Lecture Notes
Essentials in Mathematics
42

Felix Schwenninger
Nathanael Skrepek

The Art of ε

An Introduction to Analysis

Limited Edition
now with proofs!

Eichhütter

Felix Schwenninger Nathanael Skrepek

The Art of ε

An Introduction to Analysis

First edition

Eichhütter

Author
Felix Schwenninger
Department of Applied Mathematics
University of Twente

Author
Nathanael Skrepek
Department of Applied Mathematics
University of Twente

This version is from August 21, 2026.

Colophon
Typeset in LaTeX. First edition.

From *Rules for the Direction of the Mind* Rule II

> *. . . This furnishes us with an evident explanation of the great superiority in certitude of arithmetic and Geometry to other sciences. The former alone deal with an object so pure and uncomplicated, that they need make no assumptions at all which experience renders uncertain, but wholly consist in the rational deduction of consequences. They are on that account much the easiest and clearest of all, and possess an object such as we require, for in them it is scarce humanly possible for anyone to err except by inadvertence. And yet we should not be surprised to find that plenty of people of their own accord prefer to apply their intelligence to other studies, or to Philosophy. The reason for this is that every person permits himself the liberty of making guesses in the matter of an obscure subject with more confidence than in one which is clear, and that it is much easier to have some vague notion about any subject, no matter what, than to arrive at the real truth about a single question however simple that may be.*

René Descartes

Preface

This work is used as the lecture notes for the courses *Analysis 1 & Analysis 2* which are currently held at the University of Twente in the *first* semester[1] of the Applied Mathematics Bachelor program. The course builds on a series of four lectures in *Foundations of Mathematics* which take place in the very first week Prof. Frederic Schuller. We will also refer to both the notation and concepts in these lectures.

Clearly, there are many excellent and more extensive references for a first course in Analysis. It should be emphasized that **this reader does not follow** the "Calculus–Real Analysis"-approach as common in Anglo-Saxon countries as well as for engineering study fields worldwide; that approach usually means that a non-rigorous *calculus* course precedes a rigorous *(real) analysis* course. **Instead**, these notes follow the strictly logical, but nonetheless accessible, build-up of the subject, common in the curriculum of traditional mathematics education in central Europe. The message is that *Calculus is indeed part of (Mathematical) Analysis* and should not—at least for mathematics students—be separated from each other or even put in reversed order. To name a few manifestations of this choice, we neither assume informal understanding of the real numbers—yet, we define them axiomatically—nor knowledge of "elementary" functions such as the exponential functions or sine and cosine (we meet them in Chapter 4). The goal is to have a self-contained presentation with as little as (time-wise) necessary "jumps" so that upon digesting the material, scholars can confidently "lean their back against a solid and 100%-trustworthy wall" of mathematical theory, ready to make the next steps.

With these comments in mind, the style of this manuscript rather follows the one in German mathematics literature, and in particular the excellent book by Kaltenbäck [5] or other standard textbooks (unfortunately only) written in German [2, 3, 7, 8], and the all-time classic, Rudin's *Principles of Mathematical Analysis* [9]. We also refer to Tao's *An epsilon of room I* [10] and his other books [11, 12]. The first half of the notes has a vaguely similar setup (not necessarily content) as Zorich [13], whereas the second part differs more significantly in that regard.

[1] which combined refers to about 32 lectures (90mins) held in 16 weeks.

These notes are still subject to regular updates and extensions. They were designed and used during the past two editions of the course and are still subject to regular changes. They may (actually will) still contain a good portion of typos and smaller inconsistencies. We are therefore very grateful for being made aware of typos, mistakes or stylistic shortcomings. Also note that they may slightly differ from what we discussed in class; they may contain more detailed arguments on the on-hand and also may (yet) not contain arguments (or proofs) provided in class or the tutorial sessions.

> Thus, the notes alone cannot replace attending the lectures and are also not meant to do so.

The first edition of this manuscript was written during the winter term in the academic year 2024-2025. This would not have been possible with the great help of the rest of the *Analysis*–Team: *Carlos Pérez Arancibia* and *Karsten Kruse* — for their extraordinary dedication invested in designing and correcting homework exercises, guiding students during the tutorials and contributing far beyond this to improve the course where possible. We also very much appreciate the active participation of our students, in lectures and tutorials. We can only say: Classical blackboard-lectures and engaging tutorials with presenting solutions on the board are more alive than ever before.
To *José Iglesias* we owe the presentation of Chapter 7, which originiates from joint lecture notes for *Analysis 3* at the University of Twente.

The drive to organize such a course did foster in numerous discussions, which FS had with our esteemed colleague *Frederic Schuller*, with whom we share a crystal clear view on how mathematics should be taught.

Finally, it is very questionable if any of the following lines would have ever been written without our great teacher(s), who guided us on our endeavour to master Analysis when we were students.

Enschede, August 2026 Felix Schwenninger & Nathanael Skrepek

Contents

Chapter 1

The real numbers

This chapter takes the FOUNDATIONS OF MATHEMATICS lectures as point of departure, but also follows (parts of) the approach in Zorich [13, Chapter 2]. In particular Zorich's sections 2.1 and 2.2.1–2.2.3 as well as 2.3 and 2.4.

1.1 Ordered field and the definition of $\mathbb{R}$

Definition 1.1. Let M be a set, A relation R on M (i.e., $R \subseteq M \times M$) is called a *partial ordering on* M, if R is

(i) reflexive, i.e., $\forall x \in M \,:\, x \; R \; x$,

(ii) anti-symmetric, i.e., $\forall x \in M \,:\, (x \; R \; y \wedge y \; R \; x) \implies x = y$,

(iii) transitive, i.e., $\forall x, y, z \in M \,:\, (x \; R \; y \wedge y \; R \; z) \implies x \; R \; z$.

If additionally it holds that

(iv) $\forall x, y \in M \,:\, (x \; R \; y \vee y \; R \; x)$,

then (M, R) is called linearly ordered. ✧

Definition 1.2. A field $(M, \boxplus, \boxdot)$ is called *ordered* (or *ordered field*) if there exists a partial ordering R on M such that (M, R) is linearly ordered and

(i) $\forall x, y, z \in M \,:\, x \; R \; y \implies (x \boxplus z) \; R \; (y \boxplus z)$,

(ii) $\forall x, y \in M \,:\, (0_+ \; R \; x \wedge 0_+ \; R \; y) \implies 0_+ \; R \; (x \boxdot y)$.

We also write $(M, \boxplus, \boxdot, R)$ is an ordered field. ✧

Example 1.3 (The natural order of $\mathbb{Q}$)**.** Let us consider the field of rational numbers $(\mathbb{Q}, \boxplus, \boxdot)$ as constructed in the FOUNDATIONS OF MATHEMATICS. Recall that $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*)/_{\approx}$, that is, the quotient set of $\mathbb{Z} \times \mathbb{Z}^*$, where $\mathbb{Z}$ denotes the integers and $\mathbb{Z}^* = \mathbb{Z} \setminus \{0_{\mathbb{Z}}\}$ and

$$(p, q) \approx (\bar{p}, \bar{q}) :\Longleftrightarrow p \odot \bar{q} = \bar{p} \odot q, \qquad (p, q) \in (\mathbb{Z} \times \mathbb{Z}^*)$$

and where $(\mathbb{Z}, \oplus, \odot)$ is the ring of the integers. Let $(p, q) \in (\mathbb{Z} \times \mathbb{Z}^*)$ and consider the equivalence classes $[(p, q)]$ and $[\tilde{p}, \tilde{q}]$ in $\mathbb{Q}$. Without loss of generality, we will in the following always assume that $q, \tilde{q} \in \mathbb{N}^*$ (note that $(-p, -q) \approx (p, q)$ and $(-\tilde{p}, -\tilde{q}) \approx (\tilde{p}, \tilde{q})$)[1]. We define the relation $\leq$ in $\mathbb{Q}$ by

$$[(p, q)] \leq [(\tilde{p}, \tilde{q})] \quad :\Longleftrightarrow \quad \exists n \in \mathbb{N} : \tilde{p} \odot q = S^n(p \odot \tilde{q}),$$

where $S\colon \mathbb{Z} \to \mathbb{Z}$ is the "successor" mapping, see also F01d (Finger Practice Exercise 01d from the Linear structures lecture)[2]. Note that this definition formalizes the idea of defining the ordering on $\mathbb{Z}$ (or $\mathbb{N}$), which we could have done separately. It is a routine, but not completely trivial task to show that $\leq$ indeed is well-defined (!) and satisfies the properties of a partial ordering. The fact $(\mathbb{Q}, \boxplus, \boxdot, \leq)$ defines an ordered field is also left as exercise. ✧

Definition 1.4 (The real numbers)**.** The real numbers $\mathbb{R}$ are an ordered field $(\mathbb{R}, +, \cdot, \leq)$ which satisfies the

> **Completeness Axiom.** For any non-empty subsets X and Y of $\mathbb{R}$, with $x \leq y$ for all $x \in X$ and $y \in Y$, there exists $c \in \mathbb{R}$ such that $x \leq c \leq y$ for all $x \in X$ and $y \in Y$. ✧

Remark 1.5.

(i) We point out that—in a similar philosophy as in the Foundations of Mathematics lectures for $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$—we could have constructed the real numbers on a purely set-theoretic level. Yet, this procedure is a bit more technical than the cases that have already been addressed, which is why we deliberately choose to follow the axiomatic definition presented here. Nevertheless, at some point in time, we will comment on a specific construction to show the existence of an ordered field which satisfies the completeness axiom — this is called the *consistency* of this definition. On the other hand, one should comment on the *uniqueness* of such an ordered field. This is also postponed to some exercises. We also refer to the corresponding comments in Zorich, page 39-40.

(ii) Although the points raised in (i) may sound unsatisfactory, we emphasize that the axiomatic definition of $\mathbb{R}$ means that its defining properties (the ordered field and the completeness axiom) are the "only rules" we are at this stage allowed to work with when we deal with real numbers. ✧

---

[1] this is actually abuse of notation; more precisely, this should read $q, \tilde{q} \in \operatorname{int}(\mathbb{N})$

[2] there we only defined it as mapping from $\mathbb{N}$ to $\mathbb{N}$, but the definition can be "extended" to $\mathbb{Z}$ in a straight-forward way.

> *See also Section 2.1.1 in [13], where the presentation is more elaborate as the notions of "field" and "ordering" are not explicitly used, but their defining properties are spelled out explicitly instead.*

1.2 Direct consequences

Let us collect the following rather direct consequences from the defining properties of $\mathbb{R}$, Definition 1.4. Most of the proofs are left to exercises and self-preparation.

Fact 1.6 ($\mathbb{R}$ as a field)**.** *The following facts follow from the fact that $\mathbb{R}$ is a field.*

(i) $\exists!$ *neutral element* $0 = 0_+$ *with respect to (w.r.t.) "+", called (the)* zero (element) *in* $\mathbb{R}$*; and*
$\exists!$ *neutral element* $1 = 1.$ *w.r.t. "·", called (the)* one (element) *in* $\mathbb{R}$.

(ii) $\forall x \in \mathbb{R}\ \exists! y \in \mathbb{R} : x + y = 0_+$. *We write* $-x = y$ *and call it "additive inverse of* x*", and*
$\forall x \in \mathbb{R} \setminus \{0_+\}\ \exists! y \in \mathbb{R} : x \cdot y = 1$. *We write* $x^{-1} = \frac{1}{x} = y$ *and call it "multiplicative inverse of* x*".*

(iii) $\forall a, b \in \mathbb{R}\ \exists! x \in \mathbb{R} : a + x = b$*; this* x *is given by* $x = b + (-a) =: b - a$*;*
$\forall a \in \mathbb{R} \setminus \{0_+\}\ \forall b \in \mathbb{R}\ \exists! x \in \mathbb{R} : a \cdot x = b$*; this* x *is given by* $x = b \cdot a^{-1} =: \frac{b}{a}$.

(iv) $\forall x \in \mathbb{R} : x \cdot 0_+ = 0_+ \cdot x = 0_+$.

(v) $\forall x, y \in \mathbb{R} : (x \cdot y = 0) \implies (x = 0) \vee (y = 0)$.

(vi) $\forall x \in \mathbb{R} : -x = (-1) \cdot x$, $(-1) \cdot (-x) = x$ *and* $(-x) \cdot (-x) = x \cdot x =: x^2$.

Since $\leq$ is assumed to be a partial ordering on $\mathbb{R}$, it is reflexive. We will use the following ("common") notation to exclude the equality case for $x \leq y$: We define, for $x, y \in \mathbb{R}$,

$$x < y \quad :\Longleftrightarrow \quad (x \leq y \wedge x \neq y)$$

If $x < y$ we say that "x is strictly less than y" and we use the short-hand notation $x > y$ (read "x strictly greater than y") for $y < x$. Similarly, we say "x is less (or) equal to y" if $x \leq y$ and "x is greater (or) equal to y" if $y \leq x$.

Fact 1.7 (ordering on the field $\mathbb{R}$)**.** *For all* $x, y, z, w \in \mathbb{R}$ *we have that*

(i) *exactly one of the following three cases holds:* $(x < y) \dot\vee (y < x) \dot\vee (x = y)$.

(ii) $(x < y \wedge y \leq z) \implies (x < z)$, *and* $(x \leq y \wedge y < z) \implies (x < z)$

(iii) $(x < y) \implies x + z < y + z$
$0 < x \implies -x < 0$,

$$(x \leq y) \wedge (w \leq z) \implies (x + w \leq y + z)$$
$$(x < y) \wedge (w \leq z) \implies (x + w < y + z)$$

(iv) $(0 < x) \wedge (0 < y) \implies (0 < xy)$ *and* $(0 > x) \wedge (0 > y) \implies (0 < xy)$
$(0 < x) \wedge (0 > y) \implies (0 > xy)$
$(x < y) \wedge (0 < z) \implies (xz < yz)$ *and* $(x < y) \wedge (z < 0) \implies (xz > yz)$

(v) $0 < 1$

(vi) $0 < x \implies 0 < x^{-1}$ *and* $(0 < x < y) \implies (0 < y^{-1} < x^{-1})$

See also [13, Section 2.1.2d-e].

1.3 The completeness-axiom and consequences I

See also Zorich [13] Section 2.1.3.

Definition 1.8 (bounded (above/below) sets, (upper/lower) bound, maximal/minimal elements)**.** Let M be a subset of $\mathbb{R}$. Then we say that

- M is *bounded (from) above*, if $\exists c \in \mathbb{R} \forall x \in M : x \leq c$; any such c is called an *upper bound of M*;
- M is *bounded (from) below*, if $\exists c \in \mathbb{R} \forall x \in M : c \leq x$; any such c is called a *lower bound of M*;
- M is *bounded*, if M is bounded above and bounded below;
- $c \in \mathbb{R}$ is called *maximal in M* or *maximal element of M*, if c is an upper bound of M and $c \in M$;
- $c \in \mathbb{R}$ is called *minimal in M* or *minimal element of M*, if c is a lower bound of M and $c \in M$. ✧

We point out that a subset of $\mathbb{R}$ need neither be bounded (above/below) and neither does a bounded (above/below) set need to have a maximal/minimal element. The latter leads us to the following definition.

Definition 1.9 (Supremum and Infimum)**.** Let M be a non-empty subset of $\mathbb{R}$. If M is bounded above, then the smallest upper bound of M is called the *supremum of M* and denoted by $\sup M$. If M is bounded below, then the greatest lower bound of M is called *infimum of M* and denoted by $\inf M$. ✧

> Yet, it is not clear why the supremum and the infimum should exist for a bounded above/below set M. This is however a consequence of the completeness-axiom, as we will see soon (in Theorem 1.11).

If a subset of $\mathbb{R}$ is not bounded above, it is a common convention to denote its supremum with the symbol $+\infty$, and similarly the infimum by $-\infty$ in case a set is not bounded below. Furthermore, we may define $\sup \emptyset := -\infty$ and $\inf \emptyset := +\infty$. In fact, we could give this a formal meaning by defining and ordering on the "extended real numbers" $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

Fact 1.10. *We want to take a closer look at the notion of supremum and infimum.*

- *The definitions of the supremum and infimum can be expressed using quantifiers:*

$$s = \sup M \;:\Longleftrightarrow\; \forall x \in M \,:\, (x \le s) \wedge \big(\forall c < s\, \exists x' \in M : c < x'\big)$$
$$i = \inf M \;:\Longleftrightarrow\; \forall x \in M \,:\, (i \le x) \wedge \big(\forall c > i\, \exists x' \in M : x' < c\big)$$

or in other words,

$$\sup M = \min\{c \in \mathbb{R} \mid c \textit{ is upper bound of } M\},$$
$$\inf M = \max\{c \in \mathbb{R} \mid c \textit{ is lower bound of } M\},$$

if they exist.

- *If a set has a maximal (minimal) element, then it is unique and equals the supremum (infimum).*
  Pf.: If $c, \tilde{c} \in \mathbb{R}$ are maximal (minimal) elements of a set M, then, since $c, \tilde{c} \in M$, $c \le \tilde{c}$ and $\tilde{c} \le c$, which (by anti-symmetry of $\le$) implies that the must be equal. By definition, the maximal (minimal) element c is an upper (lower) bound of M. Since $c \in M$, any $x < c$ cannot be an upper bound of M, thus $c = \sup M$.

