x R_n'(t)\ \mathrm{d}t = n!(R_n(x) - R_n(p)) = n! R_n(x). \qquad \square$$

Using the intermediate value theorem of integration (Theorem 6.51) it follows that there exists an ξ between x and p such that

$$R_n(x) = \frac{f^{[n+1]}(\xi)}{n!}\int_p^x (x-t)^n\ \mathrm{d}t = \frac{f^{[n+1]}(\xi)}{(n+1)!}(x-p)^{n+1}.$$

For real valued functions the formula above is true even if $f^{[n+1]}$ is not continuous as the next theorem shows.

Theorem 7.4 (Lagrange form of the remainder term). *Let $D \subseteq \mathbb{R}$ and interval and $f \in C^{[n]}(D)$ (i. e., f is n-times continuously differentiable in D), and assume that $f^{[n]}$ is differentiable. For p and $x \in D$ there exists ξ between p and x (excluding p and x) such that*

$$R_n(x) = \frac{f^{[n+1]}(\xi)}{(n+1)!}(x-p)^{n+1}. \tag{7.7}$$

Proof. Let $x \in D$. For $x = p$ there is nothing to show. Now assume $x > p$. (The proof for $x < p$ is analogous.) By assumption R_n (defined in (7.5)) is $(n+1)$ times differentiable on D. By the generalised mean value theorem (Theorem 6.24) there exist $p < \xi_{n+1} < \dots < \xi_1 < x$ such that

$$\begin{aligned}
\frac{R_n(x)}{(x-p)^{n+1}} &= \frac{R_n(x) - R_n(p)}{(x-p)^{n+1} - 0^{n+1}} &&= \frac{1}{n+1}\,\frac{R_n'(\xi_1)}{(\xi_1 - p)^n} \\
&= \frac{1}{n+1}\,\frac{R_n'(\xi_1) - R_n'(p)}{(\xi_1-p)^n - 0^n} &&= \frac{1}{(n+1)n}\,\frac{R_n''(\xi_2)}{(\xi_2-p)^{n-1}} \\
&= \dots &&= \frac{1}{(n+1)!}\,\frac{R_n^{[n]}(\xi_n)}{\xi_n - p} \\
&= \frac{1}{(n+1)!}\,\frac{R_n^{[n]}(\xi_n) - R_n^{[n]}(p)}{(\xi_n - p) - 0} &&= \frac{1}{(n+1)!}\,\frac{R_n^{[n+1]}(\xi_{n+1})}{1} = \frac{f^{[n+1]}(\xi_{n+1})}{(n+1)!}.
\end{aligned}$$

Setting $\xi = \xi_{n+1}$ shows (7.7). □

Remark 7.5. Formula (7.6) is also true for complex functions f, but the Lagrange form (7.7) holds only for real valued functions f (because the proof uses the generalized mean value theorem).

Definition 7.6. Let X, Y normed vector spaces, $D \subseteq X$, $f, g : D \to Y$ and p limit point of $D_{\frac{f}{g}}$. The Landau symbols O and o are defined by

(i) $f(x) = \mathrm{O}(g(x))$, $x \to p$, if

$$\exists\, \delta > 0 \quad \exists\, C > 0 \quad \forall\, x \in D: \quad \|x - p\| < \delta \implies \|f(x)\| \le C\|g(x)\|,$$

(ii) $f(x) = \mathrm{o}(g(x))$, $x \to p$, if $\displaystyle\lim_{x\to p} \frac{\|f(x)\|}{\|g(x)\|} = 0$.

Using the Landau symbols, Theorem 7.4 says

$$(f - R_n)(x) = \mathrm{O}((x-p)^{n+1}) \quad \text{and} \quad (f - R_n)(x) = \mathrm{o}((x-p)^n) \quad \text{for} \quad x \to p,$$

that is, $f - R_n$ *vanishes of order* $(x-p)^n$ as $x \to p$.

Remark 7.7. The radius of convergence of the Taylor series of an arbitrarily often differentiable function can be 0. If the Taylor series converges on an interval, it does not necessarily converge to f. But if f has a power series representation, then it is its Taylor series.

The Taylor series of the exponential function and sin and cos are the power series given in Definition 5.49. Another important example is the *binomial series*.

Definition 7.8. For $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}$ the *generalised binomial coefficients* are defined to be

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}, \qquad \binom{\alpha}{0} := 1.$$

For $\alpha \in \mathbb{N}_0$ this definition coincide with Definition 2.18. As in the proof of Proposition 2.19 it can be shown that

$$\binom{\alpha-1}{k-1} + \binom{\alpha-1}{k} = \binom{\alpha}{k}.$$

Example 7.9. For $\alpha \in \mathbb{R}$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \qquad |x| < 1. \tag{7.8}$$

Proof. For $\alpha \in \mathbb{N}_0$, the assertion is already proved in Theorem 2.22. Now assume that $\alpha \notin \mathbb{N}_0$. The power series in (7.8) has radius of convergence $R = 1$ because

$$\left| \binom{\alpha}{n} \binom{\alpha}{n+1}^{-1} \right| = \left| \frac{n+1}{\alpha - n} \right| \to 1 \quad \text{for } n \to \infty.$$

Let $f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$, $|x| < 1$. Then

$$\begin{aligned}
(1+x)f'(x) &= (1+x) \sum_{n=1}^{\infty} n \binom{\alpha}{n} x^{n-1} = (1+x) \sum_{n=1}^{\infty} \alpha \binom{\alpha-1}{n-1} x^{n-1} \\
&= \alpha(1+x)\Big(1 + \sum_{n=2}^{\infty} \binom{\alpha-1}{n-1} x^{n-1}\Big) = \alpha\Big(1 + \sum_{n=2}^{\infty} \Big(\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1}\Big) x^n\Big) \\
&= \alpha\Big(1 + \sum_{n=1}^{\infty} \binom{\alpha}{n} x^n\Big) = \alpha \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = \alpha f(x).
\end{aligned}$$

Let $\varphi(x) = \frac{f(x)}{(1+x)^\alpha}$, $|x| < 1$. By the result above we find that φ is constant because

$$\varphi'(x) = \frac{(1+x)^\alpha f'(x) - f(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} = 0.$$

Since $f(0) = 1 = (1+0)^\alpha$ it follows that $\varphi = 1$, hence $f(x) = (1+x)^\alpha$, $|x| < 1$. □

Special cases:

(i) $\alpha = -1$: $\quad (1+x)^{-1} = \sum_{k=0}^{\infty} \binom{-1}{k} x^k = \sum_{k=0}^{\infty} (-)^k x^k \qquad$ (geometric series),

(ii) $\alpha = -2$: $\quad (1+x)^{-2} = \sum_{k=0}^{\infty} \binom{-2}{k} x^k = \sum_{k=0}^{\infty} (-)^k (k+1) x^k$,

(iii) $\alpha = \frac{1}{2}$: $\quad \sqrt{1+x} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k = 1 + \frac{1}{2}x - \frac{1}{2 \cdot 4}x^2 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^3 \mp \cdots$,

(iv) $\alpha = -\frac{1}{2}$: $\quad \frac{1}{\sqrt{1+x}} = \sum_{k=0}^{\infty} \binom{-1/2}{k} x^k = 1 - \frac{1}{2}x + \frac{1 \cdot 3}{2 \cdot 4}x^2 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^3 \pm \cdots$,

From (ii), for instance, it follows that

$$\sqrt{2} = \sqrt{1+1} = 1 + \frac{1}{2} - \frac{1}{2 \cdot 4} + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6} - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8} \pm \cdots$$

because the series converges converges also for $x = 1$ by the Leibniz criterion (Theorem 4.54) and is equal to $\sqrt{2}$ by Abel's theorem (Theorem 5.48).

Calculating with Taylor series

Example 7.10. $\lim_{x\to 0} \dfrac{1-\cos x}{x^2} = \dfrac{1}{2}$.

Proof. Let $f(x) = 1 - \cos(x)$. Since $f(0) = f'(0) = 0$ and $f''(0) = 1$ it follows from equation (7.7) that $f(x) = \frac{1}{2}x^2 + x^3\varphi(x)$ where $\varphi(x)$ is bounded by $\frac{1}{4!}\|f'''\|_\infty = \frac{1}{4!}$. Therefore

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0}\big(\frac{1}{2} + x\varphi(x)\big) = \frac{1}{2}. \qquad \square$$

Note that the limit can also be found by applying l'Hospital's rule twice.

Theorem 7.11. *Let $D \subseteq \mathbb{R}$ be an interval, $p \in D$ and $f, g : D \to \mathbb{R}$. If f and g are n-times differentiable in 0 then*

$$j_0^n(f+g) = j_0^n f + j_0^n g, \qquad j_0^n(fg) = j_0^n\Big(j_0^n f \cdot j_0^n g\Big). \tag{7.9}$$

If $f, g \in C^\infty(D)$, then

$$j_0(f+g) = j_0 f + j_0 g, \qquad j_0(fg) = j_0 f \cdot j_0 g \quad (\textit{Cauchy product}).$$

Proof. The first formula in (7.9) follows immediately from the linearity of the differentiation (Theorem 6.8). For the second formula, we define $\widetilde{f}$ and $\widetilde{g}$ by

$$f = j_0^n f + \widetilde{f}, \qquad g = j_0^n g + \widetilde{g}.$$

Obviously $\widetilde{f}^{[k]}(0) = 0$ and $\widetilde{g}^{[k]}(0) = 0$ for $0 \le k \le n$. It follows that

$$f \cdot g = j_0^n f \cdot j_0^n g + (\widetilde{f} g + \widetilde{g} j_0^n f).$$

Since the derivatives of order $0 \le k \le n$ of the terms in brackets are 0, it follows that

$$j_0^n(f\cdot g) = j_0^n(j_0^n f \cdot j_0^n g). \qquad \square$$

Example 7.12. The Taylor series of $\dfrac{\ln(1+x)}{1+x}$ at 0 is $\displaystyle\sum_{n=1}^\infty (-)^{n+1}\Big(\sum_{k=1}^n \frac{1}{k}\Big)x^n$.

Proof. The Taylor series of $\ln(1+x)$ and $(1+x)^{-1}$ at 0 are for $|x| < 1$

$$(1+x)^{-1} = \sum_{n=0}^\infty (-)^n x^n, \qquad \ln(1+x) = \int_0^x (1+t)^{-1}\,\mathrm{d}t = \sum_{n=0}^\infty \frac{(-)^n}{n+1}\,x^{n+1}.$$

Therefore we obtain the desired Taylor series as the Cauchy product of the two series:

$$\begin{aligned}\Big(\sum_{n=0}^\infty (-)^n x^n\Big)\Big(\sum_{n=0}^\infty \frac{(-)^n x^{n+1}}{n+1}\Big) &= -\sum_{n=1}^\infty\Big(\sum_{k=0}^{n-1}(-)^{n-k}\cdot\frac{(-)^{k+1}}{k+1}\Big)x^n \\ &= -\sum_{n=1}^\infty (-)^n\Big(\sum_{k=1}^n \frac{1}{k}\Big)x^n. \qquad \square\end{aligned}$$

Example 7.13. The function $\mathbb{R} \to \mathbb{R}$, $x \mapsto x(1 + x - \cos x)$ has a local minimum at 0.

Proof. The jets are

$$j(1+x-\cos x) = x + \dots \quad\Longrightarrow\quad j(x(1+x-\cos x)) = x^2 + \dots,$$

so the function behaves locally like x^2 and has therefore a local minimum at 0. □

Example 7.14. Use the method of undetermined coefficients to find the Taylor series of tan at 0.

Solution. We have to determine coefficients a_j, $j \in \mathbb{N}_0$ such that locally at 0

$$j_0 \tan x = \sum_{n=0}^{\infty} a_n x^n.$$

Since $\tan(0) = 0$ it follows that $a_0 = 0$. We know that $\tan'(x) = 1 + (\tan(x))^2$ and

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n,$$
$$1 + \Big(\sum_{n=0}^{\infty} a_n x^n\Big)^2 = 1 + \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{n} a_k a_{n-k}\Big) x^n$$

Comparison of the coefficients yield the recursion formula for the a_n:

$$a_0 = 0, \quad a_1 = 1, \quad n a_n = \sum_{k=1}^{n-2} a_k a_{n-k-1}, \quad n \geq 2.$$

In particular, $a_{2n} = 0$ for all $n \in \mathbb{N}_0$. □

The following theorem generalises the chain rule.

Theorem 7.15. *Let $D_f,\ D_g \subseteq \mathbb{R}$ be intervals, $f : D_f \to \mathbb{R}$ and $g : D_g \to \mathbb{R}$ n-times differentiable functions such that $f(D_f) \subseteq D_g$. Moreover let $p \in D_f$ and $q := f(p) \in D_g$. Then*

$$j_p^n(g \circ f) = j_0^n\big((j_q^n g) \circ (j_p^n f - q)\big).$$

If f and g are arbitrarily often differentiable, then

$$j_p(g \circ f) = j_0\big((j_q g) \circ (j_p f - q)\big).$$

Proof. Without restriction, we can assume $p = q = 0$. If g is a polynomial, then the assertion follows from Theorem 7.11. If g is not a polynomial then we define $\widetilde{g}$ by $g = j_0^n g + \widetilde{g}$. We obtain

$$\begin{aligned} j_0^n(g \circ f) &= j_0^n((j_0^n g) \circ f + \widetilde{g} \circ f) = j_0^n((j_0^n g) \circ f) + \underbrace{j_0^n(\widetilde{g} \circ f)}_{=\,0 \text{ by product and chain rule}} \\ &= j_0^n((j_0^n g) \circ j_0^n f). \end{aligned}$$
□

Example 7.16. The fourth Taylor polynomial of $f(x) = \cos(1 - \frac{1}{1+x^2})$ is $1 - \frac{1}{2}x^4$.

Proof. Using the power series representation of the cosine (Definition 5.49) and the geometric series to represent $\frac{1}{1+x^2}$ as a power series, we find

$$j_0^4 \cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4, \qquad j_0^4\big(1 - \frac{1}{1+x^2}\big) = -x^2 + x^4.$$

Therefore we obtain

$$\begin{aligned} j_0^4 f(x) &= j_0^4\big(1 - \tfrac{1}{2}(-x^2 + x^4)^2\big) + \tfrac{1}{24}(-x^2 + x^4)^4\big) \\ &= j_0^4\big(1 - \tfrac{1}{2}(x^4 + \text{higher order term}) + \tfrac{1}{24}(\text{higher order term})\big) \\ &= 1 - \frac{1}{2}x^4. \end{aligned}$$

□

7.2 Construction of differentiable functions

Definition 7.17. Let $D \subseteq \mathbb{R}$ an open interval and $f : D \to \mathbb{R}$. Then f is called *analytic* if for every point $p \in D$ there exists a $\varepsilon > 0$ such that f has a convergent power series representation (centred in p) in $B_\varepsilon(p)$.