Theorem 1.11 (The supremum-"axiom")**.** *Every non-empty bounded above subset of $\mathbb{R}$ has a unique supremum. Every non-empty bounded below subset of $\mathbb{R}$ has a unique infimum.*

Proof. The uniqueness follows from the facts stated above the theorem. Let X be a non-empty set which is bounded above and consider the set Y of all upper bounds of X, that is,

$$Y = \{y \in \mathbb{R} \mid \forall x \in X \,:\, x \le y\}.$$

Since X is bounded above, Y is non-empty. By definition of Y, for all $x \in X$ and $y \in Y$, we have that $x \le y$. Thus, by the completeness-axiom, there exists $c \in \mathbb{R}$ such that

$$\forall x \in X \;\forall y \in Y \,:\, x \le c \le y.$$

By the first inequality, c is an upper bound of X, and, by the second inequality, any other upper bound is larger or equal to c. Thus $\sup M = c$. The proof for the existence of the infimum follows with the usual modifications. ❑

Remark 1.12. It can be shown that the statement in Theorem 1.11 also implies the completeness axiom, which is why we could have chosen either of the axioms. ✧

Example 1.13. The set $\{x \in \mathbb{R} \mid 0 \le x < 1\}$ has minimum 0 and supremum 1, which is not a maximum. ✧

Example 1.14 (The existence of a positive square root)**.** We will show that

$$\forall y \in \mathbb{R} : y \geq 0 \implies (\exists!\, x \in \mathbb{R} : x \cdot x = y \wedge x \geq 0).$$

We can show (by $(a+b)^2 - 4ab = (a-b)^2 \geq 0$) that

$$\Big(x - \frac{1}{2x}(x^2 - y)\Big)^2 \geq y \tag{1.1}$$

for $x \neq 0$ and $y \in \mathbb{R}$, see Exercise 1.8. For given $y > 0$ we define $M := \{z \in [0, +\infty) \mid z^2 \leq y\}$ and choose $x := \sup M$.

> Note that for $y \leq 1$ we have that $y^2 \leq y$ and therefore $y \in M$ on the other hand for $y > 1$ we have $1 \in M$, i.e., in any case $M \neq \emptyset$ and the supremum x of M is strictly larger than 0.

Let us assume $x^2 > y$, then by (1.1) $\hat{x} = x - \frac{1}{2x}(x^2 - y) < x$ and $\hat{x}^2 \geq y$. Hence,

$$z^2 \leq y \leq \hat{x}^2 \quad \forall z \in M,$$

which gives $z \leq \hat{x}$ (otherwise a contradiction arises). Thus, $z \leq \hat{x} < x$ which contradicts that x was the supremum, which shows that our assumption cannot hold and $x \in M$.

However, $x^2 < y$ can also not hold, because then $(x + \varepsilon)^2 < y$ for a sufficiently small $\varepsilon > 0$.[3] Hence, $x^2 = y$. ✧

Theorem 1.15 (existence of n-th root)**.** *For every natural number $n \geq 1$ and real number $y \geq 0$, there exists a unique $x \geq 0$ such that* $x^n := \underbrace{x \cdot \ldots \cdot x}_{n\text{-times}} = y$.

1.4 The integers and the rationals in $\mathbb{R}$

In the Foundations of Mathematics lectures, we have already encountered the natural numbers, the integers and the rational numbers. We will now discuss how they naturally embed in $\mathbb{R}$. Recall that the natural numbers $\mathbb{N}_0$ where defined based on the infinity axiom of ZFC, while the integers $\mathbb{Z}$ where introduced as a subset (more precisely, a quotient space) of $\mathbb{N}_0 \times \mathbb{N}^*$. The rational numbers $\mathbb{Q}$, in turn, where defined as quotient space of $\mathbb{Z} \times \mathbb{Z}^*$. Using the canonical embeddings $\mathrm{int}\colon \mathbb{N}_0 \to \mathbb{Z}$, $\mathrm{rat}\colon \mathbb{Z} \to \mathbb{Q}$, we have

$$\mathrm{rat}(\mathrm{int}(\mathbb{N}_0)) \subseteq \mathrm{rat}(\mathbb{Z}) \subseteq \mathbb{Q},$$

which we will usually write in the common short-hand notation $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.

---

[3] For example $\varepsilon \leq \min(\frac{y - x^2}{2x+1}, 1)$.

We will now explain how these sets can be naturally identified with subsets in $\mathbb{R}$. We define

$$f\colon \mathbb{Z} \to \mathbb{R}, \quad [(m,n)] \mapsto \begin{cases} S^{m-n-1}(1_{\mathbb{R}}), & m > n,^{4} \\ -S^{n-m-1}(1_{\mathbb{R}}), & m < n, \\ 0_{\mathbb{R}}, & m = n, \end{cases}$$

and

$$g\colon \mathbb{Q} \to \mathbb{R}, \quad [(m,n),(q,r)] \mapsto f([m,n]) \cdot f([q,r])^{-1}. \tag{1.2}$$

Lemma 1.16. *The function* $g\colon \mathbb{Q} \to \mathbb{R}$ *(defined in (1.2)) is well-defined, injective and satisfies the properties*

$$\forall x, y \in \mathbb{Q} \; : \; g(x \boxplus y) = g(x) + g(y) \wedge g(x \boxdot y) = g(x) \cdot g(y).$$

Proof. (Tedious) exercise. ❑

Definition 1.17. Let $g\colon \mathbb{Q} \to \mathbb{R}$ be defined by (1.2). The set

- $g(\mathbb{N}^*)$ is called the *natural numbers in* $\mathbb{R}$ and we use the notation $\mathbb{N}^*$;
- $g(\mathbb{Z})$ is called the *integers in* $\mathbb{R}$ and we use the notation $\mathbb{Z}$;
- $g(\mathbb{Q})$ is called the *rational numbers in* $\mathbb{R}$ *and we use the notation* $\mathbb{Q}$.

Furthermore, we call $g(\mathbb{N})$ the *natural numbers with zero* and denote them by $\mathbb{N}_0$ and use the (short-hand) notation $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}, \mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$. ✧

Definition 1.18 (inductive set)**.** A subset X of $\mathbb{R}$ is called inductive, if for every $x \in X$ it holds that $x + 1 \in X$. ✧

Lemma 1.19. $\mathbb{N}^*$ *is the smallest inductive set in* $\mathbb{R}$ *containing the element* 1*, that is,* $\mathbb{N}^*(= g(\mathbb{N}^*))$ *is the intersection of all inductive subsets of* $\mathbb{R}$ *containing the element* $1 = 1_{\mathbb{R}}$*.*

Proof. For the moment we still distinguish the notation "$\mathbb{N}^*$" (as defined in the FOUNDATIONS OF MATHEMATICS) and $g(\mathbb{N}^*)$ (we will shortly identify those). By definition in the Foundations of Mathematics (see the Axiom of Infinity), the natural numbers are the smallest non-empty set which for each element in it also contain its "successor" ($x \in \mathbb{N}^* \implies \{x\} \in \mathbb{N}^*$).

Let M be an inductive set in $\mathbb{R}$ containing 1. It follows from the definition of g and its properties that $S := \{x \in \mathbb{N}^* \mid g(x) \in M\} \cup \{\emptyset\}$ is a set containing the emptyset (the "0") as well as every successor of an element in the set. Since $\mathbb{N}^*$ is the smallest such set by definition, it follows that $S = \mathbb{N}^*$ and thus $g(\mathbb{N}^*) \subseteq M$. This shows that $g(\mathbb{N}^*)$ is the smallest inductive set containing $1_{\mathbb{R}}$. ❑

Remark 1.20. Some clarifications about our notation are in order.

(i) Since g is injective, there is no ambiguity in using the notation $\mathbb{N}_0, \mathbb{N}^*, \mathbb{Z}$ and $\mathbb{Q}$.

[4] Note that here $n > m$ is a short-hand notation for $\exists p \in \mathbb{N}^* \; : \; n = m + p$.

(ii) We point out that in this course, the “natural numbers” will usually refer to the *positive integers*, thus excluding zero. To be consistent with the linear structures course, we will stick to the notation $\mathbb{N}^*$ and use $\mathbb{N}_0$ to stress that 0 is included. We point out that in analysis text books it is usual to use just $\mathbb{N}$ for what we denote by $\mathbb{N}^*$.

(iii) Instead of referring to the construction of $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$ from the FOUNDATIONS OF MATHEMATICS course, we could have also defined $\mathbb{N}^*, \mathbb{Z}$ and $\mathbb{Q}$ directly as subsets of our, axiomatically defined, real numbers (see Zorich Section 2.2). Yet, the intuition of both approaches is the same: roughly speaking, we define $\mathbb{N}^*$ to be the set

$$\{1, 1+1, 1+1+1, \dots\},$$

$\mathbb{Z}$ to be the union of $\mathbb{N}$ and the set of all additive inverses of $\mathbb{N}$ and $\mathbb{Q}$ to be all elements of $\mathbb{R}$ of the form $p \cdot q^{-1}$ where p is in $\mathbb{Z}$ and q is in $\mathbb{Z}^*$. ✧

Convention. From now on we identify the natural numbers/integers/rational numbers as constructed in the FOUNDATIONS OF MATHEMATICS with $g(\mathbb{N}_0) = \mathbb{N}_0, g(\mathbb{N}^*) = \mathbb{N}^*, g(\mathbb{Z})$ and $g(\mathbb{Q})$ and write $\mathbb{N}_0, \mathbb{N}^*, \mathbb{Z}$ and $\mathbb{Q}$ from now onwards. Furthermore, we will make use of the common notation $2 = 1+1, 3 = 1+1+1, \dots$.

Corollary 1.21 (Principle of mathematical induction)**.** *Any subset of* $\mathbb{N}^*$ *which contains* 1 *and is inductive equals* $\mathbb{N}^*$*.*

Example 1.22. We use mathematical induction to show that *the sum of the first n natural numbers equals* $n \cdot (n+1) \cdot 2^{-1}$, i.e.,

$$1 + 2 + \cdots + n = n \cdot (n+1) \cdot 2^{-1}.$$

More precisely, we show that the function defined by $s(1) := 1$ and $s(n) := n + s(n-1)$ for all $n \in \mathbb{N}^*$ with $n > 1$, equals the function $\tilde{s} \colon \mathbb{N}^* \to \mathbb{R}, n \mapsto n \cdot (n+1) \cdot 2^{-1}$.

To do so we show that the set $M = \{n \in \mathbb{N}^* \mid s(n) = \tilde{s}(n)\}$ contains 1 (*induction assumption*) and is inductive (*induction step*). By definition of s and $\tilde{s}$ we see $\tilde{s}(1) = 1 = s(1)$. To see that M is inductive, let $n \in M$ and consider $n+1$. Thus, we assume $s(n) = \tilde{s}(n)$ and we want to show that $s(n+1)$ equals $\tilde{s}(n+1)$. By definition of s, $s(n+1) = n+1+s(n)$, and the latter equals $n+1+n \cdot (n+1) \cdot 2^{-1}$ since $n \in M$. Using distributivity, we get

$$\begin{aligned} s(n+1) &= n+1+n\cdot(n+1)\cdot 2^{-1} = (n+1)\cdot(1+n\cdot 2^{-1}) = (n+1)\cdot(2+n)\cdot 2^{-1} \\ &= \tilde{s}(n+1). \end{aligned}$$

Thus, M is inductive and by Corollary 1.21 coincides with $\mathbb{N}^*$, which implies $s(n) = \tilde{s}(n)$ for all $n \in \mathbb{N}^*$. ✧

Definition 1.23. Any real number $x \in \mathbb{R}$ which is not rational, i.e., $x \notin \mathbb{Q}$, is called *irrational*. ✧

Proposition 1.24. *There exists an irrational number in* $\mathbb{R}$. *More precisely, the square root of* $2 = 1 + 1$ *is not rational.*

Proof. By Example 1.14 (and more generally Theorem 1.15), we now that a unique $x \in \mathbb{R}$ exists with $x \geq 0$ and $x^2 = 2$. It remains to show that $x \notin \mathbb{Q}$. Suppose that $x \in \mathbb{Q}$. Then there exists $p \in \mathbb{N}$ and $q \in \mathbb{N}^*$ such that $x = p \cdot q^{-1}$. It follows that $p^2 = 2 \cdot q^2$. We will employ the following simple facts of natural numbers.

(i) $\forall m \in \mathbb{N}^* \; \forall n \in \mathbb{N}_0 \; : \; m \neq 1 \implies m \cdot n \neq 1$

(ii) $\forall m \in \mathbb{N}^* \; \exists r \in \mathbb{N}_0 \; : \; (m = 2r + 1) \;\dot{\vee}\; (m = 2r)$

(iii) $\forall m \in \mathbb{N}^* \; : \; (\exists r \in \mathbb{N}^* \; : \; m = 2r) \iff (\exists s \in \mathbb{N}^* \; : \; m^2 = 2s) \iff (\exists t \in \mathbb{N}^* \; : \; m^2 = 4t)$

(the first item can be used to show the others). By the definition of $\mathbb{Q}$, we can assume that exactly one of p or q is divisible by 2 (*think about what this should mean according to our setting*), otherwise we get a direct contradiction to (ii) above from $p^2 = 2q^2$. It follows that exactly one of p^2 and q^2 is divisible by 4 and exactly one is not divisible by 2, by (ii) and (iii). But this contradicts $p^2 = 2 \cdot q^2$. ❑

Fact 1.25. *The following facts can be shown by using Lemma 1.19 (see also the proofs in Zorich, Section 2.2.2) or by relating to our definition of* $\mathbb{N}^*$ *through the function* g.

(i) $\forall n, m \in \mathbb{N}^* \implies n + m \in \mathbb{N}^*$
Pf: Clear from the properties of g, *Lemma 1.16*

(ii) $\forall n \in \mathbb{N}^* \; : \; n \neq 1 \implies n - 1 \in \mathbb{N}^*$

(iii) $\forall n \in \mathbb{N}_0 \; : \; \min\{x \in \mathbb{N}_0 \mid n < x\} = n + 1$

(iv) $\forall n, m \in \mathbb{N}_0 \; : \; n < m \implies (n + 1) \leq m$

(v) $\forall n \in \mathbb{N}_0 \; \nexists x \in \mathbb{N}_0 \; : \; n < x < n + 1$

(vi) $n - 1$ *is immediate predecessor of* n *for all* $n \in \mathbb{N}^*$, $n \neq 1$

(vii) *every subset of* $\mathbb{N}^*$ *has a minimal element.*

1.5 The Archimedian property of $\mathbb{R}$

Fact 1.26. *We list the following straightforward boundedness properties of subsets of* $\mathbb{Z}$.

(i) *Every bounded above, non-empty subset* M *of* $\mathbb{N}^*$ *or* $\mathbb{Z}$ *has a maximal element*
Pf.: Let $s = \sup M$ *which exists as* M *is bounded above, by Theorem 1.11. As* s *is the least upper bound, there exists* $n \in M$ *such that*

$s-1 < n \leq s$. Rearranging this inequality chain gives $n \leq s < n+1$. By Fact 1.25 (v), we conclude that $s = n$.

(ii) *$\mathbb{N}^*$ is not bounded above.*
Pf.: If $\mathbb{N}^*$ was bounded above, then by the first item, there exists a maximal element n in $\mathbb{N}^*$. But this contradicts the fact that $n+1 \in \mathbb{N}$ and $n < n+1$.

(iii) *$\mathbb{Z}$ is not bounded below.*

(iv) *Any non-empty, bounded below subset of $\mathbb{N}^*$ has a minimal element.*

(v) *$\mathbb{Z}$ is neither bounded above nor below.*

Proposition 1.27 (Principle of Archimedes)**.** *Let $x \in \mathbb{R}$ and $h > 0$. Then there exists a unique $k \in \mathbb{Z}$ such that $(k-1)h \leq x < kh$.*

Proof. Consider the set $M = \{n \in \mathbb{Z} \mid x < nh\}$, which is bounded below. Therefore there exists a minimal element k in the set, by Fact 1.26. By the definition of the minimal element and M, it holds that

$$k = \min M \iff k-1 \notin M \wedge k \in M \iff (k-1)h \leq x < kh,$$

which settles the proof including the uniqueness of k (by the uniqueness of the minimum). ❑

Fact 1.28. *The Principle of Archimedes yields the following.*

(i) $\forall \varepsilon > 0\ \exists n \in \mathbb{N}^* : 0 < \frac{1}{n} < \varepsilon$
Pf.: Take $x = 1$ and $h = \varepsilon$ and apply the Archimedian Principle.

(ii) $\forall x \in \mathbb{R} : (\forall n \in \mathbb{N}^* : 0 \leq x < \frac{1}{n}) \implies (x = 0)$

(iii) $\forall a, b \in \mathbb{R} : (a < b) \implies (\exists r \in \mathbb{Q} : a < r < b)$
Pf.: Let $a, b \in \mathbb{R}$ such that $a < b$. By the first fact, $\exists n \in \mathbb{N}^* : \frac{1}{n} < b - a$. By the Archimedian principle (with $h = \frac{1}{n}$) there exists $k \in \mathbb{Z}$ such that $\frac{k-1}{n} \leq a < \frac{k}{n}$. All three inequalities together imply that $a < \frac{k}{n} < b$.

(iv) $\forall x \in \mathbb{R}\ \exists! z \in \mathbb{Z} : z \leq x < z+1$ *(this z is also denoted by $\lfloor x \rfloor$ and called the* integer part *of x, the function $\lfloor \cdot \rfloor$ is called the* floor*-function).*

1.6 The completeness-axiom and consequences II

In the following we need the notion of *a sequence of (sub)sets* $A_1, A_2, A_3, \ldots$ of some set X, which is defined as function

$$A \colon \mathbb{N}^* \to \mathcal{P}(X),$$

where $\mathcal{P}(\mathbb{R})$ refers to the power-set of X. Anyway, we will use the short-hand notation $A_1, A_2, A_3, \ldots$

Definition 1.29. Let X be a set. A sequence $A\colon \mathbb{N}^* \to \mathcal{P}(X)$ of sets is called *nested* if $A\colon \mathbb{N}^* \to \mathcal{P}(X)$ is decreasing with respect to the inclusion ordering on X, that is $\forall n \in \mathbb{N}^* : A_n \supseteq A_{n+1}$. We also use the shorthand notation $A_1 \supseteq A_2 \supseteq \dots$. ✧

We want to recall the notion of intervals from school.

Definition 1.30. For $a, b \in \mathbb{R}$ we introduce the following sets

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

Every of those previous sets is less specifically just called *interval* or in contrast to the next type of interval *bounded interval.* The numbers a and b are called the *boundary points* of the interval. Additionally, we introduce *unbounded intervals*

$$\begin{aligned}
\textit{closed unbounded interval:} \qquad & [a,\infty) = \{x \in \mathbb{R} \mid x \geq a\}, \\
\textit{closed unbounded interval:} \qquad & (-\infty,a] = \{x \in \mathbb{R} \mid x \leq a\}, \\
\textit{open unbounded interval:} \qquad & (a,\infty) = \{x \in \mathbb{R} \mid x > a\}, \\
\textit{open unbounded interval:} \qquad & (-\infty,a) = \{x \in \mathbb{R} \mid x < a\}.
\end{aligned}$$

Also closed and open unbounded intervals are less specifically just called *intervals.*

Note that bounded intervals can be empty, e.g., if $b < a$. Hence, if an interval is empty, then we call it an *empty interval.* Moreover, we say $[a, b]$ is a *trivial interval*, if $a = b$.

> If an interval is neither an empty interval nor a trivial interval, then we call it a *nontrivial interval.*

For every bounded interval I with boundary points a and b, we define the *length* of I by

$$|I| = \max\{b - a, 0\}.$$ ✧

Remark 1.31. Let us remark immediate consequences of the previous definition.

(i) Note that if $b < a$, then all of (a, b), $(a, b]$, $[a, b)$, $[a, b]$ are empty. Therefore, an empty interval qualifies as open, closed and semi-open.

(ii) Every trivial interval contains exactly one number.

(iii) Every nontrivial interval contains infinitely many numbers. On the other hand every interval that contains more than one point is nontrivial. ✧

Theorem 1.32 (Nested intervals principle)**.** *For any nested sequence of nonempty closed intervals $I_1 \supseteq I_2 \supseteq \dots$ (in $\mathbb{R}$) there exists $c \in \mathbb{R}$ which lies in all I_k, for $k \in \mathbb{N}^*$—in other words, $c \in \bigcap_{k\in\mathbb{N}^*} I_k$.*

If, in addition, for every $\varepsilon > 0$ there exists $k \in \mathbb{N}^*$ such that $|I_k| < \varepsilon$, then c is unique.

Proof. Let $I_k = [a_k, b_k]$, $k \in \mathbb{N}^*$ and consider the sets $A = \{a_k \in \mathbb{R} \mid k \in \mathbb{N}^*\}$ and $B = \{b_k \in \mathbb{R} \mid k \in \mathbb{N}^*\}$, which are obviously non-empty. Since we assumed that the intervals are nested, it follows

$$\forall a \in A\, \forall b \in B \,:\, a \leq b.$$

Thus, by the completeness axiom, there exists $c \in \mathbb{R}$ such that for all $a \in A, b \in B$ we have that $a \leq c \leq b$, which implies that $c \in I_k$ for all $k \in \mathbb{N}^*$.

Under the additional assumption, suppose that c_1 and c_2 are different elements in the intersection of all I_k's. Let $\varepsilon = |c_1 - c_2| > 0$. Then there exists $k \in \mathbb{N}^*$ such that $|I_k| < \varepsilon/2$, which implies that I_k cannot contain both c_1 and c_2 (if it would, then $|I_k| = \max I_k - \min I_k \geq |c_1 - c_2| = \varepsilon$, (*why?*)). Contradiction. ❑

1.7 Countable and uncountable sets

Definition 1.33 (finite, countable and uncountable set)**.** Let M be a set. We say that

- M is finite, if there exists a bounded above subset N of $\mathbb{N}^*$ and a bijection $f\colon N \to M$.
- M is countable, if there exists a bijection $f\colon \mathbb{N}^* \to M$.
- M is at most countable, if M is either finite or countable.
- M is infinite, if M is not finite.
- Let M be infinite. M is uncountable if M is not countable. ✧

Clearly, $M = \mathbb{N}^*$ is countable, as f can be chosen to be the identity mapping. We can also construct a bijection $f\colon \mathbb{N}^* \to \mathbb{Z}$, showing that $\mathbb{Z}$ is countable. Furthermore, $\mathbb{Q}$ turns out to be countable (*think about why*).

Proposition 1.34. *The following two assertions hold.*

(i) *Every infinite subset of a countable set is countable.*

(ii) *Finite or countable unions of sets which are at most countable, are at most countable.*
That means: Let X be a set of sets. Then

$$\Big(X \text{ is at most countable } \wedge\ (\forall M \in X \,:\, M \text{ is at most countable})\Big) \implies \bigcup X \text{ is at most countable.}$$

Proof. See Exercise 1.26. ❑

Theorem 1.35. *The real numbers are uncountable.*

Proof. We show that the interval $[0,1]$ is uncountable by using the nested interval principle Theorem 1.32. ❑

1.8 Exercises

Exercise 1.1. Let H be the set of all humans. We define the relation R on H by

$$x \, R \, y \quad :\Longleftrightarrow \quad x \text{ is a cousin of } y \ \vee \ x = y.$$

Question	True	False
Is R reflexive?	☐	☐
Is R anti-symmetric?	☐	☐
Is R symmetric?	☐	☐
Is R transitive?	☐	☐
Is R a partial order?	☐	☐
Is R an equivalence relation?	☐	☐

✧

Exercise 1.2. Let $\mathbb{F}$ be the field consisting of the set $\{0,1,2\}$ and the operations $(\boxplus, \boxdot)$ defined by:

$$\begin{array}{c|ccc} \boxplus & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \qquad \text{and} \qquad \begin{array}{c|ccc} \boxdot & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

(a) Show that $\mathbb{F}$ is indeed a field (Definition of field from the foundation lecture).

(b) Show that the relation $R \subseteq \mathbb{F} \times \mathbb{F}$ defined by

$$R := \{(0,0),(1,1),(2,2),(0,1),(0,2),(1,2)\}$$

does not make $\mathbb{F}$ an ordered field.

(c) Show that there is no relation R such that $(\mathbb{F}, R)$ is an ordered field. ✧

Exercise 1.3. Prove the remaining items of Fact 1.6. ✧

Exercise 1.4. Let $(\mathbb{F}, +, \cdot)$ be a field and $p, q \in \mathbb{F}$. Moreover, let $f\colon \mathbb{F} \to \mathbb{F}$ be a function defined by $f(x) := x^2 + px + q$ and $x_1 \in \mathbb{F}$ a zero of f, i.e., $f(x_1) = 0$.

(a) Show that $x_2 := -p - x_1$ is also a zero of f.

Hint: Note that $f(x) = f(x) - f(x_1)$ and $(x^2 - x_1^2) = (x - x_1)(x + x_1)$.

(b) Show that there is no further zero of f, i.e., f has maximal two zeros. ✧

Exercise 1.5. Prove the remaining items of Fact 1.7. ✧

Exercise 1.6. Let $x, y \geq 0$. Show that $x^2 \leq y^2$ implies $x \leq y$. ✧

Exercise 1.7. Let $x, y \in \mathbb{R}$ such that $x, y \geq 0$ and $x^2 < y$. We define

$$\varepsilon := \min\bigl((2x+1)^{-1}(y - x^2), 1\bigr).$$

Show that

$$(x + \varepsilon)^2 \leq y.$$

Use this to conclude that the supremum x of the set $\{z \in \mathbb{R} \mid z^2 \leq y\}$ does not satisfy $x^2 < y$. ✧

Exercise 1.8. Let x and y be strictly positive real numbers such that $x^2 > y$. Prove (by using the axioms and the facts seen in the lecture) that

$$\Bigl(x - (2x)^{-1}\left(x^2 - y\right)\Bigr)^2 \geq y.$$

Use this to conclude that the supremum x of the set $\{z \in \mathbb{R} \mid z^2 \leq y\}$ does not satisfy $x^2 > y$. ✧

Exercise 1.9. Let A and B be two bounded subsets of $\mathbb{R}$. Prove that

$$\sup(A \cup B) = \max(\sup A, \sup B).$$

Is there any corresponding result for $A \cap B$? ✧

Exercise 1.10. Consider the function $|\cdot|\colon \mathbb{R} \to \mathbb{R}$ defined by

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

This function is called *absolute value.* Prove that for $\{a, b\} \subseteq \mathbb{R}$:

(a) $|a \cdot b| = |a| \cdot |b|$,

(b) $|a + b| \leq |a| + |b|$,

(c) $|a - b| \geq \bigl||a| - |b|\bigr|$. ✧

Exercise 1.11. Let S and T be nonempty subsets of $\mathbb{R}$ with the following property: $s \leq t$ for all $s \in S$ and $t \in T$.

(a) Show that S is bounded above and T is bounded below.

(b) Prove $\sup S \leq \inf T$.

(c) Give an example of such sets S and T where $S \cap T$ is nonempty.[5]

(d) Give an example of sets S and T where $\sup S = \inf T$ and $S \cap T$ is the empty set.[6] ✧

Exercise 1.12. Prove that if $\{a, b\} \subseteq \mathbb{R}$ and

$$a < b + \varepsilon$$

for all $\varepsilon > 0$, then $a \leq b$. ✧

Exercise 1.13. Let $(\mathbb{F}, +, \cdot)$ be a field such that $2 := 1 + 1 \neq 0$ and $4 := 1 + 1 + 1 + 1 = 2 \cdot 2 \neq 0$. Moreover, let $p, q \in \mathbb{F}$ and $f \colon \mathbb{F} \to \mathbb{F}$, $x \mapsto x^2 + px + q$. Show that f has a zero if and only if there exists a $y \in \mathbb{F}$ such that $y^2 = p^2/4 - q$. Show that in that case the zeros are given by $-p/2 + y$ and $-p/2 - y$.
Hint: Completing the square. ✧

Exercise 1.14. Let $A + B$ be the set of numbers of the form $a + b$ and $A \cdot B$ the set of numbers of the form $a \cdot b$, where $a \in A \subseteq \mathbb{R}$ and $b \in B \subseteq \mathbb{R}$. Determine whether it is always true that

(a) $\sup(A + B) = \sup A + \sup B$,

(b) $\sup(A \cdot B) = \sup A \cdot \sup B$. ✧

Exercise 1.15. Let $-A$ be the set of numbers of the form $-a$, where $a \in A \subseteq \mathbb{R}$. Show that $\sup(-A) = -\inf A$. ✧

Exercise 1.16. Prove each of the following statements:

(a) $\forall m \in \mathbb{N}^* \; \forall n \in \mathbb{N}_0 \; : \; m \neq 1 \Longrightarrow m \cdot n \neq 1$

(b) $\forall m \in \mathbb{N}^* \; \exists r \in \mathbb{N}_0 \; : \; (m = 2r + 1) \mathbin{\dot\vee} (m = 2r)$

(c) $\forall m \in \mathbb{N}^* \; : \; (\exists r \in \mathbb{N}^* \; : \; m = 2r) \Longleftrightarrow \big(\exists s \in \mathbb{N}^* \; : \; m^2 = 2s\big)$
$\Longleftrightarrow \big(\exists t \in \mathbb{N}^* \; : \; m^2 = 4t\big)$ ✧

Exercise 1.17. Prove that for all $a \in \mathbb{R}$ and each $n \in \mathbb{N}^*$ there exists a rational r_n such that $|a - r_n| < 1/n$. ✧

Exercise 1.18. Show that the following statements are true:

(a) Either $n \in \mathbb{N}_0$ is divisible by 2 or $n + 1$ is divisible by 2.

[5] You have to prove that S and T have these properties; just giving S and T is not enough.

[6] see Footnote 5

(b) $2^n > n$ for all $n \in \mathbb{N}^*$.

(c) $2^n > n^2$ whenever $n \in \mathbb{N}^*$ and $n \geq 5$.

(d) $2^n > n^3$ whenever $n \in \mathbb{N}^*$ and $n \geq 10$.

(e) $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$ ✧

Exercise 1.19. Suppose that A and B are sets and that B is uncountable. If there exists a function f which takes A onto B (i.e., $f\colon A \to B$ is surjective), prove that A is uncountable. ✧

Exercise 1.20. Let a and b be non-negative rational numbers such that $\sqrt{a} + \sqrt{b}$ is rational. Show that $\sqrt{a}$ and $\sqrt{b}$ are both rational. ✧

Exercise 1.21. Let $a \in \mathbb{R}$ and $M := \{r \in \mathbb{Q} \mid r < a\}$ as a subset of $\mathbb{R}$. Show that $\sup M = a$. ✧

Exercise 1.22. Show by mathematical induction that for all $n \in \mathbb{N}^*$

$$\sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k} = \sum_{k=1}^{n} \frac{1}{n+k}.$$ ✧

Exercise 1.23. Let $a \in \mathbb{N}^*$. Prove that:

(a) For all $n \in \mathbb{N}^*$, the number $(a+1)^n - 1$ is divisible by a.

(b) For all even $n \in \mathbb{N}^*$, the number $(a-1)^n - 1$ is divisible by a. ✧

Exercise 1.24. Let $\mathbb{F}$ be a finite field. Show that:

(a) $\exists n \in \mathbb{N}^* : \sum_{j=1}^n 1 = 0$.

Hint: Assume that there is no such n and show that then $\sum_{j=1}^n 1$ has to be different for all $n \in \mathbb{N}^*$.

(b) There is no relation R on $\mathbb{F}$ such that $(\mathbb{F}, R)$ is an ordered field. ✧

Exercise 1.25. Show that there is no relation R on $\mathbb{C}$ that makes $(\mathbb{C}, R)$ an ordered field. ✧

Exercise 1.26. Prove that:

(a) Every infinite subset of a countable set is countable.

(b) Finite or countable unions of sets which are at most countable, are at most countable.

That means: Let X be a set of sets. Then

$$\big(X \text{ is at most countable} \ \wedge\ (\forall M \in X : M \text{ is at most countable})\big) \implies \bigcup X \text{ is at most countable}$$ ✧

Exercise 1.27. For $n \in \mathbb{N}^*$, let $n!$ (*read "n factorial"*) denote the product $1 \cdot 2 \cdot 3 \cdots n$ and let $0! := 1$, i.e., $n! := n \cdot (n-1)!$ for $n > 0$ and $0! := 1$. Furthermore, we define

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \qquad \text{for} \qquad k = 0, 1, \ldots, n.$$

The binomial theorem asserts that for all $a, b \in \mathbb{R}$ it holds that

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

(a) Verify the binomial theorem for $n = 1, 2$, and 3.

(b) Show $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$ for $k = 1, 2, \ldots, n$.

(c) Prove the binomial theorem using mathematical induction and item (b).

(d) Use the binomial theorem to prove that $(a+b)^n \geq a^n + b^n$ for all $n \in \mathbb{N}^*$ and $a, b \geq 0$.

(e) Use the binomial theorem to prove that $(1 + 1/n)^n \geq 2$ for all $n \in \mathbb{N}^*$. ✧

Exercise 1.28. Show that if $\mathbb{R}$ and $\mathbb{R}'$ are two models of the set of real numbers and $f\colon \mathbb{R} \to \mathbb{R}'$ is a mapping such that $f(x+y) = f(x) + f(y)$ and $f(x \cdot y) = f(x) \cdot f(y)$ for any $x, y \in \mathbb{R}$, then

(a) $f(0) = 0'$.

(b) $f(1) = 1'$ if $f(x) \not\equiv 0'$, which we shall henceforth assume.