By definition, every analytic function is a C^∞ function, but not every C^∞ function is analytic as the following example shows:

Theorem 7.18. *The function*

$$\varphi : \mathbb{R} \to \mathbb{R}, \qquad \varphi(x) = \begin{cases} \exp(-\frac{1}{x^2}), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

lies in $C^\infty(\mathbb{R})$ and $\varphi^{[n]}(0) = 0$ for all $n \in \mathbb{N}$. In particular, the Taylor series of φ at 0 converges in all of $\mathbb{R}$ but it is equal to φ only in the point 0.

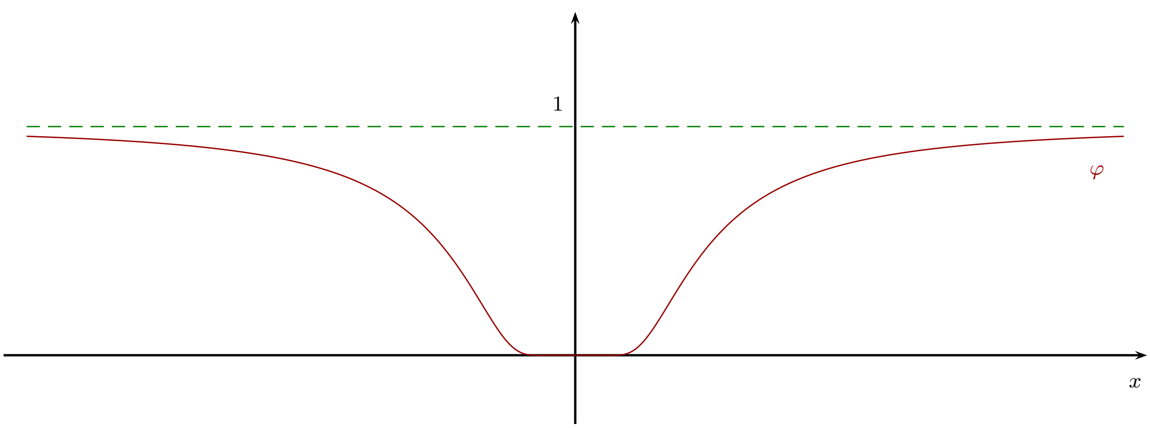


FIGURE 7.1: The non-analytic C^∞ function φ (see Theorem 7.18). Although the plot gives another impression, φ has an isolated global minmum at 0.

Proof. Step 1: For all $k \in \mathbb{N}_0$ exists a polynomial P_k (of degree $n = 3k$) such that

$$\varphi^{[k]}(x) = P_k(x^{-1}) \exp(-x^{-2}), \qquad x \neq 0.$$

We prove the assertion by induction on k. It is clearly true for $k = 0$. Assume now that we know the assertion already shown for some $k \in \mathbb{N}_0$. Then we find that

$$\begin{aligned} \varphi^{[k+1]}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\big(P_k(x^{-1}) \exp(-x^{-2})\big) \\ &= \underbrace{\big(-x^{-2}P_k'(x^{-1}) + 2x^{-3}P_k(x^{-1})\big)}_{:=P_{k+1}(x^{-1})} \exp(-x^{-2}). \end{aligned}$$

Obviously, P_{k+1} is a polynomial in x^{-1} of degree $\deg(P_{k+1}) = \deg(P_k) + 3 = 3(k+1)$. *Step 2*: $\lim\limits_{x\to 0} x^{-k}\exp(-x^{-2}) = 0$ *for all* $k \in \mathbb{N}_0$. We show the assertion by induction on k. It clearly holds for $k = 0$. For $k = 1$ it follows by l'Hospital's rule:

$$\lim_{x\to 0} x^{-1}\,\mathrm{e}^{-\frac{1}{x^2}} = \lim_{x\to 0} \frac{x^{-1}}{\mathrm{e}^{\frac{1}{x^2}}} = \lim_{x\to 0} \frac{-x^{-2}}{-2x^{-3}\,\mathrm{e}^{\frac{1}{x^2}}} = \lim_{x\to 0} x\,\mathrm{e}^{-\frac{1}{x^2}} = 0.$$

Now assume that the assertion holds for all $0 \le k \le n$ for some $n \in \mathbb{N}_0$. Then, again with the help of l'Hospital's rule and the induction hypothesis, we obtain

$$\begin{aligned} \lim_{x\to 0} x^{-n-1}\,\mathrm{e}^{-\frac{1}{x^2}} &= \lim_{x\to 0} \frac{x^{-n-1}}{\mathrm{e}^{\frac{1}{x^2}}} = \lim_{x\to 0} \frac{-(n+1)x^{-n-2}}{-2x^{-3}\,\mathrm{e}^{\frac{1}{x^2}}} \\ &= \frac{n+1}{2}\lim_{x\to 0} x^{-(n-1)}\,\mathrm{e}^{-\frac{1}{x^2}} = 0. \end{aligned}$$

Step 3: *All derivatives of* φ *in* 0 *exist and* $\varphi^{[k]}(0) = 0$ *for all* $k \in \mathbb{N}_0$.
Again, the assertion is proved by induction. For $k = 0$ the assertion follows directly from the definition of φ. If $k = 1$ then

$$\varphi^{[k]}(0) = \varphi'(0) = \lim_{x\to 0} \frac{\varphi(x) - \varphi(0)}{x - 0} = \lim_{x\to 0} \frac{\varphi(x)}{x} = \lim_{x\to 0} x^{-1}\exp(-x^{-2}) = 0.$$

Assume that the assertion is true for some $k \in \mathbb{N}$. Then, by induction hypothesis and the results of step 1 and step 2,

$$\begin{aligned} \varphi^{[k+1]}(0) &= \lim_{x\to 0} \frac{\varphi^{[k]}(x) - \varphi^{[k]}(0)}{x - 0} = \lim_{x\to 0} \frac{\varphi^{[k]}(x)}{x} \\ &= \lim_{x\to 0} x^{-1}P_k(x^{-1})\,\exp(-x^{-2}) = 0. \end{aligned}$$

It follows that all coefficients of the Taylor series of φ in 0 are 0, therefore its radius of convergence is ∞. Since $\varphi(x) = 0$ if and only if $x = 0$ it follows that $\varphi(x) = j_0^\infty\varphi(x)$ if and only if $x = 0$. □

Theorem 7.19. *Let* $r, \varepsilon > 0$. *Then there exists a function* $\psi \in C^\infty(\mathbb{R})$ *such that* $0 \le \psi \le 1$ *and*

$$\psi(x) = 1 \iff |x| \le r, \qquad \psi(x) = 0 \iff |x| \ge r + \varepsilon.$$

Proof. We use the function φ from Theorem 7.18 to construct ψ. First we define the function

$$\mu : \mathbb{R} \to \mathbb{R}, \qquad \mu(x) = \begin{cases} \mathrm{e}^{-\frac{1}{x^2}}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

By the theorem above, $\mu \in C^\infty(\mathbb{R})$ (but it is not analytic in 0). Next we define

$$\mu_\varepsilon : \mathbb{R} \to \mathbb{R}, \qquad \mu_\varepsilon(x) = \frac{\mu(x)}{\mu(x) + \mu(\varepsilon - x)}.$$

The function μ_ε satisfies $0 \le \mu_\varepsilon \le 1$ and

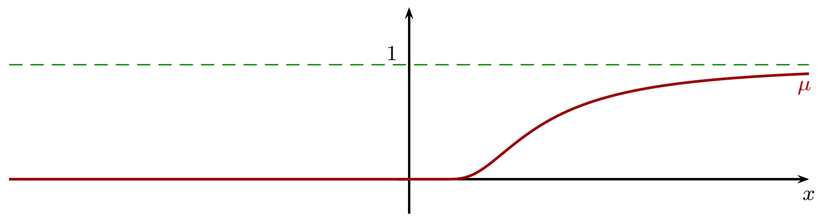
$$\mu_\varepsilon(x) = 0 \iff x \le 0, \qquad \mu_\varepsilon(x) = 1 \iff x \ge \varepsilon.$$

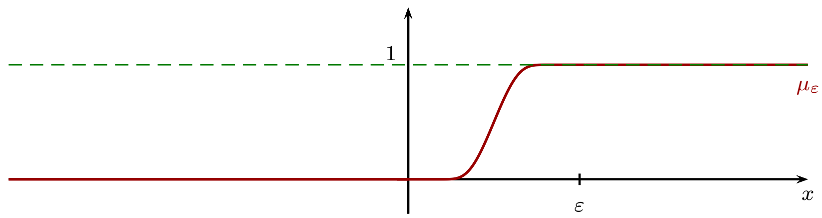
Finally,

$$\psi : \mathbb{R} \to \mathbb{R}, \qquad \psi(x) = 1 - \mu_\varepsilon(|x| - r)$$

has the desired properties (note that ψ is differentiable of arbitrary order at 0 because it is locally constant at 0).

□





The existence of a function as in Theorem 7.19 implies that if a function is known only locally at a point, nothing can be deduced about the global behaviour of the function.

For instance, let f, g be arbitrary functions on $\mathbb{R}$ and ψ as in Theorem 7.19. Let

$$h = (1-\psi)f + \psi g.$$

Then $h(x) = f(x)$ for $|x| \geq r + \varepsilon$ and $h(x) = g(x)$ for $|x| \leq r$.
The next theorem says that every power series is the Taylor series of a C^∞ function. Of course, the radius of convergence may be 0 and the Taylor series does not need to represent anywhere apart from the point in which we calculate the Taylor expansion.

Theorem 7.20 (Borel's theorem). *Let $(c_n)_{n\in\mathbb{N}_0} \subseteq \mathbb{R}$. Then there exists a function $f \in C^\infty(\mathbb{R})$ such that*

$$j_0 f(x) = \sum_{n=0}^{\infty} c_n x^n.$$

Proof. Let ψ be the function of Theorem 7.19 with $r = \varepsilon = \frac{1}{2}$. For $a \geq 1$ define

$$\xi_a : \mathbb{R} \to \mathbb{R}, \qquad \xi_a(x) = x \cdot \psi(ax).$$

Then $-a^{-1} \leq \xi_a \leq a^{-1}$ and

$$\xi_a(x) = x \iff |ax| \leq \frac{1}{2}, \qquad \xi_a(x) = 0 \iff |ax| \geq 1.$$

Note that $\xi_a(x) = x$ for $|x|$ sufficiently small. We construct f as the series

$$f(x) = \sum_{k=0}^{\infty} c_k \big(\xi_{a_k}(x)\big)^k.$$

We have to show that the a_k can be chosen such that for all $n \in \mathbb{N}_0$ the series of the formal derivatives is uniformly convergent. Then f is arbitrarily often differentiable at 0 and the formal nth derivatives of f equal its nth derivative (Theorem 6.69). That the Taylor series of f equals the given power series is then clear because $\xi_{a_k}(x) = x$ locally at 0.

Set $\xi := \xi_1$. Since for all $k \in \mathbb{N}_0$ and all $n \in \mathbb{N}_0$ the function $(\xi^k)^{[n]}$ is arbitrarily often differentiable, it is bounded on the compact interval $[-1, 1]$ (Theorem 5.30). Outside of the interval it is zero, therefore there exist constants M_{nk} such that

$$\|(\xi^k)^{[n]}\|_\infty \leq M_{nk}.$$

For $a \geq$ the chain rule yields

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}\, \xi^k_{a_k}(x) = a_k^{n-k} (\xi^k)^{[n]}(a_k x), \qquad k,\, n \in \mathbb{N}_0,$$

hence

$$\left\| \frac{\mathrm{d}^n}{\mathrm{d}x^n}\, \xi^k_{a_k} \right\|_\infty \leq a_k^{n-k} M_{nk}, \qquad k,\, n \in \mathbb{N}_0,$$

For fixed $k \in \mathbb{N}_0$ we can find $a_k > 1$ such that

$$|c_k| a_k^{n-k} M_{nk} < 2^{-k} \quad \text{for all } n \leq k.$$

Hence, for fixed $n \in \mathbb{N}_0$, we have that

$$\left\| \frac{\mathrm{d}^n}{\mathrm{d}x^n}\, c_k \xi^k_{a_k} \right\|_\infty < 2^{-k} \quad \text{for all } k \geq n.$$

Therefore, the series of the formal derivatives of f converges. □

7.3 Dirac sequences

Definition 7.21. A *Dirac sequence* is a sequence of continuous functions $(\delta_n)_{n\in\mathbb{N}}$ on $\mathbb{R}$ such that

(D1) $\delta_n \geq 0, \quad n \in \mathbb{N}$.

(D2) $\displaystyle\int_{-\infty}^{\infty} \delta_n(x)\ \mathrm{d}x = 1, \quad n \in \mathbb{N}$.

(D3) For all $\eta, \varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$\int_{-\infty}^{-\eta} \delta_n(x)\ \mathrm{d}x + \int_{\eta}^{\infty} \delta_n(x)\ \mathrm{d}x < \varepsilon, \quad n \geq N.$$

Example 7.22. Let $\delta : \mathbb{R} \to [0, \infty)$ an arbitrary Riemann integrable function such that $\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1$. Then the sequence of functions $(\delta_n)_{n\in\mathbb{N}}$ defined by

$$\delta_n : \mathbb{R} \to \mathbb{R}, \quad \delta_n(x) = n\delta(nx)$$

is a Dirac sequence.

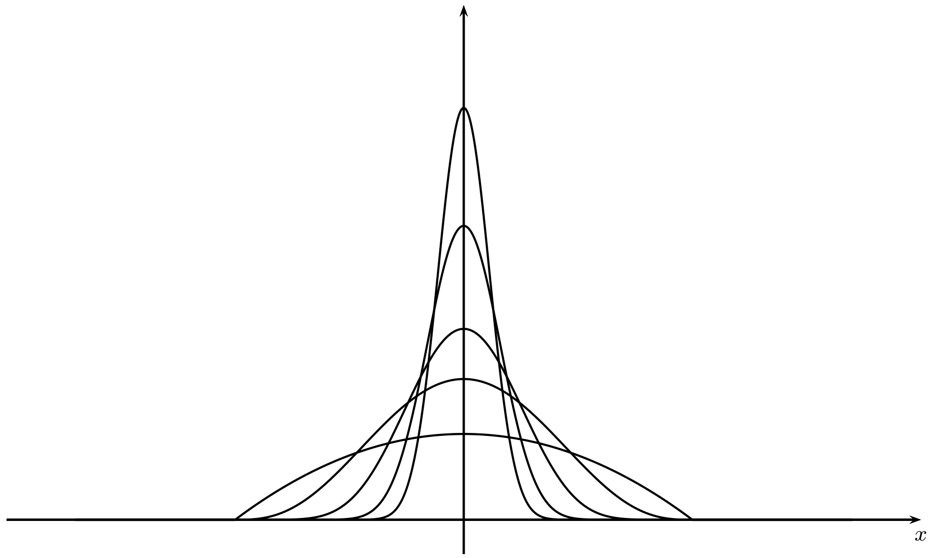
Proof. Property (D1) is clear. Property (D2) follows with the substitution $t = nx$:

$$\int_{-\infty}^{\infty} \delta_n(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} n\,\delta(nx)\,\mathrm{d}x = \int_{-\infty}^{\infty} \delta(t)\,\mathrm{d}t = 1.$$

Let $\eta, \varepsilon > 0$. Since δ is positive and integrable, there exists an $R > 0$ such that $\int_R^\infty \delta(x)\ \mathrm{d}x < \frac{\varepsilon}{2}$. Therefore for $n\eta > R$:

$$\int_{\eta}^{\infty} \delta_n(x)\,\mathrm{d}x = \int_{\eta}^{\infty} n\,\delta(nx)\,\mathrm{d}x = \int_{n\eta}^{\infty} \delta(t)\,\mathrm{d}t < \frac{\varepsilon}{2}.$$

Analogously for $\int_{-\infty}^{-\eta} \delta_n(x)\ \mathrm{d}x$. □



Definition 7.23. Let $f, g : \mathbb{R} \to \mathbb{R}$ be integrable functions. Then the *convolution* $f * g$ of f and g is defined by

$$f * g : \mathbb{R} \to \mathbb{R}, \quad (f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\ \mathrm{d}t.$$

Note that the convolution is commutative since the transformation $s = x - t$ yields

$$(f * g)(x) = -\int_{\infty}^{-\infty} f(x-s)g(s)\ \mathrm{d}s = \int_{-\infty}^{\infty} f(x-s)g(s)\ \mathrm{d}s = (g * f)(x).$$

Theorem 7.24 (Dirac approximation). *Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and locally integrable (i. e., for every compact set $K \subseteq \mathbb{R}$ the restriction $f|_K$ is integrable). Let $K \subseteq \mathbb{R}$ a compact interval such that $f|_K$ is continuous in K. Let $(\delta_n)_{n\in\mathbb{N}}$ be a Dirac sequence and define*

$$f_n(x) := (\delta_n * f)(x) = \int_{-\infty}^{\infty} f(t)\delta_n(x-t)\ \mathrm{d}t = \int_{-\infty}^{\infty} f(x-t)\delta_n(t)\ \mathrm{d}t.$$

Then $f_n|_K \to f|_K$ uniformly.

Proof. For the proof of the uniform convergence we have to estimate

$$\begin{aligned} f_n(x) - f(x) &= \int_{-\infty}^{\infty} f(x-t)\delta_n(t)\ \mathrm{d}t - f(x) \\ &= \int_{-\infty}^{\infty} f(x-t)\delta_n(t)\ \mathrm{d}t - \int_{-\infty}^{\infty} f(x)\delta_n(t)\ \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \left(f(x-t) - f(x)\right)\delta_n(t)\ \mathrm{d}t \end{aligned}$$

independently of x for $x \in K$. Let $\varepsilon > 0$. Since f is uniformly continuous on K, there exists a $\eta > 0$ such that

$$|t| < \eta,\ x \in D \quad \Longrightarrow \quad |f(x-t) - f(x)| < \varepsilon.$$

(Note that $x-t$ does not necessarily belong to K.) By assumption, f is bounded. Let $M \in \mathbb{R}$ such that $|f| < M$ and choose $N \in \mathbb{N}$ such that

$$\int_{-\infty}^{\eta} \delta_n(t)\ \mathrm{d}t + \int_{\eta}^{\infty} \delta_n(t)\ \mathrm{d}t < \frac{\varepsilon}{2M}, \qquad n \geq N.$$

Then

$$|f_n(x) - f(x)| \leq \int_{-\infty}^{-\eta} + \int_{-\eta}^{\eta} + \int_{\eta}^{\infty} |f(x-t) - f(x)|\delta_n(t)\ \mathrm{d}t.$$

Since $|f(x-t) - f(x)| < 2M$ it follows that

$$\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty} |f(x-t) - f(x)|\delta_n(t)\ \mathrm{d}t < 2M\frac{\varepsilon}{2M} = \varepsilon,$$

and $|f(x-t) - f(x)| < \varepsilon$ for $|x-t| < \eta$ implies

$$\int_{-\eta}^{\eta} |f(x-t) - f(x)|\delta_n(t)\ \mathrm{d}t < \varepsilon.$$

We have shown that $\|f_n|_K - f|_K\|_\infty < \varepsilon$, $n \geq N$. Since ε was arbitrary, the theorem is proved. □

In the proof we have used that there exists an $\eta > 0$ such that $|f(x-t) - f(x)| < \varepsilon$ for all $x \in K$ and $|t| < \eta$. Such an η can be found as the minimum of η_D, η_+ and η_- where η_D is such that $|f(x-t) - f(x)| < \varepsilon$ for all $x \in K$ and $|t| < \eta$ such that $x - t \in K$ (uniform continuity of f in K), and $\eta_\pm$ are such that $|f(x_\pm) - f(y)| < \varepsilon/2$ for all $y \in \mathbb{R}$ such that $|x-y| < \eta$ where $x_\pm$ are the endpoints of K.

Proposition 7.25. *There exists a Dirac sequence $(\delta_n)_{n\in\mathbb{N}}$ such that the restrictions $\delta_n|_{[-1,1]}$ are polynomials.*

Proof. Define the sequence $(\delta_n)_{n\in\mathbb{N}}$ by $\delta_n : \mathbb{R} \to \mathbb{R}$ with

$$\delta_n(x) = \begin{cases} c_n^{-1}(1-x^2)^n, & |x| \leq 1, \\ 0, & |x| > 1, \end{cases} \qquad \text{where} \quad c_n := \int_{-1}^{1} (1-x^2)^n \, \mathrm{d}x.$$

Obviously, all δ_n are continuous on $\mathbb{R}$, $\delta_n \geq 0$ and $\int_{-\infty}^{\infty} \delta_n(x)\, \mathrm{d}x = 1$. It remains to show property (D3) of Dirac sequences. First we estimate the constants c_n:

$$c_n = 2\int_0^1 (1-x^2)^n\ \mathrm{d}x = 2\int_0^1 (1+x)^n(1-x)^n\ \mathrm{d}x \geq 2\int_0^1 (1-x)^n\ \mathrm{d}x = \frac{2}{n+1}.$$

Now for $0 < \eta < 1$

$$\begin{aligned} \int_{\eta}^{\infty} \delta_n(x)\ \mathrm{d}x &= \int_{\eta}^{1} c_n^{-1}(1-x^2)^n\ \mathrm{d}x \leq \frac{n+1}{2}\int_{\eta}^{1} (1-\eta^2)^n\ \mathrm{d}x \\ &= \frac{n+1}{2}(1-\eta^2)^n(1-\eta), \end{aligned}$$

which tends to 0 for $n \to \infty$. If $\eta \geq 1$ then $\int_\eta^\infty \delta_n(x)\ \mathrm{d}x = 0$ by definition of δ_n. □

Theorem 7.26 (Weierstraß approximation theorem). *A continuous function on a compact interval is the uniform limit of polynomials on the compact interval.*

Proof. First we show the assertion for continuous functions f such that

$$f : [0,1] \to \mathbb{R}, \qquad f(0) = f(1) = 0.$$

We can extend f continuously to $\mathbb{R}$ by setting $f(x) = 0$ for $x \in \mathbb{R} \setminus [0,1]$. Let $(\delta_n)_{n\in\mathbb{N}}$ as in Proposition 7.25. By Theorem 7.24 the sequence $f_n := \delta_n * f$ converges to f uniformly. We have to show that the restrictions $f_n|_{[0,1]}$ polynomials. Since $\delta_n|_{[-1,1]}$ are polynomials of degree $2n$, there exists a representation

$$\delta_n(x-t) = g_0(t) + g_1(t)x \cdots + g_{2n}(t)x^{2n}, \qquad x,\, t \in [0,1].$$

Hence for all $x \in [0,1]$:

$$\begin{aligned} f_n(x) &= \int_{-\infty}^{\infty} f(t)\delta_n(x-t)\ \mathrm{d}t = \int_0^1 f(t)\delta_n(x-t)\ \mathrm{d}t \\ &= \int_0^1 f(t)\big(g_0(t) + g_1(t)x \cdots + g_{2n}(t)x^{2n}\big)\ \mathrm{d}t = a_0 + a_1 x \cdots + a_{2n}x^{2n} \end{aligned}$$

with coefficients $a_j := \int_0^1 f(t)g_j(t)\ \mathrm{d}t$.
Now let $g : [a,b] \to \mathbb{R}$ an arbitrary continuous functions. Let $\varphi : [0,1] \to \mathbb{R},\ \varphi(y) = a + y(b-a)$, and define

$$f : [0,1] \to \mathbb{R}, \qquad f(x) = (g \circ \varphi)(x) - (g(a) - x(g(a) - g(b)).$$

By what we have prove so far, there exists a sequence $(P_n)_{n\in\mathbb{N}}$ of polynomials that converges uniformly to f. For $n \in \mathbb{N}$ and $x \in [a,b]$ define $Q_n(x) := P_n(\varphi^{-1}(x)) + (g(a) - \varphi^{-1}(x)(g(a) - g(b))$. Then $(Q_n)_{n\in\mathbb{N}}$ converges uniformly to g on $[a,b]$. □

The theorem can be generalised to the so-called Stone-Weierstraß theorem, see Theorem 8.38. (See [Rud76, Theorem 7.32]).

Chapter 8

Basic Topology

8.1 Topological spaces

Recall that a metric space (X, d) is a set X together with a function

$$d : X \times X \to [0, \infty),$$

such that $d(x, y) = d(y, x)$, $d(x, y) = 0 \iff x = y$ and $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$. For arbitrary $r > 0$ and $a \in X$ we defined

$$B_r(a) := \{x \in X \,:\, d(a, x) < r\} =: \textit{open ball with centre at } a \textit{ and radius } r,$$
$$K_r(a) := \{x \in X \,:\, d(a, x) \leq r\} =: \textit{closed ball with centre at } a \textit{ and radius } r.$$

If $Y \subseteq X$, then (Y, d_Y) with $d_Y = d|_{Y \times Y}$ is also a metric space. d_Y is called the induced metric.
A sequence $(x_n)_{n\in\mathbb{N}}$ is called convergent if there exists an $p \in X$ such that

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} : \quad n \geq N \implies x_n \in B_\varepsilon(p).$$

A sequence $(x_n)_{n\in\mathbb{N}}$ is called a Cauchy sequence if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$m, n \geq N \implies d(x_m, x_n) < \varepsilon.$$

We showed that every convergent sequence sequence is a Cauchy sequence, but not every Cauchy sequence converges. A metric space in which every Cauchy sequence converges is called a complete metric space.

Metric spaces are special cases of the more abstract concept of topological spaces.

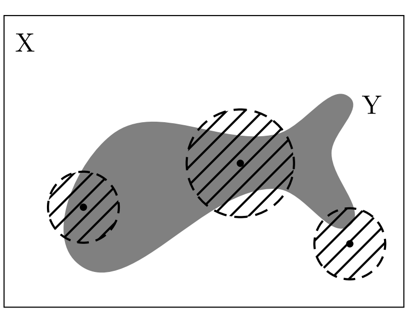


FIGURE 8.1: Examples for balls in the metric space X . . .

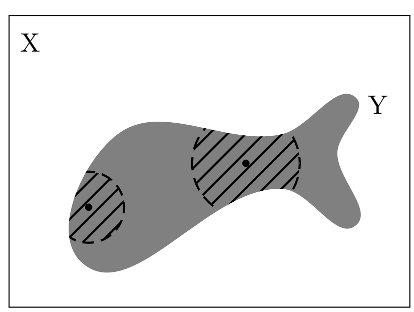


FIGURE 8.2: . . . and the induced balls in the subspace Y. (The right lower X-ball in the left picture does not induce a ball in Y because its centre does not belong to Y.)

Definition 8.1. A *topological space* $(X, \mathcal{O})$ is a set X together with a subset $\mathcal{O} \subseteq \mathbb{P}(X)$ of the power set of X such that

(i) $X,\ \emptyset \in \mathcal{O}$.

(ii) $U_1, \dots, U_n \in \mathcal{O} \implies U_1 \cap \dots \cap U_n \in \mathcal{O}$.

(iii) $U_\lambda \in \mathcal{O},\ \lambda \in \Lambda \implies \bigcup_{\lambda\in\Lambda} U_\lambda \in \mathcal{O}$.

$\mathcal{O}$ is called the *topology* of X and its members are called *open sets* of X.

By definition, X and $\emptyset$ are open, the finite intersection of open sets is open, the arbitrary union of open sets is open.

Definition 8.2. Let $(X, \mathcal{O})$ be a topological space and let $p \in X$. A set V is called a *neighbourhood of p* if there exists an open set U such that $p \in U \subseteq V$.

Lemma 8.3. *Let $(X, \mathcal{O})$ be a topological space and $U \subseteq X$.*

(i) *U is open if and only if for each $p \in U$ there exists an open set V such that $p \in V \subseteq U$.*

(ii) *U is open if and only if it is a neighbourhood of each $p \in U$.*

Proof. (i) Assume that U is open and let $p \in U$. Then we can choose $V = U$.
Now assume that for each $p \in U$ there exists an open V_p such that $p \in V_p \subseteq U$. Then $U = \bigcup_{p\in U} V_p$, hence U is open as union of open sets.
(ii) Follows immediately from (i). □

Definition 8.4. Let $(X, \mathcal{O})$ be a topological space and $Y \subseteq X$. Then X induces *subspace topology* on Y by

$$U \subseteq Y \text{ is open in } Y \quad\iff\quad \exists V \in \mathcal{O}:\ U = V \cap Y.$$

It is not hard to see that Y with the induced topology is indeed a topological space.