(c) $f(m) = m'$ where $m \in \mathbb{Z}$ and $m' \in \mathbb{Z}'$ are either

$$\left(S_{\mathbb{R}}^{|m|}(1) \;\text{ and }\; S_{\mathbb{R}'}^{|m|}(1')\right) \quad \text{or} \quad \left(-S_{\mathbb{R}}^{|m|}(1) \;\text{ and }\; -S_{\mathbb{R}'}^{|m|}(1')\right).$$

The mappings $S_{\mathbb{R}}$ and $S_{\mathbb{R}'}$ are the successor functions in the respective sets $\mathbb{R}$ and $\mathbb{R}'$, i.e., for $n \in \mathbb{N}$, $n > 1$ and $x \in \mathbb{R}$

$$S_{\mathbb{R}}^{n} = S_{\mathbb{R}} \circ S_{\mathbb{R}}^{n-1} \quad \text{and} \quad S_{\mathbb{R}}^{1}(x) = S_{\mathbb{R}}(x) = x + 1,$$

and accordingly for $\mathbb{R}'$.

(d) The mapping $f\colon \mathbb{Z} \to \mathbb{Z}'$ is injective and preserves the order.

(e) $f\left(\frac{m}{n}\right) = \frac{m'}{n'}$, where $m, n \in \mathbb{Z}, n \neq 0, m', n' \in \mathbb{Z}', n' \neq 0', f(m) = m'$, $f(n) = n'$. Thus $f\colon \mathbb{Q} \to \mathbb{Q}'$ is a bijection that preserves order.

(f) $f\colon \mathbb{R} \to \mathbb{R}'$ is a bijective mapping that preserves order. ✧

Chapter 2

Sequences and limits

Definition 2.1. Let X be a set. A function $x\colon \mathbb{N}^* \to X$ is called *sequence in* X. We often use the alternative notation $(x_n)_{n\in\mathbb{N}^*}$ or short-hand (x_n).[1] The set of sequences in X is denoted by $X^{\mathbb{N}^*}$. If $X = \mathbb{R}$, we will also say that $x = (x_n)_{n\in\mathbb{N}^*}$ is a sequence of real numbers. ✧

One may also understand vectors as "sequences of finite length", but we won't use this analogy here. In this chapter we will mostly study sequences of *real numbers*[2], but, for convenience of later chapters (and in particular Analysis 2), we start with more general sequences which includes sequences in $\mathbb{R}$ as special case. For the moment, it is fine to always have this special case in mind. For the more general framework we need the notion of a distance function, which formalizes the well-known properties of the function $(x, y) \mapsto |x - y|$ for reals x, y.

Definition 2.2. Let X be a set. A function $d\colon X \times X \to [0, \infty)$ is called a *distance function* or *metric* on X, if

(i) $\forall x, y \in X \,:\, d(x, y) = 0 \iff x = y$,

(ii) $\forall x, y \in X \,:\, d(x, y) = d(y, x)$,

(iii) $\forall x, y, z, \in X \,:\, d(x, z) \leq d(x, y) + d(y, z)$,

and the pair (X, d) is called a *metric space*[3]. For every metric space (X, d) and $x \in X$ we define the *ball* around x with radius $r > 0$ by

$$\mathrm{B}_r(x) \coloneqq \{y \in X \mid d(x, y) < r\}.$$ ✧

Example 2.3. We provide a few folklore examples for metric spaces.

[1] Please note that in [13] the notation "$\{x_n\}$" is used. We refrain from using that on as we want to point out that a sequence is a function(!) and not a set.

[2] Note that Zorich [13] uses the notion of *numerical sequence*.

[3] Note that the word *space* has no further meaning! (e.g., X need not be a vector space).

(a) The most important one: $X = \mathbb{R}$ and $d(x, y) = d_2(x, y) := |x - y|$ defined for every $x, y \in \mathbb{R}$.

> We refer to this metric as *the standard metric* or *the canonical metric* on $\mathbb{R}$.

(b) The simplest one: Let $X \neq \emptyset$ and $d = d_0$ defined by

$$d_0(x, y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y, \end{cases} \quad \text{for} \quad x, y \in X.$$

(c) One abstract one: Let (X, d) be a metric space and $Y \subseteq X$. Let $d|_{Y\times Y} : Y \times Y \to [0, \infty), (x, y) \mapsto d(x, y)$ (which is called *the restriction of d to the subset $Y \times Y$*). Then $(Y, d_{Y\times Y})$ is also a metric space.

(d) The most important one (general): Let $n \in \mathbb{N}^*$ and $X = \mathbb{R}^n$ and

$$d(\mathbf{x}, \mathbf{y}) := d_2(\mathbf{x}, \mathbf{y}) := \left(\sum_{j=1}^{n} |x_j - y_j|^2 \right)^{\frac{1}{2}}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We will postpone the proof that this d_2 indeed satisfies the properties of a metric to later. ✧

Again, for what follows in this chapter it is enough to only think of Example 2.3 (a), that is, *sequences in* $\mathbb{R}$ and with respect to the "canonical distance function" given by the absolute value, to digest the concepts. We emphasize that if we speak of "sequences in $\mathbb{R}$" without specifying further the metric, we consider the metric from that example, by convention. Yet, we will try to state results as general as possible. Note already a key difference between sequences in general metric spaces (X, d) and in $\mathbb{R}$: The set X may not have any structure allowing for 'adding elements in X', which means that it is not clear how to defined sums of sequences in X. In contrast, since $\mathbb{R}$ is a field, sequences in $\mathbb{R}$ can be added and multiplied in the canonical way of multiplying functions, i.e., for $x, y \in \mathbb{R}^{\mathbb{N}^*}$ define $x + y$ and $x \cdot y$ by

$$(x + y)(n) = x(n) + y(n) \quad \text{and} \quad (x \cdot y)(n) = x(n) \cdot y(n) \quad \text{for all} \quad n \in \mathbb{N}^*$$

or equivalently in the index notation

$$(x + y)_n = x_n + y_n \quad \text{and} \quad (x \cdot y)_n = x_n \cdot y_n \quad \text{for all} \quad n \in \mathbb{N}^*.$$

Definition 2.4 (Convergent-, Cauchy- and bounded sequence)**.** Let (X, d) be a metric space. A sequence $(x_n)_{n\in\mathbb{N}^*}$ is called

- *convergent*, if

$$\exists x \in X \ \forall \varepsilon > 0 \ \exists N \in \mathbb{N}^* \ \forall n \in \mathbb{N}^* \ : \ (n \geq N) \implies d(x_n, x) < \varepsilon;$$

- a *Cauchy* sequence, if

$$\forall \varepsilon > 0\, \exists N \in \mathbb{N}^*\, \forall m, n \in \mathbb{N}^* \,:\, (m, n \geq N) \implies d(x_n, x_m) < \varepsilon;$$

- *bounded*, if

$$\exists c \in X\, \exists M > 0\, \forall n \in \mathbb{N}^* \,:\, d(x_n, c) < M. \qquad ✧$$

Remark 2.5. Note that we can alternatively write the condition for convergence in terms of ε balls: The sequence $(x_n)_{n \in \mathbb{N}}$ converges to x, if and only if

$$\forall \varepsilon > 0\, \exists N \in \mathbb{N}^*\, \forall n \in \mathbb{N}^* \,:\, (n \geq N) \implies x_n \in \mathrm{B}_\varepsilon(x).$$

This is in particular for future generalizations a helpful perspective. ✧

Example 2.6. Let $(x_n)_{n \in \mathbb{N}^*}$ be defined by $x_n = \frac{1}{n}$ for all $n \in \mathbb{N}^*$. We claim that (x_n) converges to 0 in the metric space $(\mathbb{R}, d_2)$, where $d_2(x, y) = |x - y|$. To prove this from the definition, we need, for a given $\varepsilon > 0$ find an $N = N_\varepsilon \in \mathbb{N}^*$ such that $|\frac{1}{n} - 0| < \varepsilon$ for all $n \geq Nast$. This follows from Fact 1.28 (i) by the Archmedian principle. ✧

Proposition 2.7. *Let (X, d) be a metric space and let (x_n) be a sequence in X. Then the following holds.*

(i) (x_n) *converges* $\implies$ (x_n) *Cauchy sequence* $\implies$ (x_n) *bounded.*

(ii) *Limits are unique, i.e., $x, y \in X$ with $x_n \overset{n\to\infty}{\to} x$ and $x_n \overset{n\to\infty}{\to} y$, then $x = y$.*

(iii) $x_n \overset{n\to\infty}{\to} x$ *in* (X, d) $\iff$ $d(x_n, x) \overset{n\to\infty}{\to} 0$ *in* $\mathbb{R} = (\mathbb{R}, d_2)$.

Fact 2.8 (Facts on sequences in $\mathbb{R} = (\mathbb{R}, d_2)$)**.** *Let $x_n \to x$ and $y_n \to y$ in $\mathbb{R}$ (with standard metric d_2).*

(i) *(sums, products, quotients)*

- $x_n + y_n \to x + y$ *as* $n \to \infty$,
- $x_n \cdot y_n \to x \cdot y$ *as* $n \to \infty$,
- *if $y_n \neq 0$ for all $n \in \mathbb{N}^*$ and $y \neq 0$, then $\frac{x_n}{y_n} \to \frac{x}{y}$.*

(ii) *Let $x_n \to x$ and $y_n \to y$ as $n \to \infty$. If there exists $N \in \mathbb{N}^*$ such that*

- *$x_n \leq y_n$ for all $n \geq N$, then $x \leq y$;*
- *$x_n < y_n$ for all $n \geq N$, then $x \leq y$.*

(iii) *(Sandwich/Squeeze lemma): Let $(z_n)_{n \in \mathbb{N}^*}$ a sequence of real numbers. If there exists $N \in \mathbb{N}^*$ such that*

$$x_n \leq z_n \leq y_n \quad \textit{for all } n \in \mathbb{N}^*,$$

and $x_n \to x$ and $y_n \to x$ as $n \to \infty$, then $z_n \to x$ as $n \to \infty$.

(iv) *we point out that for all* $a, b \in \mathbb{R}$ *and* $\varepsilon > 0$,

$$|a-b| < \varepsilon \quad \iff \quad -\varepsilon < a-b < \varepsilon \quad \iff \quad b-\varepsilon < a < b+\varepsilon.$$

This equivalence is very useful when dealing with convergence proofs.

The following result answers the question whether Cauchy sequences of real numbers (considered in the standard metric d_2) converge.

Theorem 2.9 (Cauchy criterion, Completeness of $\mathbb{R}$)**.** *In the metric space* $\mathbb{R} = (\mathbb{R}, d_2)$ *a sequence is a Cauchy sequence if and only if the sequence converges.*

Proof. $\boxed{\Leftarrow}$: By Proposition 2.7, every convergence sequence is a Cauchy sequence.

$\boxed{\Rightarrow}$: Let $(x_n)_{n\in\mathbb{N}^*}$ be a Cauchy sequence in $\mathbb{R}$ with the metric d_2. By Proposition 2.7, we know already that the sequence is bounded. Therefore, we can proceed with the following definition: Let

$$a_n = \inf_{k\geq n} x_k := \inf\{x_k \mid k \geq n\},$$
$$b_n = \sup_{k\geq n} x_k := \sup\{x_k \mid k \geq n\},$$

for $n \in \mathbb{N}^*$. Since the sequence $(x_n)_{n\in\mathbb{N}^*}$ is bounded (as a Cauchy sequence), $(a_n)_{n\in\mathbb{N}^*}$ and $(b_n)_{n\in\mathbb{N}^*}$ are bounded sequences of real numbers. By construction, $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$ for all $n \in \mathbb{N}^*$. Therefore, the intervals $I_n = [a_n, b_n]$ are nested. By the nested interval principle, we conclude that there exists $x \in \bigcap_{n\in\mathbb{N}^*} I_n$, that is $x \in I_n$ for all $n \in \mathbb{N}^*$. We claim that $x_n \to x$ as $n \to \infty$. By definition of a_n and b_n, we have that

$$a_n \leq x_n \leq b_n, \quad \forall n \in \mathbb{N}^*,$$

where we used the definition of inf and sup. Therefore,

$$|x - x_n| < b_n - a_n \qquad \forall n \in \mathbb{N}^*$$

and even $|x - x_k| < b_n - a_n$ for all $k \geq n$. We finally show that for every $\varepsilon > 0$ there exists $N \in \mathbb{N}^*$ such that $b_N - a_N < 2\varepsilon$. This indeed follows from the property that $(x_n)_n$ is a Cauchy sequence as for every $\varepsilon > 0$ there exists $N \in \mathbb{N}^*$ such that for all $k \geq N$ (set $m = N$),

$$x_N - \varepsilon \leq x_k \leq x_N + \varepsilon.$$

This implies

$$x_N - \varepsilon \leq \inf_{k\geq N} x_k = a_N \leq b_N = \sup_{k\geq N} x_k \leq x_N + \varepsilon,$$

which gives $b_N - a_N < 2\varepsilon$. ❑

Remark 2.10. Consider the metric space $(\mathbb{C}, d_2)$, where $d_2(x, y) = |x - y|$ and $|\cdot|$ denotes the absolute value of a complex number. Denoting by $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ the real and imaginary parts of $z \in \mathbb{C}$, we have by the inequality

$$\max\{|\mathrm{Re}(z)|, |\mathrm{Im}(z)|\} \leq |z| \leq \sqrt{2}\max\{|\mathrm{Re}(z)|, |\mathrm{Im}(z)|\}, \quad z \in \mathbb{C},$$

that a sequence of complex numbers $(z_n)_{n\in\mathbb{N}^*}$ converges (is a Cauchy sequence) in $(\mathbb{C}, d_2)$ if and only if both the sequences $(\mathrm{Re}(z_n))_{n\in\mathbb{N}^*}$ and $(\mathrm{Im}(z_n))_{n\in\mathbb{N}^*}$ converge (are Cauchy sequences) in $(\mathbb{R}, d_2)$. Therefore, by Theorem 2.9, $(\mathbb{C}, d_2)$ is *complete*, that is, a sequence converges if and only if it is a Cauchy sequence. ✧

Example 2.11. The metric space $(\mathbb{Q}, d_2)$ is not complete because there exist sequences (*why?*) in $\mathbb{Q}$ which are Cauchy sequences, but not convergent. ✧

2.1 Monotonic sequences and Euler's constant

Definition 2.12. A sequence (x_n) in $\mathbb{R}$ is called

- *decreasing/increasing*, if for all $n \in \mathbb{N}^*$, $x_n \geq x_{n+1}$ or $x_n \leq x_{n+1}$.
- *monotonic* if the sequence is either increasing or decreasing.
- *strictly decreasing/increasing*, if (x_n) is decreasing/increasing and the inequalities are strict for all n.
- *bounded above/below*, if the set $\{x_n \mid n \in \mathbb{N}^*\}$ is bounded above/below.

✧

The above definition can (and will later) given more generally for functions $f\colon S \to \mathbb{R}$ where $S \subseteq \mathbb{R}$.

Theorem 2.13 (monotone convergence in $\mathbb{R}$)**.** *Every increasing/decreasing sequence in $\mathbb{R}$, which is bounded above/below converges to the supremum/infimum of $\{x_n \mid n \in \mathbb{N}^*\}$.*

Proof. Assume that $\{x_n \mid n \in \mathbb{N}^*\}$ is increasing and bounded above and let $s \in \mathbb{R}$ be the supremum. Then, by the definition of the supremum and the increasing property for every $\varepsilon > 0$ there exists $N \in \mathbb{N}^*$ such that

$$x_n \leq s < x_n + \varepsilon$$

for all $n > N$. Rearranging terms in the inequalities gives

$$0 \leq s - x_n < \varepsilon, \quad n > N.$$

But this implies that $x_n \to s$ as $n \to \infty$.

The case of a decreasing sequence which is bounded below follows analogously or by multiplying the sequence by -1. ❑

Definition 2.14. A sequence $(x_n)_{n\in\mathbb{N}^*}$ in $\mathbb{R}^{\mathbb{N}^*}$ is said to *converge to* $+\infty$, if for every $c > 0$ there exists $N \in \mathbb{N}^*$ such that for all $n \geq N$ it holds that $x_n > c$. Analogously, we say that $(x_n)_{n\in\mathbb{N}^*}$ *converges to* $-\infty$ if $(-x_n)_{n\in\mathbb{N}^*}$ converges to $+\infty$. ✧

> **Convention.** We will always specifically indicate if we mean "convergence to $\pm\infty$". In particular, if we say "$(x_n)_{n\in\mathbb{N}^*}$ converges", we will **not** mean that the sequence can converge to $\pm\infty$.