Definition 8.5. A topological space $(X, \mathcal{O})$ is called a *Hausdorff space* (or T_2 space) if for all $x, y \in X$ with $x \neq y$ there exist neighbourhoods V_x of x and V_y of y such that $V_x \cap V_y = \emptyset$.

Examples 8.6. (i) Let X be an arbitrary set and define $\mathcal{O} = \{\emptyset, X\}$. Then $(X, \mathcal{O})$ is a topological space. It is Hausdorff if and only if $|X| \leq 1$.

(ii) Let X be an arbitrary set and define $\mathcal{O} = \mathbb{P}(X)$. Then $(X, \mathcal{O})$ is a Hausdorff space.

(iii) Every subspace of a Hausdorff space is again a Hausdorff space.

The topology in example (i) is called the *trivial topology*, the topology in example (ii) is called the *discrete topology*.

Example 8.7 (Topology induced by a metric). Let (X, d) be a metric space. Then d induces a topology on X: a set $U \subseteq X$ is open if and only if

$$\forall p \in U \quad \exists \varepsilon > 0 \quad B_\varepsilon(p) \subseteq U.$$

It can be shown (Exercise 8.2) that X with the induced topology is indeed a topological space with the Hausdorff property, that the open balls are open sets and that the closed balls are closed sets (see Definition 8.8).

Whenever we speak of a metric space as a topological space, we refer to the topology induced by the metric. For example, the topology in $\mathbb{R}$ is generated by the open intervals. Note, however, that not every open set is an open interval.

Definition 8.8. Let $(X, \mathcal{O})$ be a topological space and let $A \subseteq X$. A point p is called a *boundary point* of A if for every neighbourhood U_p of p

$$A \cap U_p \neq \emptyset \quad \text{and} \quad (X \setminus A) \cap U_p \neq \emptyset.$$

The set of all boundary points of A is denoted by ∂A.
The *interior* of A is $A^\circ := A \setminus \partial A$.
The *closure* of A is $\overline{A} := A \cup \partial A$.
The set A is called *closed* if it contains all its boundary points (i. e. $\partial A \subseteq A$).

Note that a boundary point does not necessarily belong to A. For example, if A is an open set, then $A \cap \partial A = \emptyset$.

Proposition 8.9. *Let $(X, \mathcal{O})$ be a topological space. Then a subset $A \subseteq X$ is closed if and only if $X \setminus A$ is open.*

Proof. Assume that A is closed and let $p \in X \setminus A$. If every neighbourhood of p would have non-empty intersection with A, then $p \in \partial A \subseteq A$ (note that every neighbourhood of p contains $p \notin A$). Therefore there exists an neighbourhood of p that has empty intersection with A, hence $X \setminus A$ is open.
Now assume that $X \setminus A$ is open and let $p \in X \setminus A$. Then there exists a neighbourhood U of p such that $U \cap A = \emptyset$, hence $p \notin \partial A$ □

It follows that X and $\emptyset$ are closed, the finite union of closed sets is open, the arbitrary intersection of closed sets is open.

Remark 8.10. Let X be a topological space and $Y \subseteq X$. Then a set $A \subseteq Y$ is closed in Y if and only if there exists $B \subseteq X$ such that B is closed in X and $B \cap Y = A$.
Indeed, since A is closed in Y, $Y \setminus A$ is open in Y, hence, by definition, there exists an $U \in X$, such that U is open in X and $U \cap Y = Y \setminus A$. Then $B := X \setminus U$ is closed in X and $B \cap Y = A$.

Lemma 8.11 (de Morgan's laws). *Let X and Λ be sets and $M_\lambda \subseteq X$, $\lambda \in \Lambda$. Then*

$$X \setminus \bigcup_{\lambda \in \Lambda} M_\lambda = \bigcap_{\lambda \in \Lambda} (X \setminus M_\lambda), \qquad X \setminus \bigcap_{\lambda \in \Lambda} M_\lambda = \bigcup_{\lambda \in \Lambda} (X \setminus M_\lambda).$$

Proposition 8.12. *Let $(X, \mathcal{O})$ be a topological space and $A \subseteq X$. Then A° is the union of all open sets that are contained in A and $\overline{A}$ is the intersection of all closed subsets of X that contain A*

Corollary 8.13. (i) *The interior of a set is open. A is open if and only if $A^\circ = A$.*
(ii) *The closure of a set is closed. A is closed if and only if $\overline{A} = A$.*

Proof of Proposition 8.12. Let B be the union of all open subsets of A. Then B is open and we have to show that $B = A^\circ$.
If $p \in A^\circ$ there exists an open neighbourhood U of p that $U \cap (X \setminus A) = \emptyset$, that is, $p \in U \subseteq A$. In particular p lies in the union of all open sets contained in A, so we have shown $A^\circ \subseteq B$.
Obviously, B is open and contained in A, therefore, for each $p \in B$ there exists an open set U such that $p \in U \subseteq B \subseteq A$, hence $p \in A^\circ$ which proves $B \subseteq A^\circ$.
The second part of the proposition follows from Proposition 8.9 and de Morgan's laws. □

Note that there are sets that are neither closed nor open, for example $[0,1) \subseteq \mathbb{R}$. Moreover, a set can be both closed and open, for examples, in a non-empty topological space with the discrete topology every set is open and closed.

Proposition 8.14. *Let (X,d) be a metric space with topology induced by the metric d. Let $A \subseteq X$. Then $p \in \overline{A}$ if and only if there exists a sequence $(a_n)_{n\in\mathbb{N}} \subseteq A$ that converges to p.*

Proof. Assume that $p \in \overline{A}$. Then $B_{\frac{1}{n}}(p) \cap A \neq \emptyset$ for all $n \in \mathbb{N}$. In particular we can choose a sequence $(a_n)_{n\in\mathbb{N}} \subseteq A$ such that $d(a_n,p) < \frac{1}{n}$.
Now assume that $p \notin \overline{A}$. Since $X \setminus \overline{A}$ is open, there exists an $\varepsilon > 0$ such that $B_\varepsilon(p) \cap A = \emptyset$, i. e.$d(a,x) \geq \varepsilon$ for all $a \in A$. Therefore there exists no sequence in A that converges to p. □

Definition 8.15. Let $(X,\mathcal{O})$ be a topological space and let $A \subseteq X$. A point $p \in X$ is called a *limit point* of A if $U \cap (A \setminus \{p\}) \neq \emptyset$ for every neighbourhood U of p. A set A is called *perfect* if it contains all its limit points.

Definition 8.16. The set A is said to be dense in X if $\overline{A} = X$.

For example, $\mathbb{Q}$ is dense in $\mathbb{R}$.

Recall that a function $f : X \to Y$ between metric spaces is called continuous if and only if for every $\varepsilon > 0$ and $p \in X$ there exists an $\delta > 0$ such that $f(B_\delta(p)) \subseteq B_\varepsilon(f(p))$, that is, the preimage of a neighbourhood of $f(p)$ is a neighbourhood of p.

Definition 8.17. A function $f : X \to Y$ between topological spaces is called *continuous* at $p \in X$ if and only if for every neighbourhood U of $f(p)$ the set $f^{-1}(U)$ is a neighbourhood of p. The function f is called *continuous* if it is continuous in every $p \in X$.

Proposition 8.18. *A function $f : X \to Y$ is continuous if and only if preimages of open sets are open.*

Proof. Assume that f is continuous and let $V \subseteq Y$ be open. Let $p \in f^{-1}(V)$. Then V is a neighbourhood of $f(p)$. Since f is continuous, $f^{-1}(V)$ is a neighbourhood of p, hence it contains an open set U such that $p \in U \subseteq f^{-1}(V)$. By Lemma 8.3 (i) $f^{-1}(V)$ is open.

Now assume that $f^{-1}(V)$ is open for every open set $V \subseteq Y$. Let $p \in X$ and V a neighbourhood of $f(p)$. Then V contains an open neighbourhood V' of $f(p)$ and $f^{-1}(V')$ is open by assumption. Therefore $f^{-1}(V)$ contains a neighbourhood of p which implies that f is continuous in p. □

If X is a topological space, then $\mathrm{id} : X \to X$ is continuous. Compositions of continuous functions are continuous.

Definition 8.19. A *homeomorphism* between two topological spaces X and Y is an bijective function $f : X \to Y$ such that both f and f^{-1} are continuous.

Note that in general the continuity of f does not imply the continuity of f^{-1}. For example, $f : \mathbb{N} \to \mathbb{Q},\ f(x) = x$ is continuous, f^{-1} is not (when $\mathbb{N}$ carries the discrete topology and $\mathbb{Q}$ the topology induced by its metric).

Definition 8.20. Let $(X,\mathcal{O}_X)$ and $(Y,\mathcal{O}_Y)$ be topological spaces. Then we can define a topology $\mathcal{O}$ on $X \times Y$ as follows: A subset $W \subseteq X \times Y$ is called open if and only if it is the union of sets of the form $U \times V$ with U open in X and V open in Y. Obviously, the projections

$$\mathrm{pr}_X : X \times Y \to X, \quad (x, y) \mapsto x, \qquad \mathrm{pr}_Y : X \times Y \to Y, \quad (x, y) \mapsto y,$$

are continuous when $X \times Y$ carries the product topology.

8.2 Compact sets

Definition 8.21. Let X be a topological space and $A \subseteq X$. A family $\mathcal{U} = (U_\lambda)_{\lambda\in\Lambda}$ of open sets in X is called an *open cover* of A if $A \subseteq \bigcup_{\lambda\in\Lambda} U_\lambda$. The open cover $\mathcal{U}$ is called finite if Λ is a finite set. An open cover $\mathcal{U}$ *contains* an open cover $\mathcal{V}$ if every member of $\mathcal{V}$ is also a member of $\mathcal{U}$.

A Hausdorff space X is called *compact* if every open cover of X contains a finite subcover. A subset A of X is called compact if and only if it is compact in the topology induced by X. Obviously, this is the case if and only if every cover of A with open sets in X contains a finite subcover.

Examples. (i) The empty set is compact in every topological space.

(ii) Let X be a Hausdorff space and $M \subseteq X$ a finite subset. Then M is compact.

(iii) The set $M = (0, 1)$ is not compact in $\mathbb{R}$ with the usual topology.

Proof. For $n \in \mathbb{N}$ let $U_n := (\frac{1}{n}, 1)$. Then $(0, 1) \subseteq \bigcup_{n=2}^{\infty} U_n$ and $(U_n)_{n\in\mathbb{N}}$ contains no finite subcover of $(0, 1)$. □

Remark 8.22. A Hausdorff space is compact if and only if the following is true: If $\mathcal{A} = (A_\lambda)_{\lambda\in\Lambda}$ is a family of closed sets such that $\bigcap_{\lambda\in\Lambda} A_\lambda = \emptyset$, then there exists a finite set $\Gamma \subset \Lambda$ such that $\bigcap_{\lambda\in\Gamma} A_\lambda = \emptyset$.

Next we show that all compact sets are closed, and that closed subsets of compact sets are compact.

Theorem 8.23. *Let X be a Hausdorff space and $A \subseteq X$. Then*

(i) *X compact, A closed in X $\implies$ A is compact.*

(ii) *A compact $\implies$ A is closed.*

Proof. (i) Let $\mathcal{A} = (A_\lambda)_{\lambda\in\Lambda}$ a family of closed subsets of A such that $\bigcap_{\lambda\in\Lambda} A_\lambda = \emptyset$. Then every A_λ is also closed in X (see Remark 8.10), therefore there exists a finite set $\Gamma \subset \Lambda$ such that $\bigcap_{\lambda\in\Gamma} A_\lambda = \emptyset$. Hence A is compact by Remark 8.22.

(ii) We show that $X \setminus A$ is open. Let $p \in X \setminus A$. Since X is Hausdorff space, for every $a \in A$ there exist open neighbourhoods U_a of a and V_a of p such that $U_a \cap V_a = \emptyset$. Since A is compact, there exist $a_1, \dots, a_n \in A$ such that $A \subseteq \bigcup_{j=1}^{n} U_{a_j}$. Let $V = \bigcap_{j=1}^{n} V_{a_j}$. Then V is open, $p \in V$ and

$$V \cap A \subseteq V \cap \bigcup_{j=1}^{n} U_{a_j} = \bigcup_{j=1}^{n} V \cap U_{a_j} = \emptyset. \qquad \square$$

Note that the implication (ii) is not necessarily true if X is not Hausdorff. For example, if $X = \{1, 2\}$ with the trivial topology. Then X is compact and the subset $\{1\}$ is compact but not closed.

Definition 8.24. Let (X, d) be a metric space. A set $M \subseteq X$ is called *totally bounded* if for every $\varepsilon > 0$ there exist $x_1, \dots, x_n \in M$ such that $M \subseteq \bigcup_{j=1}^{n} B_\varepsilon(x_j)$.

If M is totally bounded, then M is bounded. The reverse implication is not necessarily true.

Proposition 8.25. *Let (X,d) be a metric space and $A \subseteq X$. Then A is totally bounded if and only if every sequence in A contains a Cauchy subsequence.*

Proof. If $A = \emptyset$, then the assertion is clear.

Assume that A is totally bounded and let $x = (x_n)_{n\in\mathbb{N}} \subseteq A$. Since A is totally bounded there exist $q_1, \dots q_n \in X$ such that $A \subseteq \bigcup_{j=1}^n B_1(q_j)$. Therefore there exists in $j \in \{1, \dots n\}$ and a subsequence $x_1 = (x_{1,n})_{n\in\mathbb{N}}$ of x such that $x_1 \subseteq B_1(p_1)$ (with $p_1 = q_j$). By the same argument, applied to the sequence x_1, there exists a $p_2 \in A$ and a subsequence x_2 of x_1 such that $x_2 \subseteq B_{\frac{1}{2}}(p_2)$. We can now choose inductively a sequence $(p_j)_{j\in\mathbb{N}} \subseteq X$ and a sequence of subsequences $x_k = (x_{k,n})_{n\in\mathbb{N}}$ of x such that p_1 and x_1 as chosen above, $x_k \subseteq B_{\frac{1}{k}}(p_k)$ and x_{k+1} is a subsequence of x_k for all $k \in \mathbb{N}$. Then the sequence $(x_{n,n})_{n\in\mathbb{N}}$ is a subsequence of x and, by construction, it is a Cauchy sequence.