Proposition 2.15 (Bernoulli inequality)**.** *Let $x > -1$ and $n \in \mathbb{N}^*$. Then*

$$(1+x)^n \geq 1 + nx.$$

Proof. See Exercise 2.6. ❑

Example 2.16 (Euler's number)**.** The sequence $x_n = (1+\frac{1}{n})^n$, $n \in \mathbb{N}^*$ converges as we shall see below. Indeed, by using Bernoulli's inequality, we can show that the sequence $(y_n)_{n\in\mathbb{N}^*}$ given by

$$y_n = \left(1+\frac{1}{n}\right)^{n+1}, \quad n \in \mathbb{N}^*,$$

is decreasing, see Exercise 2.11 (a). Since the sequence is also bounded below (e.g., by 0), Theorem 2.13 implies that $y_n \to y$ as $n \to \infty$ for some $y \in \mathbb{R}$. We now conclude that also $x_n \to y$ since $x_n = y_n(1+\frac{1}{n})^{-1}$, $n \in \mathbb{N}^*$ and $(1+\frac{1}{n})^{-1} \to 1$ as $n \to \infty$. We call the limit *Euler's number* and denote it by e, i.e.,

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

2.2 Subsequences, Limit superior and Limit inferior

Definition 2.17. Let $x = (x_n)_{n\in\mathbb{N}^*}$ be a sequence in a set X and $k\colon \mathbb{N}^* \to \mathbb{N}^*$ be a strictly increasing function, that is, $k(n) < k(n+1)$ for all $n \in \mathbb{N}^*$ (equivalently, we can say that k is a strictly increasing sequence in $\mathbb{N}^*$). Then the function

$$x \circ k = (x_{k(n)})_{n\in\mathbb{N}^*} = (x_{k_n})_{n\in\mathbb{N}^*}$$

is called *a subsequence of* x. ✧

Example 2.18. Let us consider two simple instances of subsequences.

(a) Let $x_n = \frac{1}{n}$, $n \in \mathbb{N}^*$. Then

$$k(n) = 2n+1, \qquad \text{and} \qquad \tilde{k}(n) = 2n, \quad n \in \mathbb{N}^*,$$

define the subsequences $(x_{k(n)})_{n\in\mathbb{N}^*} = (\frac{1}{2n+1})_{n\in\mathbb{N}^*}$ and $(x_{\tilde{k}(n)})_{n\in\mathbb{N}^*} = (\frac{1}{2n})_{n\in\mathbb{N}^*}$.

(b) Let $x_n = (-1)^n$, $n \in \mathbb{N}^*$. Then the subsequences $(x_{k(n)})_{n\in\mathbb{N}^*}$ defined by

$$k(n) = 2n + 1, \qquad \text{and} \qquad \tilde{k}(n) = 2n, \quad n \in \mathbb{N}^*,$$

are the constant -1 and constant 1 sequence, respectively, and therefore converge—in contrast to the original sequence. ✧

Fact 2.19. *Let (X, d) be a metric space.*

(i) *A sequence in (X, d) converges if and only if every subsequence converges. In that case all limits coincide (think about* why?*).*

(ii) *A subsequence can converge although the original sequence does not converge.*

(iii) *A sequence that converges to $\pm\infty$ does not converge.*

Proposition 2.20. *Every sequence of real numbers has a monotonic subsequence.*

Proof. See Exercise 2.21 ❑

Theorem 2.21 (Bolzano–Weierstraß)**.** *Every bounded sequence in $\mathbb{R} = (\mathbb{R}, d_2)$ has a convergent subsequence.*

Proof. By Proposition 2.20 every sequence has a monotonic subsequence. Since, by assumption, the sequence was bounded, the subsequence is bounded as well. Thus, by monotone convergence, Theorem 2.13, the subsequence converges is convergent. ❑

For a sequence $(x_n)_{n\in\mathbb{N}^*}$ of real numbers, define the sequence $(z_k)_{k\in\mathbb{N}^*}$ by

$$z_k := \sup\{x_n \mid n \geq k\}, \quad k \in \mathbb{N}^*,$$

where we allow for the "value" $+\infty$. The sequence is decreasing, that is $z_k \geq z_{k+1}$ for all $k \in \mathbb{N}^*$ (with the convention that $c \leq \infty$ for every $c \in \mathbb{R} \cup \{\infty\}$), as the supremum in the definition of z_{k+1} is taken over a subset of the corresponding one for z_k. Therefore, by Theorem 2.13, the sequence $(z_k)_{k\in\mathbb{N}^*}$ either

- converges (with respect to the standard metric) to a real number, or
- converges to $-\infty$ (with respect to the standard metric), or
- is constant $+\infty$, that is, $z_k = \infty$ for all $k \in \mathbb{N}^*$.

This (and the analogous considerations for $z_k = \inf\{x_n \mid n \geq k\}$) justifies the following definitions.

Definition 2.22 (Limit superior and limit inferior)**.** Given a sequence $(x_n)_{n\in\mathbb{N}^*}$ in $\mathbb{R}$ (with the standard metric). Let

$$z_k := \sup_{n\geq k} x_n := \sup\{x_n \mid n \geq k\} \in \mathbb{R} \cup \{\infty\}, \quad k \in \mathbb{N}^*.$$

The limit (including the value $+\infty$) of $(z_k)_{k\in\mathbb{N}^*}$ is called *limit superior* of the sequence $(x_n)_{n\in\mathbb{N}^*}$ and is denoted by

$$\limsup_{n\to\infty} x_n \coloneqq \lim_{k\to\infty} z_k \in \mathbb{R}\cup\{-\infty,+\infty\}$$

Analogously, we define the *limit inferior* of $(x_n)_{n\in\mathbb{N}^*}$ as

$$\liminf_{n\to\infty} x_n \coloneqq \lim_{k\to\infty} \inf_{n\geq k} x_n \coloneqq \lim_{k\to\infty} \inf\{x_n \mid n\geq k\} \in \mathbb{R}\cup\{-\infty,+\infty\}. \quad ✧$$

Example 2.23. Let us regard the following examples for the limit superior and inferior.

(a) Let $x_n = (-1)^n$ for $n\in\mathbb{N}^*$. Then

$$z_k = \sup_{n\geq k} x_n = 1,$$

for all $k\in\mathbb{N}^*$ and thus $\limsup_{n\to\infty} x_n = 1$. Similarly, we see that $\liminf_{n\to\infty} x_n = -1$. Recall that $(x_n)_{n\in\mathbb{N}^*}$ is not convergent. Also note that we could change finitely many components of the sequence without affecting the results for the limit superior and limit inferior.

(b) Let $(x_n)_{n\in\mathbb{N}^*}$ be a convergent series in $\mathbb{R}$ (with the standard metric) with limit $x\in\mathbb{R}$. Then $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = x$, as we shall see in Theorem 2.25 below. ✧

Definition 2.24. Let (X,d) be a metric space and let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in X. We say that $x\in X$ is a *partial limit* of $(x_n)_{n\in\mathbb{N}^*}$ if there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}^*}$ converging to x.
If $(X,d) = (\mathbb{R},d_2)$, we also say that $x\in\{-\infty,+\infty\}$ is a partial limit if there exists a subsequence converging to $-\infty$ or $+\infty$ respectively. ✧

Theorem 2.25. *Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathbb{R}$ and $L\in\mathbb{R}\cup\{-\infty,\infty\}$. Then*

(i) *$L = \limsup_{n\to\infty} x_n$ if and only if L is the largest partial limit of $(x_n)_{n\in\mathbb{N}^*}$.*

(ii) *$L = \liminf_{n\to\infty} x_n$ if and only if L is the smallest partial limit of $(x_n)_{n\in\mathbb{N}^*}$.*

(iii) *We have that*

$$\inf_{n\in\mathbb{N}^*} x_n \leq \liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n \leq \sup_{n\in\mathbb{N}^*} x_n.$$

Furthermore, $x_n\to L$ as $n\to\infty$ if and only if $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = L$.

Proof. We show (ii) for $L\in\mathbb{R}$, the rest is self-study (see also [13, pages 92–94]). Let $L = \liminf_{n\to\infty} x_n$. We aim to find a subsequence of $(x_n)_{n\in\mathbb{N}^*}$

which converges to L. By definition of the infimum and the fact that $L = \lim_{k\to\infty} \inf_{n\geq k} x_n$, we find a strictly increasing sequence of natural numbers, that is a function $k\colon \mathbb{N}^* \to \mathbb{N}^*$, such that

$$\inf_{n\geq k(l)} x_n \leq x_{k(l)} \leq \inf_{n\geq k(l)} x_n + \frac{1}{l}, \qquad l \in \mathbb{N}^*.$$

Since the left and the right-hand side converges to L as $l \to \infty$, we have by the squeeze lemma that $(x_{k(l)})_{l\in\mathbb{N}^*}$ converges to L. We have thus proved that L is a partial limit. To show it is the smallest partial limit, suppose that there exists a subsequence $(x_{m(l)})_{l\in\mathbb{N}^*}$ converging to $\tilde{L} < L$. Therefore, there exists $\varepsilon > 0$ (choose $\varepsilon = \frac{1}{2}(L - \tilde{L})$) and $N \in \mathbb{N}^*$ such that

$$\forall l > N \,:\, L - x_{n_l} > \varepsilon.$$

Therefore, for every $k \in \mathbb{N}^*$ there exists $n \geq k$ such that $x_n < L - \varepsilon$. Thus also $\inf_{n\geq k} x_n \leq L - \varepsilon$ for all $k \in \mathbb{N}^*$, which implies that $L := \lim_{k\to\infty} \inf_{n\geq k} x_n \leq L - \varepsilon$, which is a contradiction. ❑

2.3 Series

Definition 2.26. Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathbb{R}$. The sequence $(s_n)_{n\in\mathbb{N}^*}$ defined by

$$s_n = \sum_{j=1}^{n} a_j := a_1 + a_2 + \cdots + a_n, \qquad n \in \mathbb{N}^*,$$

is called *a series of real numbers* and short-hand denoted by $\sum_{j=1}^{\infty} a_j$ (this is just a formal notation at this moment) and s_n is called the *n-th partial sum of the series.* We say that $\sum_{j=1}^{\infty} a_j$ *converges* if the sequence $(s_n)_{n\in\mathbb{N}^*}$ converges (to a real number)[4]. The limit, if exists, is also denoted by $\sum_{j=1}^{\infty} a_j$ and also called *infinite sum*[5]. If the series does not converge, we say that it *diverges.* The a_j's are called *terms of the series.* ✧

More generally, we can define series on vector spaces and define their convergence if the vector space is metric. We will however, focus on a special case of metrics in that case; normed spaces, which are special metric spaces, see the tutorials.

Definition 2.27 (Series in normed vector spaces)**.** Let $(X, \|\cdot\|)$ be a normed vector space (over $\mathbb{R}$ or $\mathbb{C}$). Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in X. The sequence $(s_n)_{n\in\mathbb{N}^*}$ defined by

$$s_n = \sum_{j=1}^{n} a_j := a_1 + a_2 + \cdots + a_n, \qquad n \in \mathbb{N}^*,$$

[4]As mentioned before, this excludes the case that (s_n) converges to $\pm\infty$

[5]thus, we write $\sum_{j=1}^{\infty} a_j$ for both the formal object of a series as well as its limit. While these are two different objects, this won't typically lead to confusion.

is called *a series* in X and short-hand denoted by $\sum_{j=1}^{\infty} a_j$ (this is just a formal notation at this moment) and s_n is called the *n-th partial sum of the series.* We say that $\sum_{j=1}^{\infty} a_j$ *converges* if the sequence $(s_n)_{n\in\mathbb{N}^*}$ converges to an element $s \in X$. The limit, if exists, is also denoted by $s = \sum_{j=1}^{\infty} a_j$ and is also called *infinite sum.* ✧

A particularly important example of the above definition is the field of complex numbers which becomes a metric space with the metric $d(x, y) = |x - y|$, $x, y \in \mathbb{C}$. Even more, $(\mathbb{C}, |\cdot|)$ is a normed space.

Fact 2.28. *Note that a series is just the sequence of partial sums $(s_n)_{n\in\mathbb{N}^*}$. Hence, we can apply the results that we have already derived for sequences.*

(i) *If $\sum_{j=1}^{\infty} a_j$ converges, then the sequence $(a_j)_{j\in\mathbb{N}^*}$ of its terms converges to 0.*

Pf.: Since $a_j = s_j - s_{j-1}$ for $j > 1$, $a_j \overset{j\to\infty}{\to} 0$ since (s_j) is convergent and hence a Cauchy sequence.

(ii) *Let $a, \tilde{a}$ be sequences in $\mathbb{R}^{\mathbb{N}^*}$ such that $\{k \in \mathbb{N}^* \mid a_k \neq \tilde{a}_k\}$ is finite. Then $\sum_{j=1}^{\infty} a_j$ converges if and only if the series $\sum_{j=1}^{\infty} \tilde{a}_j$ converges. However, the infinite sums/limits need not be the same.*

Pf.: Let $K = \max\{k \in \mathbb{N}^* \mid a_k \neq \tilde{a}_k\}$. Then, for $n \geq K$,

$$\sum_{j=1}^{n} a_j = \sum_{\ell=1}^{K} \tilde{a}_\ell + \sum_{\ell=K+1}^{n} \tilde{a}_\ell = \left(\sum_{\ell=1}^{K} \tilde{a}_\ell - \sum_{\ell=1}^{K} a_\ell\right) + \sum_{\ell=1}^{n} \tilde{a}_\ell,$$

which shows that the convergence of the partial sums $(s_n = \sum_{j=1}^{n} a_j)_n$ and $(\tilde{s}_n = \sum_{\ell=1}^{n} \tilde{a}_\ell)_n$ is equivalent as well as that the limits differ by $\sum_{j=1}^{K}(a_j - \tilde{a}_j)$, Fact 2.8.

(iii) *A series $\sum_{j=1}^{\infty} a_j$ in $(\mathbb{R}, d_2)$ converges if and only if*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}^*\ \forall m, n \geq N \text{ with } m \geq n\ :\ \left|\sum_{j=n}^{m} a_j\right| < \varepsilon.$$

Pf.: Since $s_m - s_{n-1} = \sum_{j=n}^{m} a_j$, $m \geq n > 1$, this assertion follows since $(\mathbb{R}, d_2)$ is complete, Theorem 2.9.

(iv) *Let $(a_j)_{j\in\mathbb{N}^*}$ be a sequence of non-negative numbers. Then $\sum_{j=1}^{\infty} a_j$ converges if and only if the sequence of partial sums $(s_n = \sum_{j=1}^{n} a_j)_{n\in\mathbb{N}^*}$ is bounded above.*

Pf.: Since $a_n \geq 0$, it follows that $s_{n+1} = s_n + a_{n+1} \geq s_n$, for every $n \in \mathbb{N}^*$. Thus, $(s_n)_n$ is increasing and hence converges (to a real number) if and only if the sequence is bounded, by Theorem 2.13 and Fact 2.19.

(v) *(**Comparison Lemma**) Let $(a_j)_{j\in\mathbb{N}^*}$ and $(b_j)_{j\in\mathbb{N}^*}$ be real sequences such that $0 \leq a_j \leq b_j$ for all but finitely many $j \in \mathbb{N}^*$. If $\sum_{j=1}^{\infty} b_j$*

converges, then $\sum_{j=1}^{\infty} a_j$ *converges. On the other hand, if* $\sum_{j=1}^{\infty} a_j$ *does not converge, then* $\sum_{j=1}^{\infty} b_j$ *cannot converge either.*

Pf.: By (ii), we can assume that $0 \le a_j \le b_j$ *for all* $j \in \mathbb{N}^*$. *If* $\sum_j b_j$ *converges, then, by (iv), the associated partial sums are bounded above. Since* $\sum_{j=1}^{n} a_j \le \sum_{j=1}^{n} b_j$, *this implies that also the partial sums of* $\sum_j a_j$ *are bounded above, thus convergent by (iv) again.*

The second statement follows analogously.

Example 2.29. The following are standard series that will reappear frequently.

(a) Let $q \in \mathbb{R}$. By the formula (convention $q^0 = 1$).

$$\sum_{j=0}^{n} q^j = \frac{1 - q^{n+1}}{1 - q}, \qquad n \in \mathbb{N}^*,$$

(see Exercise 2.29), we conclude that *geometric series* $\sum_{j=0}^{\infty} q^j$ converges if and only if $|q| < 1$. In particular, (if $|q| < 1$) we have

$$\sum_{j=0}^{\infty} q^j = \lim_{n \to \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}.$$

Furthermore, for any $m \in \mathbb{N}^*$ and $|q| < 1$, we have

$$\sum_{j=m}^{\infty} q^j = \sum_{j=0}^{\infty} q^j - \sum_{j=0}^{m-1} q^j = \frac{1}{1-q} - \frac{1 - q^m}{1 - q} = \frac{q^m}{1 - q}.$$

(b) The *harmonic series* $\sum_{j=1}^{\infty} \frac{1}{j}$ does not converge, since the partial sums are not a Cauchy sequences.