Now assume that A is not totally bounded. We will construct a sequence $(x_n)_{n\in\mathbb{N}} \subseteq A$ that contains no Cauchy sequence. Since A is not totally bounded there exists an $\varepsilon > 0$ such that $A \not\subseteq \bigcup_{j=1}^k B_\varepsilon(p_j)$ for every finite sequence $(p_j)_{j=1}^k \subseteq A$. Choose x_1 arbitrary in A. By assumption on A, we can choose $x_k \in A$ inductively such that $x_{k+1} \notin \bigcup_{j=1}^k B_\varepsilon(x_j)$ and $A \not\subseteq \bigcup_{j=1}^{k+1} B_\varepsilon(x_j)$. By construction $d(x_m, x_n) \geq \varepsilon$ for all $m, n \in \mathbb{N}$, hence $(x_n)_{n\in\mathbb{N}}$ does not contain a Cauchy sequence. □

Definition 8.26. A topological space X is called *sequentially compact* if every sequence in X contains a convergent subsequence.

The following is a generalisation of the Bolzano-Weierstraß theorem (Theorem 4.40). Note that it is true in an arbitrary topological space (not necessarily a metric space).

Theorem 8.27. *Every compact metric space is sequentially compact.*

Proof. Let X be a compact metric space and $x = (x_n)_{n\in\mathbb{N}}$ be a sequence in X. Assume that x does not contain an convergent subsequence. Then for every $\varepsilon > 0$ and every $y \in X$ the open ball $B_\varepsilon(y)$ contains only finitely many members of the sequence x. Since X is compact there exist $y_1, \dots, y_n$ such that $x \subseteq X \subseteq \bigcup_{j=1}^n B_\varepsilon(y_j)$ implying that x has only finitely many members. □

Corollary 8.28. *Let (X,d) be a metric space and $A \subseteq X$ compact. Then A is closed and totally bounded.*

In a metric space, the reverse of Theorem 8.27 is true.

Theorem 8.29. *Let (X,d) be a metric space and $A \subseteq X$ sequentially compact. Then A is compact.*

Proof. Since A is sequentially compact, it is totally bounded by Proposition 8.25. Now let $\mathcal{U} = (U_\lambda)\lambda \in \Lambda$ be an open cover of A. We have to show that $\mathcal{U}$ contains a finite cover of A. First we show the existence of a $\delta > 0$ such that for every $y \in X$ the ball $B_\delta(y)$ is contained in a U_λ.
Assume that no such δ exists. Then there exists a sequence $y = (y_n)_{n\in\mathbb{N}}$ such that there exists no $\lambda \in \Lambda$ with $B_{\frac{1}{n}}(y_n) \subseteq U_\lambda$. Since A is sequentially compact, y contains a convergent subsequence; without restriction we can assume that y itself converges to some $y_0 \in A$. Since $\mathcal{U}$ is an open cover of A, we can choose $\delta > 0$ and $\lambda_0 \in \Lambda$ such that $y_0 \in B_\delta(y_0) \subseteq U_{\lambda_0}$. Since the sequence y converges to y_0, we can choose n large enough such that $d(y_n, y_) < \frac{\delta}{2}$ and $\frac{1}{n} < \frac{\delta}{2}$, see Figure 8.3. It follows that $B_{\frac{1}{n}}(y_n) \subseteq B_\delta(y_0) \subseteq U_{\lambda_0}$. □

Corollary 8.30. *Every interval of the form $[a,b]$ is compact in $\mathbb{R}$ since it is sequentially compact by the Bolzano-Weierstraß theorem (Theorem 4.40).*

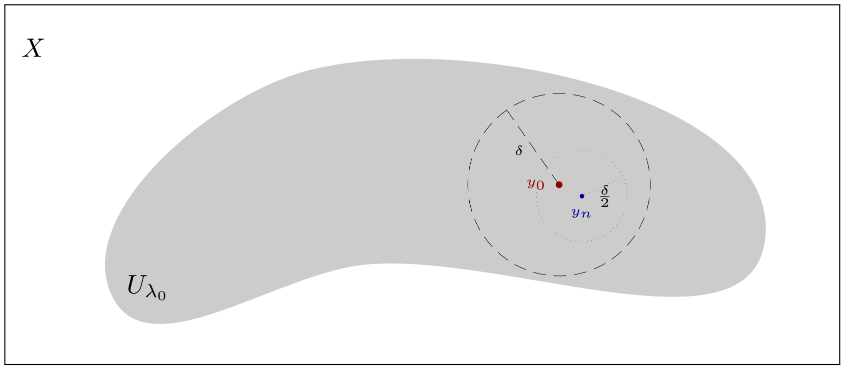


FIGURE 8.3: A sequentially compact metric space is compact (Theorem 8.29).

For the proof of the Heine-Borel theorem (Theorem 8.33) we use the following two auxiliary lemmata.

Lemma 8.31 (Tube lemma). *Let K be compact space, X a topological space, $p \in X$, $U \subseteq X \times K$ open in the product topology such that $\{p\} \times K \subseteq U$. Then there exists an open set $V \subseteq X$ such that $V \times K \subseteq U$.*

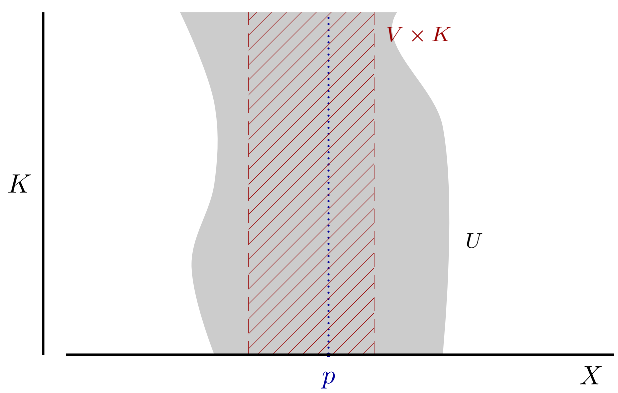


FIGURE 8.4: Tube lemma (Lemma 8.31).

Proof. Let $k \in K$. Then there exist an open neighbourhood W_k of k and an open neighbourhood V_k of p such that $(p,k) \in V_k \times W_k \subseteq U$. Since K is compact, there are $k_1, \dots, k_n \in K$ such that $K \subseteq \bigcup_{j=1}^{n} W_{k_j}$. Let $V = \bigcap_{j=1}^{n} V_{k_j}$. Then $V \times K \subseteq \bigcup_{j=1}^{n} (V_{k_j} \times W_{k_j}) \subseteq U$. □

Lemma 8.32. (i) *The product of finitely many Hausdorff spaces is again a Hausdorff space.*

(ii) *The product of finitely many compact spaces is again a compact space.*

Proof. It suffices to show the assertion for two topological spaces X and Y.
(i) Assume that X and Y are Hausdorff spaces and let $(x_1, y_1) \neq (x_2, y_2) \subseteq X \times Y$. If $x_1 \neq x_2$ then there are disjoint open neighbourhoods U_1 of x_1 and U_2 of x_2. Therefore $U_1 \times Y$ and $U_2 \times Y$ are disjoint neighbourhoods of (x_1, y_1) and (x_2, y_2). If $x_1 = x_2$, then $y_1 \neq y_2$ and as before we can find disjoint neighbourhoods of (x_1, y_1) and (x_2, y_2).

(ii) Let $\mathcal{U} = (U_\lambda)_{\lambda\in\Lambda}$ be an open cover of $X \times Y$. Let $p \in X$. Since $\{p\} \times Y$ is compact for every $p \in Y$, there exist $\lambda_1, \dots, \lambda_n$ such that $\{p\} \times Y \subseteq \bigcup_{j=1}^n U_{\lambda_j}$. By the tube lemma there exist an open neighbourhood V_p of p such that $V_p \times Y \subseteq \bigcup_{j=1}^n U_{\lambda_j}$. Since X is compact, it can be covered by finitely many such V_p, hence $\mathcal{U}$ contains a finite subcover of $X \times Y$. □

Theorem 8.33 (Heine-Borel theorem). *A subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed.*

Proof. Let $A \subseteq \mathbb{R}$. If A is compact, then it is bounded and closed by Theorem 8.23. Now assume that A is bounded and closed. Since A is bounded, it lies in a closed cube $C = [a_1, b_1] \times \dots \times [a_n, b_n]$. By Lemma 8.32, C is compact, hence A is compact by Theorem 8.23. □

Note that this equivalence is not true for an arbitrary metric space (X, d) because the bounded metric $d' := \min\{1, d\}$ induces the same topology on X. Hence boundedness says nothing about compactness. For arbitrary metric spaces we have the following characterisation:

Theorem. *Let (X, d) be a metric space. A subset A of X is compact if and only if it is complete and totally bounded.*

In the rest of this section we prove theorems for continuous functions on compact sets.

Theorem 8.34. *Let X a compact space, Y a topological space and $f : K \to Y$ a continuous function. Then $f(X)$ is compact.*

Proof. Let $\mathcal{U} = (U_\lambda)_{\lambda\in\Lambda}$ be an open cover of $f(X)$. Then $(f^{-1}(U_\lambda))_{\lambda\in\Lambda}$ is an open cover of X. Therefore there exist a finite subset $\Gamma \subseteq \Lambda$ such that $(f^{-1}(U_\lambda))_{\lambda\in\Gamma}$ is an open cover of X. Hence $(U_\lambda)_{\lambda\in\Gamma}$ be an open cover of $f(X)$ subordinate to $\mathcal{U}$. □

As a corollary we obtain the following theorem.

Theorem 8.35. *Let X a compact space, Y a Hausdorff space and $f : X \to Y$ a continuous bijection. Then f^{-1} is continuous.*

Proof. By Theorem 8.34 and Theorem 8.23 (ii) for every closed set A in X the set $f(A)$ is closed in Y. □

Note the we proved this theorem for intervals in Theorem 5.27.

Theorem 8.36. *Let X be a non-empty compact metric space and $f : X \to \mathbb{R}$ a continuous function. Then f is bounded and has a maximum and a minimum.*

Proof. By Theorem 8.34, $f(X)$ is a compact subset in $\mathbb{R}$, hence it is bounded and closed. Let $M := \sup\{f(x) : x \in K\}$. Then there exists a sequence $x = (x_n)_{n\in\mathbb{N}} \subseteq K$ such that $f(x_n) \to M$ for $n \to \infty$. Since K is compact, x contains a convergent subsequence. Without restriction we can assume that x converges to a $x_0 \in K$. By continuity of f we obtain $M = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty}(x_n)) = f(x_0)$, hence the range of f has a maximum. Analogously we can show that $\mathrm{R}(f)$ has a minimum. □

Theorem 8.37. *Let X, Y be metric spaces, $K \subseteq X$ compact and $f : X \to Y$ continuous in every point of K. Then f is uniformly continuous on K.*

Proof. Let $\varepsilon > 0$. We have to show the existence of a $\delta > 0$ such that for all $x \in X$ and $y \in K$

$$d(x,y) < \delta \qquad \Longrightarrow \qquad d(f(x), f(y)) < \varepsilon.$$

Continuity of f in K implies that for every $y \in K$ there exists a $\delta(y) > 0$ such that

$$d(x,y) < 2\delta(y) \qquad \Longrightarrow \qquad d(f(x), f(y)) < \varepsilon.$$

Since K is compact, there exist $y_1, \dots, y_n \in K$ such that $(B_{\delta(y_j)})_{j=1}^n$ is an open cover of K. Let $\delta = \min\{\delta(y_j) : j = 1, \dots n\}$. If $x \in X$ such that $d(x,y) < \delta$ for some $y \in K$, then there exists a j such that $d(x, y_j) \le d(x,y) + d(y, y_j) \le 2\delta_j$. Therefore $d(f(x), f(y)) \le d(f(x), f(y_j)) + d(f(y_j), f(y)) \le 2\varepsilon$. □

Theorem 8.38 (Stone-Weierstraß). *Let K be a compact Hausdorff space and $C(K)$ the space of all real or complex valued functions on K together with the supremum norm $\|\cdot\|$. Let $F \subseteq C(K)$ such that*

(i) *F contains a constant function not equal to 0;*

(ii) *F separates the points in K, i. e., for all x_1, $x_2 \in K$ exists an $f \in F$ such that $f(x_1) \neq f(x_2)$;*

(iii) *if $f \in \mathbb{F}$, then also $\overline{f} \in F$ ($\overline{f}$ denotes the complex conjugate of f, defined by $\overline{f}(x) = \overline{f(x)}$, $x \in K$).*

Then the algebra generated by F is dense in $C(K)$.

8.3 Connected sets

Definition 8.39. A topological space X is called *connected* if it is not the disjoint union of two non-empty open sets.

Equivalent formulations are:

(i) A topological space X if it is not the disjoint union of two non-empty closed sets.

(ii) X does not contain a set that is open and closed.

(iii) If $A, B \subseteq X$ are open, $A \neq \emptyset$, $X = A \cup B$, then $B = \emptyset$.

Theorem 8.40. *Let $D \subseteq \mathbb{R}$, $D \neq \emptyset$. Then*

$$D \text{ is connected} \qquad \Longleftrightarrow \qquad D \text{ is an interval.}$$

Proof. "$\Longrightarrow$" Assume that D is not an interval. Then there exist $a < x < b$ such that $a, b \in D$ and $x \notin D$. Then $D_a := D \cap (-\infty, x)$ and $D_b := D \cap (x, \infty)$ are open in D and D is the disjoint union of D_a and D_b, therefore D is not connected.

"$\Longleftarrow$" Let D be an interval and $A, B \subseteq D$ open in D such that D is the disjoint union of A and B. Assume that $A \neq \emptyset$. We have to show that B is empty. Let

$$f : D \to \mathbb{R}, \qquad f(x) = \begin{cases} 1, & x \in A, \\ 0, & x \in B. \end{cases}$$

Obviously, f is continuous. If $B \neq \emptyset$, then $\{0, 1\}$ lies in the range of f. By the intermediate value theorem (Theorem 5.24) there exists a $p \in D$ such that $f(p) = \frac{1}{2}$, in contradiction to the definition of f. □

Theorem 8.41. *Let X, Y be topological spaces and $f : X \to Y$ a continuous function. If X is connected, then $f(X)$ is connected.*

Proof. Let $U \neq \emptyset$ be subseteq of $f(X)$ such that U is open and closed. We have to show that $U = f(X)$. By assumption on U there exists an open set V and a closed set A in Y such that $U = f(X) \cap V$ and $A = f(X) \cap A$. By continuity of f it follows that

$$\emptyset \neq f^{-1}(U) = \underbrace{f^{-1}(V)}_{\text{open in } X} = \underbrace{f^{-1}(A)}_{\text{closed in } X}.$$

Since X is connected, it follows that $f^{-1}(U) = X$, hence $U = f(X)$. □

As a corollary we obtain the generalised intermediate value theorem:

Theorem 8.42 (Intermediate value theorem). *Let X be a connected topological space and $f : X \to \mathbb{R}$ a continuous function. Then $f(X)$ is an interval.*

Definition 8.43. A topological space X is called *arcwise connected* if for all $x, y \in X$ there exists a continuous function $f : [0, 1] \to X$ such that $f(0) = x$, $f(1) = y$.

Theorem 8.44. *An arcwise connected space is connected.*

Proof. Let X be a arcwise connected topological space. Assume that X is no connected. Then there exist U, V non-empty open subsets of X. Let $p \in U$ and $q \in V$ and choose a continuous function $f : [0, 1] \to X$ such that $f(0) = p$ and $f(1) = q$. Since f is continuous, the sets $U' = f^{-1}(U) \subseteq [0, 1]$ and $V' = f^{-1}(V) \subseteq [0, 1]$ are open. Define the function

$$g : [0, 1] \to \mathbb{R}, \qquad g(x) = 0 \iff x \in U', \qquad g(x) = 1 \iff x \in V'.$$

Then g is continuous because all preimages under g are open. Since $g(0) = 0$ and $g(1) = 1$, the intermediate value theorem implies that there exists an $t \in [0, 1]$ such that $g(t) = \frac{1}{2}$, in contradiction to the definition of g. □

Corollary 8.45. *Let $(V, \|\cdot\|)$ be a metric space, $\Omega \subseteq V$ a convex set. Then Ω is arcwise connected, in particular, it is connected.*

Note that there are connected spaces that are not arcwise connected.

Chapter 9

Exercises

Exercises for Chapter 2

1. For sets A, B and C show at least two of the following statements:
   (a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
   (b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
   (c) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$,
   (d) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

2. (a) Find the power sets of (a) $L = \emptyset$, (b) $M = \{0\}$, (c) $N = \{1, 2, 3\}$.
   (b) Let $N = \{1, 2, 3\}$ and consider the relation $\subseteq$ on $\mathbb{P}N$. Is $\subseteq$ reflexive, transitive, symmetric? Does $\subseteq$ define a total order on $\mathbb{P}N$?

3. (a) For sets A, B and C show:
   - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
   - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

   (b) For sets $A, B \subset X$ show:
   - $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$,
   - $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

4. Let X, Y and Z be sets and $f : X \to Y$, $g : Y \to Z$ functions.
   Show:
   (a) If g is injective, then
   $$f \text{ is injective} \iff g \circ f \text{ injective.}$$
   (b) If f is surjective, then
   $$g \text{ surjective} \iff g \circ f \text{ surjective.}$$

5. (a) Show that the countable subset of a countable set is countable or finite.
   (b) Show that the countable union of countable sets is countable.
   (c) Show that the direct product of countable sets is countable.
   (d) Find a bijection $\mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$.

(e) Show that $\mathbb{Q}$ is countable.

6. (a) Show that the power set $\mathbb{P}\mathbb{N}$ is not countable.
   (b) Let A, B be sets. Show or give a counterexample:
      (i) $\mathbb{P}(A \cap B) = \mathbb{P}A \cap \mathbb{P}B$.
      (ii) $\mathbb{P}(A \cup B) = \mathbb{P}A \cup \mathbb{P}B$.

7. Show the following formulas:
   (a) $\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2, \quad n \in \mathbb{N},$
   (b) $\sum_{k=1}^{2n} (-1)^{k+1} \frac{1}{k} = \sum_{k=1}^{n} \frac{1}{n+k}\ , \quad n \in \mathbb{N}.$

8. For $n \in \mathbb{N}_0$ y $m \in \mathbb{N}$ define

$$a(m, n) := \#\{(x_1, \ldots, x_m) \in \mathbb{N}_0^m \ : \ \sum_{j=1}^{m} x_j \le n\},$$

$$b(m, n) := \#\{(x_1, \ldots, x_m) \in \mathbb{N}_0^m \ : \ \sum_{j=1}^{m} x_j = n\}.$$

   (a) Show that $a(m, n) = b(m+1, n)$, $m \in \mathbb{N}$, $n \in \mathbb{N}_0$.
   (b) Show that $a(m, n) = \binom{n+m}{m}$, $m \in \mathbb{N}$, $n \in \mathbb{N}_0$.

   Hint: Show that $a(m, n-1) + a(m-1, n) = a(m, n)$ and use induction on $n + m$.

Exercises for Chapter 3

1

1. Let $(K, +, \cdot, >)$ be an ordered field and $a, x, x', y, y' \in K$. Show the following statements from Corollary 3.9. Justify every step.

   (iii) $x < y \implies x + a < y + a,$
   (iv) $x < y \ \wedge \ x' < y' \implies x + x' < y + y',$
   (v) $x < y \ \wedge \ a > 0 \implies a \cdot x < a \cdot y,$
   $x < y \ \wedge \ a < 0 \implies a \cdot x > a \cdot y,$
   (vi) $0 \le x < y \ \wedge \ 0 \le x' < y' \implies 0 \le x' \cdot x < y' \cdot y,$
   (ix) $x > 0 \implies x^{-1} > 0,$
   (x) $0 < x < y \implies 0 < y^{-1} < x^{-1},$
   (xi) $x > 0 \wedge y < 0 \implies xy < 0.$

2. Find the infimum and supremum of the following sets in the ordered field $\mathbb{R}$. Determine if they have a maximum and a minimum.

   (a) $\{x \in \mathbb{R} : \exists n \in \mathbb{N} \ \ x = n^2\},$

(b) $\left\{\frac{|x|}{1+|x|} : x \in \mathbb{R}\right\}$,

(c) $\left\{x \in \mathbb{R} : \exists\, n \in \mathbb{N} \;\; x = \frac{1}{n} + n\big(1 + (-1)^n\big)\right\}$,

(d) $\left\{x \in \mathbb{R} : x^2 \le 2\right\} \cap \mathbb{Q}$.

3. (a) For every $x \in \mathbb{R}_+$ there exists an $n \in \mathbb{N}_0$ with $n \le x < n+1$. (Proposition 3.19).

(b) Every interval in $\mathbb{R}$ contains a rational number. (Proposition 3.20).

(c) $\mathbb{Q}$ does not have the least upper bound property.

4. (a) Let $X \subset \mathbb{R}$, $X \neq \emptyset$, and $\xi \in \mathbb{R}$ an upper bound of X. Show that

$$\xi = \sup X \qquad \Longleftrightarrow \qquad \forall \varepsilon \in \mathbb{R}_+ \; \exists\, x_\varepsilon \in X \;\; \xi - \varepsilon < x_\varepsilon \le \xi.$$

What is the analogous statement for $\inf X$?

(b) Let $X,\ Y \subset \mathbb{R}$ non empty sets such that

$$\forall\, x \in X \;\; \exists\, y \in Y \;:\; y < x.$$

Does that imply $\inf Y < \inf X$? Proof your assertion.

5. (a) Muestre que para todo $z \in \mathbb{C} \setminus \{0\}$ existen exactamente dos números $\zeta_1, \zeta_2 \in \mathbb{C}$ tal que $\zeta_1^2 = \zeta_2^2 = z$.

(b) Sean $a, b, c \in \mathbb{C}$, $a \neq 0$. Muestre que existe por lo menos un $z \in \mathbb{C}$ tal que

$$az^2 + bz + c = 0.$$

Exercises for Chapter 4

1

1. (a) Let (X, d), $X \neq \emptyset$, be a metric space and $M \subseteq X$. Show that the following are equivalent:

(i) M is bounded.

(ii) $\exists\, x \in X \; \exists\, r > 0 : M \subseteq B_r(x)$.

(iii) $\forall\, x \in X \; \exists\, r > 0 : M \subseteq B_r(x)$.

(b) For $M \subseteq \mathbb{R}$ show that M is bounded as subset of the ordered field $(\mathbb{R}, >)$ if and only if M is bounded as subset of the metric space $(\mathbb{R}, d)$ where $d(x, y) = |x - y|$.

2. (a) Let (X, d), $X \neq \emptyset$, be a metric space and let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be sequences in X. Show: If there exists an $a \in X$ such that

$$\lim_{n\to\infty} x_n = a = \lim_{n\to\infty} y_n,$$

then

$$\lim_{n\to\infty} d(x_n, y_n) = 0.$$

Is the converse also true (proof or counterexample)?

(b) Let (X, d) be a metric space and $\rho : \mathbb{N} \to \mathbb{N}$ a bijection. Show: If $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges, then $(x_{\rho(n)})_{n\in\mathbb{N}} \subseteq X$ converges and has the same limit.

3. (a) Let $x_n = \sqrt{1+n^{-1}}$, $n \in \mathbb{N}$. Show that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$.
   (b) Do the following sequences in $\mathbb{R}$ converge? If so, find the limit. Prove your assertions.
      (i) $(a_n)_{n\in\mathbb{N}}$ with $a_n = \dfrac{n}{2^n}, \quad n \in \mathbb{N},$
      (ii) $(a_n)_{n\in\mathbb{N}}$ with $a_n = \dfrac{2^n}{n!}, \quad n \in \mathbb{N},$
      (iii) $(b_n)_{n\in\mathbb{N}}$ with $b_n = \sqrt{1+n^{-1}+n^{-2}}, \quad n \in \mathbb{N},$
      (iv) $(d_n)_{n\in\mathbb{N}}$ with $d_n = \sqrt{n^2+n+1} - n, \quad n \in \mathbb{N},$

4. Let $q \in \mathbb{R}_+$ and $x_n := \sqrt[n]{q}$, $y_n := \sqrt[n]{n}$, $n \in \mathbb{N}$. Do the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge? If so, find the limit.

5. Let $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence such that $a_n \neq 0$ for all $n \in \mathbb{N}$. Show or find a counterexample:

   (i) If there exists an $N \in \mathbb{N}$ and $q \in \mathbb{R}$, $q < 1$, such that
   $$\left|\frac{a_{n+1}}{a_n}\right| \leq q, \quad n \in \mathbb{N},\ n \geq N,$$
   then $\lim_{n\to\infty} a_n = 0$.

   (ii) If there exists an $N \in \mathbb{N}$ and $q \in \mathbb{R}$, $q \leq 1$, such that
   $$\left|\frac{a_{n+1}}{a_n}\right| < q, \quad n \in \mathbb{N},\ n \geq N,$$
   then $\lim_{n\to\infty} a_n = 0$.

6. The Fibonacci sequence $(a_n)_{n\in\mathbb{N}}$ is defined recursively by
$$a_0 = 1, \quad a_1 = 1, \quad a_{n+1} = a_n + a_{n-1}, \qquad n \in \mathbb{N}.$$
   Moreover, let $\sigma < \tau$ be the solutions of $x^2 - x - 1 = 0$ and
$$x_n = \frac{a_{n+1}}{a_n}, \qquad n \in \mathbb{N}.$$
   (a) Show that $(a_n)_{n\in\mathbb{N}}$ does not converge in $\mathbb{R}$.
   (b) $a_n = \dfrac{1}{\sqrt{5}}(\tau^{n+1} - \sigma^{n+1}), \quad n \in \mathbb{N}.$
   (c) $\lim_{n\to\infty} x_n = \sigma$.

7. If it exists, find the value of
$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \dots}}},$$
   i. e. the limit of the sequence $(x_n)_{n\in\mathbb{N}}$ with
$$x_1 := 1 \quad \text{and} \quad x_{n+1} := 1 + \frac{1}{x_n},\ n \geq 1.$$

8. (a) Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and define sequences $(y_k)_{k\in\mathbb{N}}$, $(z_k)_{k\in\mathbb{N}}$ in $\mathbb{R} \cup \{\pm\infty\}$ by
$$y_k := \sup\{x_n : n \geq k\}, \qquad z_k := \inf\{x_n : n \geq k\}, \qquad k \in \mathbb{N}.$$

Show that $(y_k)_{k\in\mathbb{N}}$ and $(z_k)_{k\in\mathbb{N}}$ converge in $\mathbb{R}\cup\{\pm\infty\}$ and that

$$\lim_{k\to\infty} y_k = \limsup x_n, \qquad \lim_{k\to\infty} z_k = \liminf x_n.$$

(b) Find a sequence $(a_n)_{n\in\mathbb{N}}$ such that

$$\inf\{a_n : n\in\mathbb{N}\} < \liminf\{a_n : n\in\mathbb{N}\} < \limsup\{a_n : n\in\mathbb{N}\} < \sup\{a_n : n\in\mathbb{N}\}.$$

In this case, must the set $\{a_n : n\in\mathbb{N}\}$ have a maximum?

9. Let $\mathbb{K}$ be an ordered field with the Archimedean property. Show that $\mathbb{K}$ has the least upper bound property if and only if every Cauchy sequence converges.

10. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in a normed space and $(b_n)_{n\in\mathbb{N}}$ be defined by

$$b_n := \frac{1}{n}\sum_{k=1}^{n} a_k.$$

Show or find a counterexample:

(a) $(a_n)_{n\in\mathbb{N}}$ converges $\implies$ $(b_n)_{n\in\mathbb{N}}$ converges.
(b) $(b_n)_{n\in\mathbb{N}}$ converges $\implies$ $(a_n)_{n\in\mathbb{N}}$ converges.

11. Cauchy's condensation test.

(a) For a monotonically decreasing sequence $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+^0$ show

$$\sum_{n\in\mathbb{N}} a_n \text{ converges} \iff \sum_{n\in\mathbb{N}} 2^n a_{2^n} \text{ converges.}$$

(b) Do the series $\displaystyle\sum_{n=1}^{\infty}(n\log_2 n)^{-1}$ and $\displaystyle\sum_{n=1}^{\infty}\frac{\sqrt{n+1}-\sqrt{n}}{n}$ converge? Prove your answer. (Use what you know from the calculus courses about the logarithm.)