(c) $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges by the comparison lemma and $\frac{1}{n^2} < \frac{1}{(n-1)n} = \frac{1}{n-1} - \frac{1}{n}$ for $n > 1$, whence

$$\begin{aligned}\sum_{n=2}^{N} \frac{1}{(n-1)n} &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{N-1} - \frac{1}{N} \\ &= 1 - \frac{1}{N} \overset{N\to\infty}{\to} 1.\end{aligned}$$

✧

Definition 2.30. A series $\sum_{j=1}^{\infty} a_j$ of real (or complex) numbers is called absolutely convergent if the series

$$\sum_{j=1}^{\infty} |a_j|$$

converges. We also write $\sum_{j=1}^{\infty} |a_j| < \infty$ in that case. ✧

Proposition 2.31. *Any absolutely convergent series $\sum_{j=1}^{\infty} a_j$ of real (or complex) numbers converges.*

Proof. Since by the triangle inequality (applied inductively),

$$\left|\sum_{j=n}^{m} a_j\right| \leq \sum_{j=n}^{m} |a_j|,$$

Fact 2.28 (iii) (the completeness of $(\mathbb{R}, d_2)$ or $(\mathbb{C}, d_2)$, see Theorem 2.9 and the subsequent remark) implies the assertion. ❑

Theorem 2.32 (Root test for absolute convergence)**.** *Let $\sum_{j=1}^{\infty} a_j$ be a real (or complex) series. Let $L := \limsup_{j\to\infty} \sqrt[n]{|a_n|}$.*

(i) *If $L < 1$, then $\sum_{j=1}^{\infty} a_j$ converges absolutely.*
In other words, if $\exists q \in (0, 1)$ and $N \in \mathbb{N}^*$ such that $|a_n| \leq q^n$ for all $n \geq N$, then $\sum_{j=1}^{\infty} |a_j| < \infty$.

(ii) *If $L > 1$, then $\sum_{j=1}^{\infty} a_j$ does not converge.*
More generally, if there exists a subsequence $(a_{k_n})_{n\in\mathbb{N}^*}$ such that $|a_{k_n}| \geq 1$ for all $n \in N^*$, then $\sum_{j=1}^{\infty} a_j$ does not converge.

Proof. Let $L < 1$. By the definition of the limit superior, there exist $q \in (L, 1)$ and $N \in \mathbb{N}^*$ such that

$$\forall n \geq N : \sqrt[n]{|a_n|} \leq q \iff |a_n| \leq q^n.$$

Therefore, the comparison lemma, Fact 2.28 (v), and Example 2.29 (a) imply that $\sum_{j=1}^{\infty} |a_n|$ converges.

Let $L > 1$. Then, by Theorem 2.25, there exists a subsequence $(a_{k_n})_{n\in\mathbb{N}^*}$ such that $\sqrt[k_n]{|a_{k_n}|} \geq q > 1$ for all $n \in \mathbb{N}^*$. Thus, by the comparison lemma and Example 2.29 (a), we conclude that the series does not converge. ❑

Corollary 2.33 (Ratio test for absolute convergence)**.** *Let $\sum_{j=1}^{\infty} a_j$ be a real (or complex) series with all but finitely many terms non-zero.*

(i) *If $\limsup_{j\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$, then $\sum_{j=1}^{\infty} a_j$ converges absolutely.*

(ii) *If $\liminf_{j\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$, then $\sum_{j=1}^{\infty} a_j$ does not converge.*

Proof. This can be proved by the root test, Theorem 2.32. see also [13]. ❑

2.4 Alternating sequences

While any absolutely converging sequence of real numbers converges, the converse may not be true, as we shall see shortly. Yet, previoulsy seen tests for convergence of sequences, mostly concerned absolute convergence. We now study a specific types of sequences for which we may conclude convergence even if they fail to be absolutely convergent; the Dirichet and Leibniz tests, see below.

Example 2.34. Consider the series
$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k}.$$
We have already seen in Example 2.29 that this series is not absolutely convergent, since the harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ does not converge. Nevertheless, the series converges as we will see in the following. Note that
$$\sum_{k=1}^{2n} \frac{(-1)^k}{k} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \cdots = \sum_{k=1}^{n} \frac{-1}{2k-1} + \frac{1}{2k} = -\sum_{k=1}^{n} \frac{1}{(2k-1)2k}$$
and every $\frac{1}{(2k-1)2k} > 0$. Moreover,
$$\sum_{k=1}^{n} \frac{1}{(2k-1)2k} \leq \sum_{k=2}^{2n} \frac{1}{(k-1)k} \leq \sum_{k=2}^{\infty} \frac{1}{(k-1)k} \overset{\text{Ex.2.29(c)}}{=} 1$$
By monotony and boundedness, we can conclude that $-\sum_{k=1}^{n} \frac{(-1)^k}{k}$ converges for $n \to \infty$. ✧

Lemma 2.35 (Summation by parts)**.** *Let $m \in \mathbb{N}^*$, $m > 1$. Let $a_1, \dots, a_m$ and $b_1, \dots, b_m$ be real numbers (or complex numbers). Then for every $n \in \mathbb{N}^*$ with $n < m$,*
$$\sum_{k=n}^{m} a_k b_k = a_m \sum_{k=n}^{m} b_k - \sum_{j=n}^{m-1} (a_{j+1} - a_j) \sum_{k=n}^{j} b_k.$$

Proof. It suffices to consider the case $n = 1$. The general case for $n > 1$ follows by setting $a_1 = \cdots = a_{n-1} = b_1 = \cdots = b_{n-1} = 0$.

Note that
$$\begin{aligned}
\sum_{j=1}^{m-1} (a_{j+1} - a_j) \sum_{k=1}^{j} b_k &= \sum_{j=1}^{m-1} a_{j+1} \sum_{k=n}^{j} b_k - \sum_{j=1}^{m-1} a_j \sum_{k=1}^{j} b_k \\
&= \sum_{j=2}^{m} a_j \sum_{k=1}^{j-1} b_k - \sum_{j=1}^{m-1} a_j \sum_{k=1}^{j} b_k \\
&= \sum_{j=2}^{m-1} a_j \sum_{k=1}^{j} b_k + a_m \sum_{k=1}^{m-1} b_k - \sum_{j=2}^{m-1} a_j b_j - \sum_{j=1}^{m-1} a_j \sum_{k=1}^{j} b_k \\
&= a_m \sum_{k=1}^{m-1} b_k - a_1 b_1 - \sum_{j=2}^{m-1} a_j b_j \\
&= a_m \sum_{k=1}^{m} b_k - \sum_{j=1}^{m} a_j b_j,
\end{aligned}$$
which shows the desired identity for $n = 1$. ❑

Theorem 2.36 (Dirichlet test)**.** *Let $(a_k)_{k\in\mathbb{N}^*}$ be a strictly decreasing sequence in $\mathbb{R}$ converging to 0 and let $(b_k)_{k\in\mathbb{N}^*}$ be a sequence in $\mathbb{R}$ such that there exists $C > 0$ with*

$$\left|\sum_{k=1}^{n} b_k\right| \leq C \quad \textit{for all} \quad n \in \mathbb{N}^*.$$

Then the series $\sum_{k=1}^{\infty} a_k b_k$ converges.

Proof. Let $n < m$ be in $\mathbb{N}^*$. By Lemma 2.35 and triangle inequality (applied several times), we have that

$$\begin{aligned}\left|\sum_{k=n}^{m} a_k b_k\right| &\leq |a_m|\left|\sum_{k=n}^{m} b_k\right| + \left|\sum_{j=n}^{m-1}(a_{j+1}-a_j)\sum_{k=n}^{j} b_k\right| \\ &\leq a_m\left|\sum_{k=n}^{m} b_k\right| + \sum_{j=n}^{m-1}(a_j-a_{j+1})\left|\sum_{k=n}^{j} b_k\right|,\end{aligned} \tag{2.1}$$

where we used that $|a_{j+1}-a_j| = a_j - a_{j+1}$ and that $a_m \geq 0$ by the assumption that $(a_k)_{k\in\mathbb{N}^*}$ is decreasing and converging to 0. Note that the sums over the b_k's can be estimated by

$$\left|\sum_{k=n}^{j} b_k\right| = \left|\sum_{k=1}^{j} b_k - \sum_{k=1}^{n-1} b_k\right| \leq \left|\sum_{k=1}^{j} b_k\right| + \left|\sum_{k=1}^{n-1} b_k\right| \leq 2C,$$

for all $j \geq n > 1$. Therefore, we get with (2.1) that

$$\left|\sum_{k=n}^{m} a_k b_k\right| \leq a_m 2C + 2C\sum_{j=n}^{m-1}(a_j - a_{j+1}) = 2C(a_m + a_n - a_m) = 2Ca_n$$

Since $a_n \to 0$ as $n \to \infty$, we find for every $\varepsilon > 0$ an $N \in \mathbb{N}^*$ such that for all $n \in \mathbb{N}^*$: $2Ca_n < \varepsilon$. By Fact 2.28 (iii), this shows that the series converges. ❑

Corollary 2.37 (Leibniz test)**.** *Let $(\tilde{a}_k)_{k\in\mathbb{N}^*}$ be a decreasing (or increasing) sequence in $\mathbb{R}$ converging to 0. Then the series*

$$\sum_{k=1}^{\infty}(-1)^k \tilde{a}_k$$

converges.

Proof. Clearly, $\left|\sum_{k=1}^{n}(-1)^k\right| \leq 1$ Thus, we can apply the Dirichlet test, Theorem 2.36, for $b_k = (-1)^k$ and $a_k = \tilde{a}_k$ (and $a_k = -\tilde{a}_k$ if $(\tilde{a}_k)_k$ is increasing), $k \in \mathbb{N}^*$. ❑

2.5 Exercises

Exercise 2.1. Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, i.e., X is a $\mathbb{F}$-vector space and $\|\cdot\| \colon X \to [0, +\infty)$ satisfies:

(i) $\|x\| = 0 \iff x = 0$,

(ii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$,

(iii) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{F}$ and $x \in X$.

Show that $d \colon X \times X \to [0, +\infty)$ defined by

$$d(x, y) := \|x - y\|$$

is a metric on X. ✧

Exercise 2.2. Show that $(\mathbb{C}, |\cdot|)$ is a normed space, where $|\cdot|$ refers to the modulus (absolute value) in $\mathbb{C}$. ✧

Exercise 2.3. Discuss the convergence of the following sequences of real numbers (with the canonical metric) by the definition of convergence:

(a) $(n)_{n\in\mathbb{N}^*}$

(b) $(\frac{1}{n^2})_{n\in\mathbb{N}^*}$

(c) $(x_n)_{n\in\mathbb{N}^*}$, where $x_n := \frac{n-1}{n}$, for all $n \in \mathbb{N}^*$.

(d) $(x_n)_{n\in\mathbb{N}^*}$, where $x_n := (-1)^n$, for all $n \in \mathbb{N}^*$. ✧

Exercise 2.4. Let $(x_n)_{n\in\mathbb{N}^*}$ be convergent sequence in $\mathbb{R}$ with limit x such that $x_n, x \geq 0$ for all $n \in \mathbb{N}^*$. Show that $\lim_{n\to\infty} \sqrt{x_n} = \sqrt{x}$. ✧

Exercise 2.5. Prove the following for sequences in $\mathbb{R}$ with the canonical metric.

(a) $\lim\limits_{n\to\infty} \left(\sqrt{n^2 + 1} - n\right) = 0.$

(b) $\lim\limits_{n\to\infty} \left(\sqrt{4n^2 + n} - 2n\right) = \dfrac{1}{4}.$

(c) $\lim\limits_{n\to\infty} \left(\sqrt{n^2 + n} - n\right) = \dfrac{1}{2}.$ ✧

Exercise 2.6. Let $z \in \mathbb{R}$ with $z > -1$ and $n \in \mathbb{N}^*$. Show that

$$(1 + z)^n \geq 1 + nz.$$

This is known as *Bernoulli's inequality.* ✧

Exercise 2.7. Let (X,d) be a metric space. Show that a sequence $(x_n)_{n\in\mathbb{N}^*}$ in X converges to $x \in X$, if and only if the sequence $(y_n)_{n\in\mathbb{N}^*}$ defined by $y_n := d(x_n, x)$ converges to 0 in $\mathbb{R}$. ✧

Exercise 2.8. Show the convergence of the following sequence in $\mathbb{R}$ and find the limit.

(a) $\left(\sqrt{n^2+n} - n\right)_{n\in\mathbb{N}^*}$

(b) $(q^n)_{n\in\mathbb{N}^*}$ for $q \in (0,1]$ (Hint: Bernoulli's inequality Exercise 2.6). ✧

Exercise 2.9. Is the function ρ given by $\rho(x,y) := \left|x^2 - y^2\right|$ a metric on

(a) $\mathbb{R}$?

(b) $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$? ✧

Exercise 2.10. Let $x_1 := 2$ and define $x_{n+1} := x_n^2 - x_n + 1$, for all $n \in \mathbb{N}^*$. Show that the sequence $(y_n)_{n\in\mathbb{N}^*}$ defined by

$$y_n := \frac{x_1 \cdot \ldots \cdot x_n}{x_{n+1}}, \quad n \in \mathbb{N}^*,$$

converges to 1.

Hint: Compute x_1, x_2, x_3, x_4 and make a claim about the form of the fraction y_n. ✧

Exercise 2.11. Let us regard the following sequences in $\mathbb{R}$ with the standard metric.

(a) Use Bernoulli's inequality (see Exercise 2.6) to show that the sequence $(x_n)_{n\in\mathbb{N}^*}$,

$$x_n := \left(1 + \frac{1}{n}\right)^{n+1}, \quad n \in \mathbb{N}^*,$$

is *strictly decreasing*, i.e., $\forall n \in \mathbb{N}^* : x_n > x_{n+1}$.

(b) Show that $\mathrm{e} := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ exists. ✧

Exercise 2.12. Let X be a set and let $d_0 \colon X \times X \to \{0,1\}$ be given by

$$d_0(x,y) := \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

(a) Show that (X, d_0) indeed defines a metric space.

(b) Characterize all convergent sequences in (X, d_0).

(c) Characterize all Cauchy sequences in (X, d_0). ✧

Exercise 2.13. Let d be a metric on $\mathbb{R}^2$ and let $d_P\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be defined as follows (we use the notation $x = (x_1, x_2), y = (y_1, y_2)$ for $x, y \in \mathbb{R}^2$)

$$d_P(x,y) := \begin{cases} d(x,y), & \text{if } \exists \lambda \in \mathbb{R} \,:\, \lambda x = y \text{ or } x = \lambda y, \\ d(x,0) + d(0,y), & \text{otherwise.} \end{cases}$$

Show that $(\mathbb{R}^2, d_P)$ is a metric space. ✧

Exercise 2.14. Let X be a set. Show that, if ρ is a metric on X, then so is $\sigma\colon X \times X \to \mathbb{R}$ given by

$$\sigma(x,y) := \frac{\rho(x,y)}{1+\rho(x,y)}.$$

Moreover, show that a sequence $(x_n)_{n\in\mathbb{N}^*}$ converges to x in (X,ρ) if and only if it converges to x in (X,σ). ✧

Exercise 2.15. Let $h \in \mathbb{R}$. Determine for what h the sequence $(n!h^n)_{n\in\mathbb{N}^*}$ converges to a real number, to infinity or does not converges at all. ✧

Exercise 2.16. Prove from the definition that each of the following sequences converges to $+\infty$ or to $-\infty$.

(a) $(n^2 - n)_{n\in\mathbb{N}^*}$

(b) $(n - 3n^2)_{n\in\mathbb{N}^*}$

(c) $(\frac{n^2+1}{n})_{n\in\mathbb{N}^*}$ ✧

Exercise 2.17. Show that:

(a) $\lim\limits_{n\to\infty} \dfrac{n}{q^n} = 0$ for $q > 1$.

(b) $\lim\limits_{n\to\infty} n^{\frac{1}{n}} = 1$.

(c) $\lim\limits_{n\to\infty} a^{1/n} = 1$ for $a > 0$. ✧

Exercise 2.18. Show that the metric space $(\mathbb{Q}, d)$, where $d(x,y) := |x-y|$, is not complete. *Casually, we do not specify the metric and just say $\mathbb{Q}$ is not complete (incomplete).* ✧

Exercise 2.19. Find the limit inferior and the limit superior of each of the following sequences.

(a) $(x_n)_{n\in\mathbb{N}^*}$ where $x_n := 3 - (-1)^n$

(b) $(x_n)_{n\in\mathbb{N}^*}$ where $x_n := (-1)^{n+1} + (-1)^n/n$

(c) $(x_n)_{n\in\mathbb{N}^*}$ where $x_n := \sqrt{1+n^2}/(2n-5)$

(d) $x_n = y_n/n$, where $\{y_n\}$ is any bounded sequence

(e) $(x_n)_{n\in\mathbb{N}^*}$ where $x_n := n\,(1+(-1)^n) + n^{-1}\,((-1)^n - 1)$

(f) $(x_n)_{n\in\mathbb{N}^*}$ where $x_n := \left(n^3+n^2-n+1\right) / \left(n^2+2n+5\right)$ ✧

Exercise 2.20. Consider the sequence $(x_n)_{n\in\mathbb{N}^*}$ defined as

$$x_n := \sqrt{2+x_{n-1}}, \quad n\in\mathbb{N}^*, \qquad \text{where } x_0 := \sqrt{2}.$$

(a) Use mathematical induction to prove that $(x_n)_{n\in\mathbb{N}^*}$ is increasing.