12. The Euler number e**.**

For $n\in\mathbb{N}$ let $a_n := \left(1+\frac{1}{n}\right)^n$ and $s_n := \sum_{k=0}^{\infty}\frac{1}{k!}$.

(a) Show that $2^k < k!$ for all $k\geq 4$ and that

$$1 \leq \left(1+\frac{1}{n}\right)^n \leq \sum_{k=0}^{n}\frac{1}{k!} < 3, \qquad n\in\mathbb{N}.$$

(b) Show that the sequences $(a_n)_{n\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ converge.
(c) Show that

$$\mathrm{e} := \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{\infty}\frac{1}{k!}, \qquad n\in\mathbb{N}.$$

13. Find the 5-adic and 7-adic representation of $\frac{61}{5}$. Proof!
That is, find $N_a, N_b \in \mathbb{Z}$ and $(a_n)_{n=-N_a}^{\infty} \subseteq \{0, 1, \ldots, 4\}$ and $(b_n)_{n=-N_b}^{\infty} \subseteq \{0, 1, \ldots, 6\}$ and such that

$$\frac{61}{5} = \sum_{n=-N_a}^{\infty} a_n 5^{-n} = \sum_{n=-N_b}^{\infty} b_n 7^{-n}.$$

14. Do the following series converge? Proof your answer.

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

(c) $\displaystyle\sum_{n=2}^{\infty} b_n$, where $b_{2m} := \frac{1}{(2m)^2}$, $b_{2m+1} = -\frac{1}{2m}$,

(d) $\displaystyle\sum_{n=1}^{\infty} \Big(a + \frac{1}{n}\Big)^n$ where $a \in \mathbb{R}$.

15. (a) For $n \in \mathbb{N}$ let $a_n := b_n := \frac{(-1)^n}{\sqrt{n+1}}$ and $c_n := \sum_{k=0}^{n} a_k b_{n-k}$. Show that $\sum_{n=0}^{\infty} a_n$ converges, but $\sum_{n=0}^{\infty} c_n$ diverges.

(b) Let $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ a monotonically decreasing sequence such that $\sum_{n=1}^{\infty} a_n$ converges in $\mathbb{R}$. Show that

$$\lim_{n\to\infty} n\, a_n = 0.$$

Exercises for Chapter 5

1

1. For $j = 1, \ldots, n$ let $(X_j, \|\cdot\|_j)$ be normed spaces over $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Recall that $(X_1 \times \cdots \times X_n, \|\cdot\|)$ with

$$\|(x_1, \ldots, x_n)\| := \|x_1\| + \cdots + \|x_n\|$$

is a normed space over $\mathbb{F}$.

(a) Show that for all $j = 1, \ldots, n$ the projection pr_j is continuous where

$$\mathrm{pr}_j : X_1 \times \cdots \times X_n \to X_j, \quad (x_1, \ldots, x_n) \mapsto x_j.$$

(b) Let $f = (f_1, \ldots, f_n) : V \to X_1 \times \cdots X_n$ where V is a normed space (that is $f_j : V \to X_j$ and $f(v) = (f_1(v), \ldots, f_n(v))$).
Show that f is continuous if and only if every f_j is continuous.

(c) Let X be a normed space, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $f : \mathcal{D}_f \to \mathbb{F}$, $g : \mathcal{D}_g \to \mathbb{F}$ continuous. Let $\mathcal{D}_{fg} = \mathcal{D}_f \cap \mathcal{D}_g$. Then $fg : \mathcal{D}_{fg} \to \mathbb{F}$, $(fg)(x) = f(x)g(x)$ is continuous. If $g(x) \neq 0$, $x \in \mathcal{D}_{fg}$, then $\frac{f}{g} : \mathcal{D}_{fg} \to \mathbb{F}$, $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$ is continuous.

2. Proof the Cauchy criterion (Theorem 5.15):
Let (X, d_X), (Y, d_Y) be metric spaces, Y complete, $f : X \supseteq \mathcal{D} \to Y$ a function and x_0 a limit point of $\mathcal{D}$. Then f has a limit in x_0 if and only if

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x, y \in D_f : \Big(0 < d_X(x, x_0) < \delta\ \wedge\ 0 < d_X(y, x_0) < \delta \implies d_Y\big(f(x), f(y)\big) < \varepsilon\Big).$$

3. Let (X, d) be a metric space and $f, g : X \to \mathbb{R}$ continuous functions. Show that the following functions are continuous:

$$S : X \to \mathbb{R}, \quad S(x) := \min\{f(x),\ g(x)\},$$
$$T : X \to \mathbb{R}, \quad T(x) := \max\{f(x),\ g(x)\}.$$

4. Where are the following functions are continuous? Proof your answer.

(a) $f : [0, \infty) \to \mathbb{R}, \quad x \mapsto \sqrt{x}$,

(b) $g : \mathbb{C} \to \mathbb{R}, \quad z \mapsto |z + \bar{z}^2|$,

(c) $h : [-1, 1] \cup \{2\} \to \mathbb{R}, \quad x \mapsto \begin{cases} -\sqrt{-x}, & -1 \le x \le 0, \\ \sqrt{x}, & 0 < x \le 1, \\ x, & x = 2. \end{cases}$

(d) $D : \mathbb{R} \to \mathbb{R}, \quad D(x) := \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$

Hint: Show that $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$, that is, for every $x \in \mathbb{R}$ there exists a sequence $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R} \setminus \mathbb{Q}$ such that $\lim_{n\to\infty} x_n = x$.

5. Prove Theorem 5.26 and Theorem 5.27:

Let $I = (a, b)$ a nonempty real interval and $f : I \to \mathbb{R}$ a function.

(a) Assume that f is continuous. Then f is injective if and only if f is strictly monotonic.

(b) If f is strictly monotonically increasing or decreasing, then it is invertible and its inverse $f^{-1} : f(I) \to \mathbb{R}$ is continuous.

6. Show that $f : [0, \infty) \to \mathbb{R}$, $x \mapsto \sqrt{x}$, is uniformly continuous but not Lipschitz continuous.

7. Do the following sequences of functions converge pointwise? Do they converge uniformely? If they converge, find the limit function.

(a) $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \begin{cases} n^2 x, & 0 \le x \le \frac{1}{n}, \\ 2n - n^2 x, & \frac{1}{n} < x \le \frac{2}{n}, \\ 0, & \text{else.} \end{cases}$

(b) $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \dfrac{nx}{1 + nx^2}$,

(c) $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \dfrac{nx}{1 + n^2 x^2}$,

(d) $f_n : \mathbb{R} \to \mathbb{R}, \quad f_n(x) = \dfrac{n^2 x}{1 + nx}$.

8. Let $\mathcal{D} \subseteq \mathbb{R}$, $f : \mathcal{D} \to \mathbb{R}$ a function and $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \setminus \{0\}$ a sequence that converges to 0. Define $f_n : \mathcal{D} \to \mathbb{R}$ by $f_n(x) = a_n f(x)$, $x \in \mathcal{D}$.

(a) $(f_n)_{n \in \mathbb{N}}$ converges pointwise to $g : \mathcal{D} \to \mathbb{R}$, $g(x) = 0$.

(b) $(f_n)_{n \in \mathbb{N}}$ converges uniformely if and only if f is bounded on $\mathcal{D}$.

9. Find the radius of convergence of

i) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n (2z)^n}{n}$, ii) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n} z^n$,

iii) $\displaystyle\sum_{n=1}^{\infty} (\sqrt{n} - 1)^n z^n$, iv) $\displaystyle\sum_{n=1}^{\infty} \frac{8^n z^{3n}}{3^n}$.

10. Show the following properties of the exponential function (Theorem 5.50):

(a) $\exp(\bar{z}) = \overline{\exp(z)}, \quad z \in \mathbb{C}$,
(b) $\exp(z+w) = \exp(z)\exp(w), \quad z, w \in \mathbb{C}$,
(c) $\exp(n) = \mathrm{e}^n, \quad n \in \mathbb{Z}$,
(d) $\exp(z) \neq 0, \quad z \in \mathbb{C}$,
(e) $|\exp(\mathrm{i}x)| = 1 \iff x \in \mathbb{R}$.

11. (a) Show the following identities for $x, y \in \mathbb{C}$:
(i) $\sin^2(x) + \cos^2(x) = 1$.
(ii) $\sin(x+y) = \cos(x)\sin(y) + \cos(y)\sin(x)$,
(iii) $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$,

(b) Show that $\{x \in \mathbb{R}_+ : \cos x = 0\} \neq \emptyset$.

Let $\pi := 2 \cdot \inf\{x \in \mathbb{R}_+ : \cos x = 0\}$.

(c) For $x \in \mathbb{R}$ show:
(i) $\sin x = 0 \iff \exists\, k \in \mathbb{Z} \quad x = k\pi$.
(ii) $\cos x = 0 \iff \exists\, k \in \mathbb{Z} \quad x = k\pi + \frac{\pi}{2}$.

Hint. Without proof you can use

$$1 - \frac{x^2}{2} \leq \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}, \qquad x \in (0, 3].$$

Exercises for Chapter 6

1

1. Show that the following functions are differentiable and find the derivative. Prove your assertions.

(a) $w : \mathbb{R}_+ \to \mathbb{R},\ x \mapsto \sqrt{x}$,
(b) $w : \mathbb{R} \to \mathbb{R},\ x \mapsto \sqrt{|x|}$,
(c) $\cos : \mathbb{R} \to \mathbb{R}, \quad \sin : \mathbb{R} \to \mathbb{R}$,

Hint. Prove Euler's formula (Theorem 5.50): $\exp(\mathrm{i}z) = \cos(z) + \mathrm{i}\sin(z)$ for $z \in \mathbb{C}$.

2. For $k \in \mathbb{N}$ let $f_k : \mathbb{R} \to \mathbb{R}$ defined by

$$f_k(x) := \begin{cases} x^k \sin \dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

For which k is f_k differentiable? For which k is f_k continuously differentiable?

3. **Exponential functions.** For fixed $a \in \mathbb{R}^+ = (0, \infty)$ define the function

$$p_a : \mathbb{C} \to \mathbb{C}, \qquad p_a(z) = \exp(z \ln(a)).$$

(a) For $a \in \mathbb{R}^+$ and $q \in \mathbb{Q}$ show

$$p_a(q) = a^q. \tag{$*$}$$

(b) Show that p_a is differentiable and find its derivative.

Recall. For $a \in \mathbb{R}^+$ and $n \in \mathbb{N}$ we have defined

$$a^n := \prod_{n=1}^{n} a, \qquad a^0 := 1, \qquad a^{\frac{1}{n}} := \text{unique positive solution of } x^n = a.$$

Therefore $a^q := \big((a^\sigma)^{\frac{1}{m}}\big)^n$ is defined for all $q = \frac{\sigma n}{m} \in \mathbb{Q}$ with $m \in \mathbb{N}$, $n \in \mathbb{N}_0$, $\sigma \in \{\pm 1\}$.

Remark. Because of the identity ($*$) one defines

$$a^z := \exp(z \ln(a)), \qquad a \in \mathbb{R}^+,\ z \in \mathbb{C}.$$

4. Let $f : [a, b] \to [a, b]$ be continuous and differentiable in (a, b) with $f'(x) \neq 1$, $x \in (a, b)$. Show that there exists exactly one $p \in [a, b]$ such that $f(p) = p$.

5. Let $f : [0, \infty) \to \mathbb{R}$ be differentiable, $f(0) = 1$ and $f'(x)f(x) \geq 0$ for all $x \in [0, \infty)$. Show that f is increasing.

6. (a) *Darboux's Theorem.* Let $f : \mathcal{D} \to \mathbb{R}$ be a differentiable function of the nonempty interval $\mathcal{D} = (a, b) \subseteq \mathbb{R}$. Show that for every $q \in \mathbb{R}$ with

$$\inf\{f'(x) : x \in \mathcal{D}\} \ < \ q \ < \ \sup\{f'(x) : x \in \mathcal{D}\}.$$

there exists a $c \in (a, b)$ such that $f'(c) = q$.

(b) Let $\mathcal{D} = (a, b)$ a nonempty interval and $f : \mathcal{D} \to \mathbb{R}$ a differentiable function with an isolated global minimum at $x_0 \in \mathcal{D}$. Is the following statement true:
There exist $c, d \in (a, b)$ such that $c < x_0 < d$ and $f'(x) \leq 0,\ x \in (c, x_0)$ and $f'(x) \geq 0,\ x \in (x_0, d)$.

7. Find all local and global extrema of

$$f : [0, \infty) \to \mathbb{R}, \quad f(x) = \frac{2 \sin x}{2 - \cos^2 x}.$$

8. Determine if the following limits exist. If they exist, find their value.

(a) $\displaystyle\lim_{x\to\infty} \Big(x - \sqrt[3]{x^3 - x^2 + 1}\Big)$,

(b) $\displaystyle\lim_{a\to\infty} \Big(1 + \frac{x}{a}\Big)^a$ with $x \in \mathbb{R}$,

(c) $\displaystyle\lim_{x\to 0} (1 + \arctan x)^{1/x}$,

(d) $\displaystyle\lim_{x\to 1} \Big(\frac{a}{1 - x^a} - \frac{b}{1 - x^b}\Big)$ with $a, b \in \mathbb{R} \setminus \{0\}$.

9. Sean $-\infty < \alpha < \beta < \infty$ y $f, g : (\alpha, \beta) \to \mathbb{R}$ funciones derivables con $g'(x) \neq 0$ en (α, β) y $\displaystyle\lim_{x\searrow\alpha} g(x) = \lim_{x\searrow\alpha} \frac{f'(x)}{g'(x)} = \infty$. Muestre que $\displaystyle\lim_{x\searrow\alpha} \frac{f(x)}{g(x)} = \infty$.

10. Let $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ and suppose that $a_n \geq 0$ for all $n \in \mathbb{N}$. Show:

$$\sum_{k=1}^{\infty} a_n \ \text{ converges} \quad \Longrightarrow \quad \sum_{k=1}^{\infty} \frac{\sqrt{a_n}}{n} \ \text{ converges.}$$

11. Let $a < c < b \in \mathbb{R}$, $\alpha : [a, b] \to \mathbb{R}$ a monotonic functions and $f, g : [a, b] \to \mathbb{R}$ bounded functions.

(a) (Theorem 6.43 (i)) Show that f is Riemann-Stieltjes integrable with respect to α if and only if the restrictions $f_1 := f|_{[a,c]}$ and $f_2 := f|_{[c,b]}$ are so and that in this case:

$$\int_a^b f(x)\ \mathrm{d}\alpha = \int_a^c f_1(x)\ \mathrm{d}\alpha + \int_c^b f_2(x)\ \mathrm{d}\alpha\,.$$

(b) Suppose there exists a set $M = \{a_1, \dots, a_n\} \subset [a, b]$ such that α is continuous in M and $f(x) = g(x)$ for all $x \in [a, b] \setminus M$. Then $f \in \mathcal{R}(\alpha)$ if and only if $g \in \mathcal{R}(\alpha)$; in this case

$$\int_a^b f(x)\ \mathrm{d}\alpha = \int_a^b g(x)\ \mathrm{d}\alpha.$$

12. Let $a \in \mathbb{R}_+$ and let $f : \mathbb{R} \to \mathbb{R}$, $f = \exp$. Use Riemann sums $s(f, P)$ and $S(f, P)$ to find $\int\limits_0^a \exp(x)\ \mathrm{d}x$.

13. (a) Does the improper integral $\int_0^\infty \frac{\sin t}{t}\ \mathrm{d}t$ exist?

(b) Does $\int_0^1 D(t)\,\mathrm{d}t$ exist, where D is the Dirichlet function

$$D : [0, 1] \to \mathbb{R}, \quad D(t) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0, 1], \\ 0 & \text{if } x \in [0, 1] \setminus \mathbb{Q}. \end{cases}$$

14. For $k, m \in \mathbb{N}$ find the integrals

$$\int_{-\pi}^{\pi} \sin(kx)\cos(mx)\,\mathrm{d}x, \qquad \int_{-\pi}^{\pi} \sin(kx)\sin(mx)\,\mathrm{d}x.$$

15. (a) Find $\lim\limits_{n\to\infty} \sqrt[n]{n!}$.

(b) Find $\lim\limits_{n\to\infty} \frac{1}{n}\sqrt[n]{n!}$.

16. For $n \in \mathbb{N}$ define

$$f_n : (0, \infty) \to \mathbb{R}, \quad f_n(x) = 2n(\sqrt[n]{2x} - 1).$$

(a) Find the pointwise limit of $(f_n)_{n\in\mathbb{N}}$.

(b) Show that $(f_n)_{n\in\mathbb{N}}$ converges uniformly on every compact interval in $(0, \infty)$.

(c) Does $(f_n)_{n\in\mathbb{N}}$ converge uniformly in $(0, \infty)$?

Hint. Write f_n as an integral.

17. Recall that $(C([0, 1]), \|\cdot\|_\infty)$ is a Banach space. Show that

$$T : C([0, 1]) \to \mathbb{C}, \qquad f \mapsto \int_0^1 f\ \mathrm{d}x$$

is a bounded linear map and find $\|T\|$. Show that T is continuous. Is it differentiable? If so, find its derivative.

Exercises for Chapter 7

1

1. (a) Use power series to find

$$\sum_{n=1}^{\infty} \frac{n}{3^{n-1}}, \qquad \sum_{n=1}^{\infty} \frac{n}{(n+1)!}.$$

(b) Find the power series representation of arctan at 0 and show that

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \cdots$$

2. (a) Sea $f : \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \longrightarrow \mathbb{R}$, $f(x) = -\log(\cos x)$. Muestre que

$$\left| f(x) - \frac{x^2}{2} \right| \le \frac{2}{3}\, |x|^3, \quad x \in \left[-\frac{\pi}{4}, \frac{\pi}{4} \right].$$

3. Let $D \subset \mathbb{R}$ be an interval, $p \in D$, $n \in \mathbb{N}_0$ and $f \in C^n(D, \mathbb{C})$. Let P be a polynomial of degree $\le n$ such that

$$P^{[k]}(p) = f^{[k]}(p), \quad k = 0, 1, \dots, n.$$

Show that $P = j_p^n f$ where $j_p^n f$ is the nth Taylor polynomial of f in p.

4. (a) Find the Taylor series at $p = 2$ and determine its radius of convergence of

$$f(x) = \frac{1}{(x-3)(x-5)}$$

(b) Find the limit (without using l'Hospital's rule)

$$\lim_{x \to 0} \frac{x - \sin x}{\mathrm{e}^x - 1 - x - x^2/2}.$$

5. (a) Let $f : \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \longrightarrow \mathbb{R}$, $f(x) = -\log(\cos x)$. Show that

$$\left| f(x) - \frac{x^2}{2} \right| \le \frac{2}{3}\, |x|^3, \quad x \in \left[-\frac{\pi}{4}, \frac{\pi}{4} \right].$$

6. (a) Show that the following function is arbitrarily often differentiable and find its Taylor series at 0. What is its radius of convergence? Where is the Taylor series equal to φ?

$$\varphi : \mathbb{R} \to \mathbb{R}, \qquad \varphi(x) = \begin{cases} \exp(-x^{-2}), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

(b) Show that the following function is arbitrarily often differentiable and find its Taylor series at 0. What is its radius of convergence?

$$g : \mathbb{R} \to \mathbb{R}, \qquad g(x) = \sum_{n=0}^{\infty} \frac{\cos(n^2 x)}{2^n}.$$

7. If $f : [-1, 1] \to \mathbb{R}$ is continuous, then

$$\lim_{t \to 0} \int_{-1}^{1} \frac{t}{t^2 + x^2}\, f(x)\, \mathrm{d}x = \pi f(0).$$

Exercises for Chapter 8

1

1. Let (X, d) a metric space and define $\mathcal{O} \subseteq \mathbb{P}X$ by

$$U \in \mathcal{O} \quad :\Longleftrightarrow \quad \forall p \in U \quad \exists \varepsilon > 0 \quad B_\varepsilon(p) \subseteq U.$$

   Show that $(X, \mathcal{O})$ is a topological space with the Hausdorff property.

   Show that for $r > 0$ and $a \in X$ the open ball $B_r(a)$ is open and the closed ball $K_r(a)$ is closed. Let $S_r(a) := \{x \in X : d(x, a) = r\}$. Show that

$$\partial B_r(a) \subseteq S_r(a) \quad \text{and} \quad \overline{B_r(a)} \subseteq K_r(a). \tag{$*$}$$

   Is equality in $(*)$ true?

2. (a) Find the interior and the closure of

$$M := \{(x, \sin x^{-1}) : x \in \mathbb{R} \setminus \{0\}\} \subseteq \mathbb{R}^2.$$

   (b) Let $(X, \mathcal{O})$ be topological space and $M \subseteq X$. Can $(\partial M)^\circ = \emptyset$ be concluded?

3. Show that every open subset of $\mathbb{R}$ is the disjoint union of at most countably many open intervals.

4. Let K, $A \subseteq \mathbb{R}^n$ such that K is compact and A is closed. Then there are $p \in K$ and $a \in A$ such that

$$|p - a| \leq |q - x|, \qquad q \in K, \ x \in A.$$

5. Let X be a topological space, A, B closed subsets of X such that $A \cup B$ and $A \cap B$ are connected. Are A and B connected?

6. Let A be a closed subset of $\mathbb{R}^n$ uach that ∂A is connected. Show that A is connected. (Hint: Use Exercise 8.5).

Bibliography

[Bla80] Christian Blatter. *Analysis. I*, volume 151 of *Heidelberger Taschenbücher [Heidelberg Paperbacks]*. Springer-Verlag, Berlin, third edition, 1980.

[Brö92] Theodor Bröcker. *Analysis. I.* Bibliographisches Institut, Mannheim, 1992. http://www.mathematik.uni-regensburg.de/broecker/index.html.

[Die69] Jean Dieudonné. *Foundations of modern analysis. Enlarged and corrected printing.* New York-London: Academic Press. XV, 387 p. , 1969.

[For06] O. Forster. *Analysis 1.* Vieweg Studium: Grundkurs Mathematik. Friedr. Vieweg & Sohn, Braunschweig, 8. edition, 2006.

[Lan51] Edmund Landau. *Foundations of Analysis. The Arithmetic of Whole, Rational, Irrational and Complex Numbers.* Chelsea Publishing Company, New York, N.Y., 1951. Translated by F. Steinhardt.

[Lan97] Serge Lang. *Undergraduate analysis.* Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997.

[Rud76] Walter Rudin. *Principles of mathematical analysis.* McGraw-Hill Book Co., New York, third edition, 1976. International Series in Pure and Applied Mathematics.

Index

Most of the informations is taken from
http://www.wikipedia.org/ and http://www.genealogy.ams.org/.

Banach, Stefan ∗ 30 March 1892 in Krakau, † 31 August 1945 in Lemberg. Polish mathematician, regarded as founder of modern functional analysis. 39

Bernoulli, Jakob ∗ 6 January 1655 in Basel; † 16 August 1705 in Basel. Swiss mathematician, after whom the Bernoulli inequality is named. 26

Bolzano, Bernard ∗ 5 October 1781 in Prague, † 18 December 1848 in Prague. Philosopher, theologian and mathematician. 46

Cantor, Georg ∗ 3 March 1845 in Saint Petersburg, Russia, † 6 January 1918 in Halle (Saale), Germany. German mathematician. Cantor made important contributions to modern mathematics; he is regarded as founder of set theory. 53

Cauchy, Augustin Louis ∗ 21 August 1789 in Paris; † 23 May 1857 in Sceaux. French mathematician, pioneer in modern analysis. 36

Dedekind, Richard ∗ 6 October 1831 in Braunschweig, † 12 February 1916 in Braunschweig. German mathematician who contributed to the axiomatic introduction of the natural and real numbers (Dedekind cuts). Moreover he is a pioneer in group theory. 27

Dirac, Paul Adrien Maurice ∗ 8 August 1902 in Bristol, † 20 Oktober 1984 in Tallahassee, Florida. British theoretical physicist with Swiss roots (his father's origins are in Saint-Maurice, Wallis). Dirac is one of the founders of quantum mechanics. His physical work was motivated by the principals of mathematical beauty: "Physical laws should have mathematical beauty". 126

Dirichlet, Johann Peter Gustav Lejeune ∗ 13 February 1805 in Düren, † 5 May 1859 in Göttingen. German mathematician, with contributions mainly to analysis and number theory. 64

Euler, Leonhard ∗ 15 April 1707 in Basel, Schweiz; † 18 September 1783 in Sankt Petersburg. Swiss mathematician. One of the most important and influential mathematicians. 46, 59

Fibonacci, Leonardo Date of birth unknown (about 1180?), † probably after 1241. Also known as Leonardo of Pisa. He is considered as one of the best mathematicians of the Middle Ages. Today, he is mostly known for the Fibonacci sequence. 35

Fréchet, Maurice René ∗ 2 September 1878 in Maligny, † 4 Juny 1973 in Paris. 91

Fresnel, August-Jean ∗ 10 May 1788, † 14 July 1827. 114

Hölder, Otto Ludwig ∗ 22 December 1859 in Stuttgart, † 29 August 1937 in Leipzig. German mathematician. 101

Hahn, Hans ∗ 27 September 1879 in Wien, † 24 July 1934 in Wien. 91

Hausdorff, Felix ∗ 8 November 1868 in Breslau, † 26 January 1942 in Bonn. German mathematician, considered as one of the founders of modern topology. He made crucial contributions to general and descriptive set theory, to measure theory, functional analysis and algebra. Moreover, using the pseudonym Paul Mongré, he is authored philosophical and literary texts. 132

Heaviside, Oliver ∗ 18 May 1850 in London, † 3 February 1925 in Homefield near Torquay. British mathematician and physicist. 64

Landau, Edmund Georg Hermann ∗ 14 February 1877 in Berlin, † 19 February 1938 in Berlin. German mathematician with important contributions to number theory. 119

Leibniz, Gottfried Wilhelm, Freiherr von ∗ 1 July 1646 in Leipzig, † 14 November 1716 in Hannover. Universal scholar. 51

Lipschitz, Rudolf ∗ 14 May 1832 in Königsberg (Preußen), † 7 October 1903 in Bonn. 65

Minkowski, Hermann ∗ 22 June 1864 in Aleksotas, formerly Russia (now Kaunas/Lithuania), † 12 January 1909 in Göttingen. 102

Peano, Giuseppe ∗ 27 August 1858 in Spinetta, Piemont, † 20 April 1932 in Turin. Italian mathematician who worked in mathematical logic, axiomatic of the natural numbers and differential equations of first order. 13

Riemann, Bernhard ∗ 17 September 1826 in Breselenz near Dannenberg (Elbe), † 20 July 1866 in Selasca at the Lago Maggiore. German mathematician who despite of his short life contributed crucially to analysis, differential geometry, mathematical physics and analytic number theory. He is considered as one of the most important mathematicians. 102

Rolle, Michel ∗ 21 April 1652 in Ambert, Basse-Auvergne, † 8 November 1719 in Paris. French mathematician. 93

Stieltjes, Thomas Jean ∗ 29 December 1856 in Zwolle, The Netherlands, † 31 December 1894 in Toulouse. Dutch mathematician who worked on the theory of continued fractions, number theory. He generalized the Riemann integral to the so-called Riemann-Stieltjes integral. 102

Weierstraß, Karl ∗ 31 October 1815 in Ostenfelde near Ennigerloh/Münsterland, † 19 February 1897 in Berlin. German mathematician. 46

Young, William Henry ∗ 20 October 1863 in London, † 7 July 1942 in Lausanne. English mathematician. 101

Greek Alphabet

A table with the hand written greek alphabet can be found at
http://www.greece.org/gr-lessons/gr-english/Gif/script.gif or at
http://www.xanthi.ilsp.gr/filog/ch1/alphabet/alphabet.asp .

α	β	γ	δ	$\varepsilon,\ \epsilon$	ζ	η	$\vartheta,\ \theta$
A	B	Γ	Δ	E	Z	H	Θ
alpha	beta	gamma	delta	epsilon	zeta	eta	theta

ι	κ	λ	μ	ν	ξ	o	π
I	K	Λ	M	N	Ξ	O	Π
iota	kappa	lambda	my	ny	xi	omikron	pi

ρ	σ	τ	υ	$\varphi,\ \varphi$	χ	ψ	ω
P	Σ	T	Υ	Φ	X	Ψ	Ω
rho	sigma	tau	ypsilon	phi	chi	psi	omega