(b) Prove that $(x_n)_{n\in\mathbb{N}^*}$ converges.

(c) Find $\lim_{n\to\infty} x_n$. ✧

Exercise 2.21. Show that every sequence in $\mathbb{R}$ has a monotonic subsequence[6].

Hint: Discuss the cases when the set $\{x_k \mid x_k \geq x_n \ \forall n > k\}$ is finite/infinite ✧

Exercise 2.22. Let $a > 0$ and $(x_n)_{n\in\mathbb{N}^*}$ defined by $x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ for $n\in\mathbb{N}^*$ and $x_1 := \frac{1}{2}(a+1)$.

(a) Show that $(x_n)_{n\in\mathbb{N}^*}$ is bounded from below.

(b) Show that $(x_n)_{n\in\mathbb{N}^*}$ is decreasing.

(c) Conclude that $\lim_{n\to\infty} x_n$ exists and find its value. ✧

Exercise 2.23. Let $\mathbb{R}$ be equipped with the standard metric. Prove or disprove:

(a) $(x_n)_{n\in\mathbb{N}^*}$ is Cauchy and $(y_n)_{n\in\mathbb{N}^*}$ is bounded. $\implies$ $(x_n y_n)_{n\in\mathbb{N}^*}$ is Cauchy.

(b) $(x_n)_{n\in\mathbb{N}^*}$ has a convergent subsequence. $\implies$ $(x_n)_{n\in\mathbb{N}^*}$ is bounded. ✧

Exercise 2.24. Let $a\in\mathbb{R}$ and $(x_n)_{n\in\mathbb{N}^*}$ be a convergent sequence in $\mathbb{R}$ such that $\lim_{n\to\infty} x_n > a$. Prove there exists $N\in\mathbb{N}^*$ such that $n > N$ implies $x_n > a$. ✧

Exercise 2.25. Let (M,d) be a metric space and $(x_n)_{n\in\mathbb{N}^*}$ a sequence in M. Show that:

(a) $(x_n)_{n\in\mathbb{N}^*}$ is Cauchy. $\implies$ $(x_n)_{n\in\mathbb{N}^*}$ is bounded.

(b) $(x_n)_{n\in\mathbb{N}^*}$ is Cauchy and has a convergent subsequence. $\implies$ $(x_n)_{n\in\mathbb{N}^*}$ is convergent. ✧

[6] that is, a subsequence that is either increasing or decreasing

Exercise 2.26. Let $a > 0$. Show that the sequence $(x_n)_{n\in\mathbb{N}^*}$ defined by $x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ for $n \in \mathbb{N}^*$ and $x_1 := b > 0$, converges to $\sqrt{a}$.
Hint: Note that this is basically Exercise 2.22. Hence, you have discuss the influence of the initial value x_1. ✧

Exercise 2.27. Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence such that $|x_{n+1} - x_n| \le 2^{-n}$ for all $n \in \mathbb{N}^*$. Prove that $(x_n)_{n\in\mathbb{N}^*}$ is a convergent sequence.

Does this also work if we assume instead $|x_{n+1} - x_n| \le \frac{1}{n}$ for all $n \in \mathbb{N}^*$? ✧

Exercise 2.28. Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathbb{R} \setminus \{0\}$ such that the limit $L := \lim_{n\to\infty}\left|\frac{x_{n+1}}{x_n}\right|$ exists.

(a) Show that if $L < 1$, then $\lim_{n\to\infty} x_n = 0$.
Hint: Select a so that $L < a < 1$ and obtain N so that $|x_{n+1}| < a|x_n|$ for $n \ge N$.

(b) Show that if $L > 1$, then $\lim_{n\to\infty}|x_n| = +\infty$.

(c) Use (a) and (b) to show that

$$\lim_{n\to\infty} a^n = \begin{cases} 0, & \text{if } |a| < 1, \\ 1, & \text{if } a = 1, \\ +\infty, & \text{if } a > 1, \\ \text{does not exist} & \text{if } a \le -1. \end{cases}$$ ✧

Exercise 2.29. Let $a, b, r \in \mathbb{R}$ and consider the sequence $(x_n)_{n\in\mathbb{N}^*}$ defined by

$$x_n := \sum_{k=0}^{n} r^k, \quad n \in \mathbb{N}^*.$$

(a) Show that

$$b^{n+1} - a^{n+1} = (b - a)\sum_{j=0}^{n} a^{n-j} b^j.$$

(b) Conclude that

$$x_n = \begin{cases} \dfrac{r^{n+1} - 1}{r - 1}, & \text{if } r \ne 1, \\ n + 1, & \text{if } r = 1, \end{cases} \quad \text{for all} \quad n \in \mathbb{N}^*.$$

(c) Prove that $(x_n)_{n\in\mathbb{N}^*}$ converges if and only if $|r| < 1$ and find the limit in this case. ✧

Exercise 2.30. Let $(x_n)_{n\in\mathbb{N}^*}$ be a real sequence and define

$$\sigma_n := \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \quad \text{for} \quad n \in \mathbb{N}^*.$$

(a) Show that

$$\liminf_{n\to\infty} x_n \le \liminf_{n\to\infty} \sigma_n \le \limsup_{n\to\infty} \sigma_n \le \limsup_{n\to\infty} x_n.$$

Hint: For the last inequality, show first that $M > N$ for $M, N \in \mathbb{N}^*$ implies

$$\sup\{\sigma_n \mid n > M\} \le \frac{1}{M}(x_1 + x_2 + \cdots + x_N) + \sup\{x_n \mid n > N\}$$

(b) Show that if $\lim_{n\to\infty} x_n$ exists, then $\lim_{n\to\infty} \sigma_n$ exists and $\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} x_n$.

(c) Prove that if $\lim_{n\to\infty} \sigma_n$ exists, $(x_n)_{n\in\mathbb{N}^*}$ is nondecreasing and $x_n \ge 0$ for all $n \in \mathbb{N}^*$, then $\lim_{n\to\infty} x_n$ exists and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} \sigma_n$.

(d) Give an example where $\lim_{n\to\infty} \sigma_n$ exists, but $\lim_{n\to\infty} x_n$ does not exist. ✧

Exercise 2.31. Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathbb{C}$ and let $(x_{k_1(n)})_{n\in\mathbb{N}^*}$, $(x_{k_2(n)})_{n\in\mathbb{N}^*}, \dots, (x_{k_m(n)})_{n\in\mathbb{N}^*}$ be subsequences that partition $\mathbb{N}^*$, i.e.,

- $\displaystyle\bigcup_{i=1}^{m} k_i(\mathbb{N}^*) = \mathbb{N}^*$ and
- $k_i(\mathbb{N}^*) \cap k_j(\mathbb{N}^*) = \emptyset$, if $i \ne j$.

Show: If all those subsequences are convergent with limits $y_1, y_2, \dots, y_m$, respectively, then there exists no other subsequence with a limit different from $y_1, y_2, \dots y_m$. ✧

Exercise 2.32. Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathbb{C}$. Show

$(x_n)_{n\in\mathbb{N}^*}$ converges
$\iff$ all subsequences of $(x_n)_{n\in\mathbb{N}^*}$ converge
$\iff$ all subsequences of $(x_n)_{n\in\mathbb{N}^*}$ converge to the same limit. ✧

Exercise 2.33. Prove that each of the following series converges and find the limit.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{\mathrm{e}^{k-1}}$ (b) $\displaystyle\sum_{k=0}^{\infty} \frac{5^{k+1} + (-3)^k}{7^{k+2}}$ ✧

Exercise 2.34. Prove that each of the following series diverges (that is, "does not converge").

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

Exercise 2.35. Represent each of the following series $\sum_{k=1}^{\infty} a_k$ as a *telescopic series*, that is write $a_k = b_k - b_{k+1}$ for all $k \in \mathbb{N}^*$ and some sequence $(b_k)_{k\in\mathbb{N}^*}$, and find its limit.

(a) $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k(k+1)}$ (b) $\displaystyle\sum_{k=1}^{\infty}\frac{12}{(k+2)(k+3)}$ ✧

Exercise 2.36. Let $(a_k)_{k\in\mathbb{N}^*}$ be a complex sequence. Prove that if $\sum_{k=1}^{\infty}|a_k|$ converges and $(b_k)_{k\in\mathbb{N}^*}$ is a bounded complex sequence, then $\sum_{k=1}^{\infty} a_k b_k$ converges. ✧

Exercise 2.37. Prove that if $\sum_{k=1}^{\infty} a_k$ is a convergent series of nonnegative numbers and $p > 1$, then $\sum_{k=1}^{\infty} a_k^p$ converges. ✧

Exercise 2.38. Show that if $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ are convergent series of nonnegative numbers, then $\sum_{k=1}^{\infty} \sqrt{a_k b_k}$ converges.
Hint: Show $\sqrt{a_k b_k} \le a_k + b_k$ for all $k \in \mathbb{N}^*$. ✧

Exercise 2.39. Let $(a_n)_{n\in\mathbb{N}^*}$ be a complex sequence.

(a) Show that $\sum_{n=1}^{\infty}|a_n|$ converges implies that $\sum_{n=1}^{\infty}|a_n|^2$ also converges.

(b) Give an example of a divergent series $\sum_{n=1}^{\infty} a_n$ for which $\sum_{n=1}^{\infty} a_n^2$ converges.

(c) Give an example of a convergent series $\sum_{n=1}^{\infty} a_n$ for which $\sum_{n=1}^{\infty} a_n^2$ diverges. ✧

Exercise 2.40. (a) Prove if (a_n) is a decreasing sequence of real numbers and if $\sum a_n$ converges, then $\lim_{n\to\infty} n a_n = 0$.

Hint: Consider $|a_{N+1} + a_{N+2} + \cdots + a_n|$ for suitable N.

(b) Use (a) to give a proof that $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges different from the one discussed in class. ✧

Exercise 2.41. Determine whether the following series converge

(a) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{n!}{n^n}$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{(n!)^2}{n^n}$ ✧

Exercise 2.42. Let $(a_n) \in \mathbb{R}^{\mathbb{N}^*}$ be decreasing and assume that $\sum_{n=1}^{\infty} a_n$ converges. Show that

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

Hint: Consider $|a_{N+1} + a_{N+2} + \cdots + a_n|$ *for suitable* N. ✧

Exercise 2.43. Define the function $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) := \sum_{k=0}^{\infty} \frac{x^k}{k!}$, $x \in \mathbb{R}$.[7]

(a) Show that the series defining $f(x)$ converges absolutely for every $x \in \mathbb{R}$.

(b) Show that $f(1) = \mathrm{e}$, where e refers to the Euler number, i.e., $\mathrm{e} := \lim_{n\to\infty}(1 + \frac{1}{n})^n$.

Hint: Show that the partial sums $s_n(1)$ of $f(1)$ satisfy $(1 + \frac{1}{n})^n \leq s_n(1) \leq \mathrm{e}$ for all $n \in \mathbb{N}^*$.

(c) Show (more generally) that for any $x > 0$, $f(x) = \lim_{n\to\infty}(1 + \frac{x}{n})^n$. *Show that the partial sums $s_n(x)$ of $f(x)$ satisfy $(1 + \frac{x}{n})^n \leq s_n \leq \lim_{n\to\infty}(1 + \frac{x}{n})^n$ for all $n \in \mathbb{N}^*$.* ✧

Exercise 2.44. Determine which of the following series $\sum_{k=1}^{\infty} a_k$ converge or converge absolutely.

(a) $a_k := \dfrac{(-2)^k}{k!}$

(b) $a_k := \begin{cases} \frac{2^{-k}}{k!}, & k \text{ is even,} \\ \frac{2^{-(k+1)}}{(k+1)!}, & k \text{ is odd.} \end{cases}$

(c) $a_k := \dfrac{(-1)^k}{\sqrt{k}}$ ✧

Exercise 2.45. Consider the series $\sum_{k=1}^{\infty} \frac{k}{2^k}$.

(a) Show that the series converges absolutely, and

(b) find the limit.

Hint: Let $s := \sum_{k=1}^{\infty} \frac{k}{2^k}$ and find an equation of the form $s = 2s + B$ for some $B \in \mathbb{R}$. ✧

Exercise 2.46. As an alternative to Exercise 2.45 you can find the limit of $\sum_{n=1}^{\infty} \frac{n}{2^n}$ also by the following strategy. Prove $\sum_{n=1}^{\infty} \frac{n-1}{2^{n+1}} = \frac{1}{2}$. Use this result to calculate $\sum_{n=1}^{\infty} \frac{n}{2^n}$.

Hint: Note $\frac{k-1}{2^{k+1}} = \frac{k}{2^k} - \frac{k+1}{2^{k+1}}$ for all $k \in \mathbb{N}^*$. ✧

Exercise 2.47. Determine for which values of $x \in \mathbb{R}$ the following series converge or converge absolutely.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{x^{2k}}{k+1}$ (b) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k x^k}{\sqrt{k^2+1}}$ ✧

[7] Here, we follow the convention that $0^0 := 1$.

Exercise 2.48. For $(a_j)_{j\in\mathbb{N}^*}$ with $a_j \neq 0$ for all $j \in \mathbb{N}^*$, show that

$$\limsup_{j\to\infty} \sqrt[j]{|a_j|} \leq \limsup_{j\to\infty} \left|\frac{a_{j+1}}{a_j}\right|.$$

Conclude that $\frac{1}{\sqrt[n]{n!}} \to 0$ as $n \to \infty$. ✧

Exercise 2.49. Let $(X, \|\cdot\|)$ be a normed vector space (see Exercise 2.1). Assume that every Cauchy sequence in $(X, \|\cdot\|)$ converges, i.e., X is complete. Then for $(x_n)_{n\in\mathbb{N}^*} \in X^{\mathbb{N}^*}$,

$$\sum_{n=1}^{\infty} \|x_n\| \text{ converges in } (\mathbb{R}, d_2) \implies \sum_{n=1}^{\infty} x_n \text{ converges in } (X, d).$$ ✧

Exercise 2.50. Calculate $\sum_{n=2}^{\infty} \sum_{m=2}^{\infty} \frac{1}{m^n}$. ✧

Chapter 3

Continuous functions

3.1 Limits of Functions

Definition 3.1 (limit point)**.** Let (X, d) be a metric space and $D \subseteq X$. An element $x \in X$ is called *limit point of D* (or *accumulation point*) if there exists a sequence in $D \setminus \{x\}$ converging to x. If $x \in D$, but x is not a limit point, we say that x is an *isolated point of D*. ✧

Fact 3.2. *Let (X, d) be a metric space, $D \subseteq X$ and $x \in X$.*

(i) *Then x is a limit point if and only if for every $\varepsilon > 0$ the set $\{y \in D : d(y, x) < \varepsilon\} \setminus \{x\}$ is non-empty.*

(ii) *A limit point of D does not need to be contained in the set D, and, conversely, an element of D is not automatically a limit point, e.g., an isolated point is not a limit point.*

Example 3.3. Let us consider a few illustrative examples:

(a) $X = \mathbb{R}$, $d = d_2$, $D = (a, b)$. Then every element in $[a, b]$ is a limit point of D.

(b) Let X be non-empty and d be some metric on X. Any finite subset D of X has no limit point and only consists of isolated points.

(c) X be non-empty and $d = d_0$ be the discrete metric. Then no subset $D \subseteq X$ has limit points. ✧

Definition 3.4. Let (X, d_X) and (Y, d_Y) be metric spaces, $D \subseteq X$ and a function $f\colon D \to Y$. For a limit point $x \in X$ of D, an element $L \in Y$ is called L *the limit of $f(t)$ as t goes to x (in D)*, denoted by

$$\lim_{t \to x} f(t) = \lim_{t \to x, t \in D} f(t) = L \quad \text{or} \quad f(t) \to L \ (t \to x),$$

if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall y \in D \ : \ 0 < d_X(x, y) < \delta \implies d_Y(L, f(y)) < \varepsilon.$$

For $Y = \mathbb{R}$ and a limit point of D, we say that $f(t)$ *goes to* $+\infty$ *($-\infty$)*, write $\lim_{t\to x} f(t) = \pm\infty$, if

$$\forall C > 0 \ \exists \delta > 0 \ \forall y \in D \ : \ d_X(x,y) < \delta \implies f(y) > C$$
$$(\forall C < 0 \ \exists \delta > 0 \ \forall y \in D \ : \ d_X(x,y) < \delta \implies f(y) < C).$$

If $X = \mathbb{R}$, we define the following variations of the limit:

- If x is a limit point of $D \cap (x,\infty)$, we call $\lim_{t\to x, t\in D\cap(x,\infty)} f(t)$, *the right-sided limit* of $f(t)$ as t goes to x, denoted by $\lim_{t\to x^+} f(t)$.
- If x is a limit point of $D \cap (-\infty, x)$, the limit $\lim_{t\to x, t\in D\cap(-\infty,x)} f(t)$ is called the *left-sided limit of $f(t)$ as t goes to x*, denoted by $\lim_{t\to x^-} f(t)$.
- If D is not bounded above (below), we call $L \in Y$ the limit of $f(t)$ as $t \to \infty$ ($t \to -\infty$), if $\exists K > 0$ such that $(K,\infty) \subseteq D$ ($(-\infty, K) \subseteq D$) and

$$\forall \varepsilon > 0 \ \exists C \in \mathbb{R} \ \forall y > C \ : \ d_Y(L, f(y)) < \varepsilon$$
$$(\forall \varepsilon > 0 \ \exists C \in \mathbb{R} \ \forall y < C \ : \ d_Y(L, f(y)) < \varepsilon). \qquad \diamond$$

Fact 3.5. *Let (X, d_X) and (Y, d_Y) be metric spaces, $D \subseteq X$, $x \in X$, $L \in Y$ and $f \colon D \to Y$.*

(i) *Then $\lim_{t\to x} f(t) = L$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}^*}$ in $D\setminus\{x\}$ converging to x, we have that $(f(x_n))_{n\in\mathbb{N}^*}$ converges to L. This corresponding statements hold for one-sided limits (left and right-sided limits) and limits to $\pm\infty$.*

Pf: We show the two direction separately.

$\boxed{\Rightarrow}$: Suppose that $\lim_{t\to x} f(t) = L$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y(L, f(y)) < \varepsilon$ for all $y \in D$ with $d_X(x,y) < \delta$. On the other hand, for every sequence $(x_n)_n$ in $D \setminus \{x\}$ converging to x, there exists $N \in \mathbb{N}^*$ such that $0 < d_X(x, x_n) < \delta$ for all $n \geq N$. Therefore, $d_Y(L, f(x_n)) < \varepsilon$, which shows that $f(x_n) \to L$ as $n \to \infty$.

$\boxed{\Leftarrow}$: To see the converse, we use contra position. Suppose that $\lim_{t\to x} f(t) \neq L$, i.e.,

$$\exists \varepsilon > 0 \ \forall \delta > 0 \ \exists y \in D \ : \ 0 < d_X(x,y) < \delta \wedge d_Y(L, f(y)) \geq \varepsilon.$$

Now choose $\delta = \frac{1}{n}$ and $x_n = y$. Then obviously $x_n \to x$ and $f(x_n) \not\to L$.

(ii) *If $Y = \mathbb{R}$, $d_Y = d_2$, $g \colon D \to Y$ and $\lim_{t\to x} f(t)$ as well as $\lim_{t\to x} g(t)$ exists, then*

$$\lim_{t\to a}(f+g)(t), \quad \lim_{t\to x}(f\cdot g)(t)$$

exists and equals $\lim_{t\to a} f(t) + \lim_{t\to a} g(t)$ and $\lim_{t\to a} f(t) \cdot \lim_{t\to a} g(t)$, respectively. If also $g(t) \neq 0$ for all $t \in D$ and $\lim_{t\to x} g(t) \neq 0$, then $\lim_{t\to x} \frac{f(t)}{g(t)}$ exists and equals $\frac{\lim_{t\to a} f(t)}{\lim_{t\to a} g(t)}$.

This follows directly from the first fact and the corresponding result on limits of sequences of real numbers.

(iii) *Let* $D \subseteq \mathbb{R}$ *and* $f\colon D \to \mathbb{R}$ *considered with the standard metric. Let* $x \in \mathbb{R}$ *be a limit point of the sets* D, $(x,\infty)\cap D$ *and* $(-\infty,x)\cap D$. *Then* $\lim_{t\to x} f(t)$ *exists if and only if both one-sided limits* $\lim_{t\to x^+} f(t)$ *and* $\lim_{t\to x^-} f(t)$ *exist and are equal, see (Problem Set 6, Practice Problem 5).*

3.2 Continuous functions

Definition 3.6. Let (X, d_X) and (Y, d_Y) be metric spaces and $f\colon X \to Y$ be a function and $x \in X$.

- We call f *continuous at* x if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in X\ :\ d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon. \tag{3.1}$$

- We say that f *is continuous*, if f is continuous at every point $x \in X$, that is

$$\forall x \in X\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in X\ :\ d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon. \tag{3.2}$$

- We call f *uniformly continuous* (on X) if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in X\ :\ d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon. \tag{3.3}$$

✧

Fact 3.7. *Let* (X, d_X) *and* (Y, d_Y) *be metric spaces.*

(i) *If* $f\colon X \to Y$ *and* x *is an isolated point of* X, *then* f *is continuous at* x *because there exists* $\delta > 0$ *such that* $\{y \in X \mid d_X(x,y) < \delta\} = \{x\}$.

(ii) *The following assertions are equivalent.*

(a) f *is continuous at* x.

(b) $\forall (x_n)_{n\in\mathbb{N}^*} \in X^{\mathbb{N}^*}\ :\ x_n \xrightarrow[n\to\infty]{d_X} x \implies f(x_n) \xrightarrow[n\to\infty]{d_Y} f(x)$.

(c) x *is an isolated point or* $\lim_{t\to x} f(t) = f(x)$.

(iii) *The* identity mapping $I\colon X \to X, x \mapsto x$ *is continuous (more precisely, one may choose* $\delta = \varepsilon$ *in the above definition).*

(iv) *Continuity is a local property: That is, for functions* $g, f\colon X \to Y$, $x \in X$ *and* $\varepsilon > 0$ *such that* $g(z) = f(z)$ *for all* $z \in X$ *with* $d(x,z) < \varepsilon$, *it holds that* f *is continuous at* x *if and only if* g *is continuous at* x.

(v) *Note that in our definition of continuity, the set* X *is the set on which the function is defined (in contrast to the definition of limits, where* f *was only defined on a subset of* X*).*

(vi) *Let* $Y = \mathbb{R}$ *and* $d = d_2$. *If* $f\colon X \to \mathbb{R}$ *is continuous at* $x \in X$ *and* $f(x) \neq 0$, *then there exists* $\delta > 0$ *such that* $f(t) \neq 0$ *for all* $t \in X$ *with* $d(x,t) < \delta$.

(vii) *Let* $Y = \mathbb{R}$, $d_Y = d_2$ *and* $f, g\colon X \to Y$ *continuous at* $x \in X$. *Then* $f+g$ *and* fg *are continuous at* x. *If* $g(x) \neq 0$, *then* $\frac{f}{g}$ *is also continuous.*

(viii) *Let* $X = \mathbb{R} = Y$, $d_X = d_Y = d_2$. *The constant function* $f\colon \mathbb{R} \to \mathbb{R}, f(x) = k$, *where* $\mathbb{R}$ *is continuous as we can directly see by definition (for any* ε, *we can choose any* δ*). Therefore, using the previous point of this fact list, we derive that every polynomial* $p\colon \mathbb{R} \to \mathbb{R}$ *with real coefficients,* $p(x) = \sum_{k=0}^{n} a_k x^k$, $a_k \in \mathbb{R}$, $n \in \mathbb{N}$, *is continuous.*

(ix) *For the discrete metric space* (M, d_0), *with* M *nonempty and some arbitrary metric space* (N, d_N) *we have that every function* $f\colon M \to N$ *is continuous. See Exercise 10.*

(x) ***Note the difference*** *between* $f\colon X \to Y$ continuous *and* uniformly continuous*: The difference in* (3.2) *and* (3.3) *is the position of the the quantifier "*$\forall x$*". Therefore, if* f *is uniformly continuous,* f *is also continuous. The converse is wrong in general, as the example* $f\colon (0,\infty) \to \mathbb{R}$, $f(x) = \frac{1}{x}$, $x \in \mathbb{R}$ *shows.*

Proposition 3.8 (Composition of continuous functions)**.** *Let* (X, d_X), (Y, d_Y), (Z, d_Z) *be metric spaces and* $f\colon X \to Y$ *be continuous at* $x \in X$ *and* $g\colon Y \to Z$ *be continuous at* $f(x) \in Y$. *Then* $g \circ f\colon X \to Z$ *is continuous at* x.

Proof. (In class we showed the proof by using the characterisation of continuity via sequences, Fact 3.7 (ii)). Here we show the claim by directly referring to the definition. Let $\varepsilon > 0$. Since g is continuous at $f(x)$, there exists $\delta > 0$ such that $d_Z(g(f(x)), g(z)) < \varepsilon$ for all $z \in Y$ with $d_Y(f(x), z) < \delta$. Since f is continuous at x, there exists $\delta' > 0$ such that $d_Y(f(x), f(y)) < \delta$ for all $y \in X$ with $d_X(x, y) < \delta'$. Therefore, $d_Z(g(f(x)), g(f(y)) < \varepsilon$ for all $y \in X$ with $d_X(x, y) < \delta'$. ❑

Definition 3.9 (bounded, closed and compact sets)**.** Let (X, d) be a metric space and $M \subseteq X$. We say that M is

- *bounded* (in (X, d)), if $\exists c \in X\, \exists R > 0 : M \subseteq \{z \in X \mid d(c, z) < R\}$;
- *closed* (in (X, d)), if M contains all limit points of M, i.e., $M = M \cup \{x \in X \mid x \text{ is limit point of } M\}$;
- *compact* (in (X, d)), if every sequence in M has a convergent subsequence with limit in M. ✧

Example 3.10. Every closed interval (according to Definition 1.30) in $\mathbb{R}$ is indeed a closed set in $(\mathbb{R}, d_2)$, because limits respect non-strict inequalities, see Fact 2.8 (ii) and choosing one sequence as one of the boundary points of the closed interval. ✧

Fact 3.11. *Let (X, d) be a metric space and $M \subseteq X$.*

(i) *M is bounded if and only if every sequence in M is bounded, see Definition 2.4.*

(ii) *M is closed if and only every convergent sequence in M has its limit in M.*

(iii) *If M is compact, then M is bounded and closed.*

(iv) *A compact set subset of $\mathbb{R}$ has a minimum and a maximum.*

Theorem 3.12 (Heine–Borel)**.** *Let $p \in \mathbb{N}^*$. A set $M \subseteq \mathbb{R}^p$ is compact in $(\mathbb{R}^p, d_2)$ if and only if M is bounded and closed (in $(\mathbb{R}^p, d_2)$).*

Proof. We prove this result first for $p = 1$ and then for an arbitrary $p \in \mathbb{N}^*$.
Step 1. Let $p = 1$. By Fact 3.11, we only have to show that a closed and bounded set $M \subseteq \mathbb{R}^p$ is compact in $(\mathbb{R}^p, d_2)$. Let $p = 1$ and consider a sequence in M. Since it is bounded, we know by the Bolzano–Weierstraß theorem, Theorem 2.21, that the sequence has a convergent subsequence. The limit lies in M because M is closed.

Step 2. For $p > 1$, we first consider the sequence of the first components of the sequence. Since this a sequence of real numbers, we now as in the case $p = 1$, that there exists a converging subsequence. Consider the indices of this subsequence and consider the corresponding sequence of the second components of the original sequence. Again, there exists a converging subsequence. Repeating this procedure for all p components of our sequence, yields a subsequence of the original sequence whose component sequences all converge. By the inequality $d_2(\mathbf{x}, \mathbf{y}) \leq \max_{i=1,\ldots,p}|x_i - y_i|$, this implies that this subsequence converges in $(\mathbb{R}^p, d_2)$. Because M is closed, the limit lies in M. Therefore, M is compact. ❑

Theorem 3.13. *Let (X, d_X), (Y, d_Y) be metric spaces and $f : X \to Y$ be continuous. If $M \subseteq X$ is compact, then $f(M)$ is compact.*

Proof. See Exercise 3.12. ❑

Corollary 3.14. *Let (X, d_X) be a metric space and $f : X \to \mathbb{R}$ be continuous. If $K \subseteq X$ is compact, then f attains its supremum and infimum on K, that is,*

$$\exists x_{\min}, x_{\max} \in K \ : \ f(x_{\min}) = \inf_{x \in K} f(x) \ \wedge \ f(x_{\max}) = \sup_{x \in K} f(x).$$

In particular, the infimum and supremum are finite and equal the minimum and maximum, respectively.

Theorem 3.15 (Intermediate value theorem I)**.** *Let $a < b$ be real numbers and $f : [a, b] \to \mathbb{R}$ be continuous, where both $[a, b]$ and $\mathbb{R}$ are equipped with the standard metric. If $f(a) \cdot f(b) < 0$, then there exists $c \in [a, b]$ such that $f(c) = 0$.*

Note that the condition $f(a) \cdot f(b) < 0$ basically says that $f(a)$ and $f(b)$ have different signs, i.e., either

$$f(a) < 0 \ \wedge \ f(b) > 0 \quad \text{or} \quad f(a) > 0 \ \wedge \ f(b) < 0.$$

Proof. W.l.o.g. we assume $f(a) < f(b)$. Otherwise we regard $-f$. Let $y \coloneqq \inf\{x \in [a,b] : f(x) \geq 0\}$. By continuity we have $f(y) \geq 0$. Moreover, if $f(y) > 0$, then by continuity there exists a $\delta > 0$ for $\varepsilon \coloneqq \frac{f(y)}{2}$ such that

$$f\big((y-\delta, y+\delta)\big) \subseteq \big(f(y)-\varepsilon, f(y)+\varepsilon\big) \subseteq (0,+\infty).$$

Hence, $f(y - \frac{\delta}{2}) > 0$, which contradicts y being the infimum of $\{x \in [a,b] : f(x) \geq 0\}$ (because $y - \delta < y$). Therefore, $f(y) = 0$. ❑

Corollary 3.16. *Let $g\colon [a,b] \to \mathbb{R}$ be a continuous function (with respect to the standard metric d_2). Then for every real number d in the interval $\big[\min\{g(a), g(b)\}, \max\{g(a), g(b)\}\big]$, there exists $c \in [a,b]$ such that $g(c) = d$.*

Proof. For $d = g(a)$ or $d = g(b)$ the claim is trivial.

Hence, we regard $d \in (\min\{g(a), g(b)\}, \max\{g(a), g(b)\})$ and define $f\colon [a,b] \to \mathbb{R}$ by $f(x) = g(x) - d$, for $x \in [a,b]$ and note that $f(a) \cdot f(b) < 0$ since $\min\{g(a), g(b)\} < \max\{g(a), g(b)\}$. Thus, by Theorem 3.15, we have that there exists $c \in [a,b]$ such that $f(c) = 0$. Thus $g(c) = d$. ❑

Example 3.17. Let $n \in \mathbb{N}^*$. The function $f\colon [0,\infty) \to [0,\infty), x \mapsto x^n$ maps $[0,\infty)$ onto $[0,\infty)$. ✧

Theorem 3.18. *Let $f\colon X \to Y$ be continuous and X be compact. Then f is uniformly continuous.*

Proof. Suppose that f is not uniformly continuous. This means that

$$\exists \varepsilon > 0 \ \forall \delta > 0 \ \exists x, y \in X \ : \ d_X(x,y) < \delta \wedge d_Y(f(x), f(y)) \geq \varepsilon.$$

This implies (set $\delta = \frac{1}{n}$, $x = x_n$ and $y = y_n$) that for this δ, there exist two sequences $(x_n)_n, (y_n)_n$ in X such that $d_X(x_n, y_n) < \frac{1}{n}$ and $d_Y(f(x_n), f(y_n)) \geq \varepsilon$ for all $n \in \mathbb{N}^*$. Since X is compact, there exists a convergent subsequence $(x_{k_n})_{n\in\mathbb{N}^*}$ (recall, that means that $k\colon \mathbb{N}^* \to \mathbb{N}^*$ is strictly increasing and we can also write $x \circ k$ for the subsequence) with limit $x \in X$. Again, by the compactness of X, the (sub)sequence $(y_{k_n})_{n\in\mathbb{N}^*} = y \circ k$ has a convergent subsequence $(y_{k'_{k_n}})_{n\in\mathbb{N}^*} = y \circ k \circ k'$ with limit $y \in X$. Also, $(x_{k'_{k_n}})_{n\in\mathbb{N}^*}$ converges to x (as a subsequence of the convergent sequence $(x_{k_n})_{n\in\mathbb{N}^*}$), and therefore

$$0 \leq d_X(x,y) \leq d_X(x, x_{k'_{k_n}}) + d_X(x_{k'_{k_n}}, y_{k'_{k_n}}) + d_X(y_{k'_{k_n}}, y) \to 0 \text{ as } n \to \infty.$$

By the squeeze theorem and the properties of metrics, we conclude that $x = y$, which further implies that $d_Y(f(x), f(y)) = 0$. This is a contradiction to $d_Y(f(x), f(y)) \geq \varepsilon$. ❑

3.3 Sequences and Series of functions

In Chapter 2 we have already defined sequences in general metric spaces. Now we consider a sequences $(f_n)_{n\in\mathbb{N}^*}$ of functions $f_n\colon X \to Y$, where, in general X and Y refer to metric spaces. In particular, the cases $X = I$ and $Y = \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, and with both sets considered with the standard metric d_2 will be most important and should serve as guiding example.

Definition 3.19. Let (X, d_X) and (Y, d_Y) be metric spaces. We define the following sets:

$$\begin{aligned} \mathrm{C}(X;Y) &:= \{f\colon X \to Y \mid f \text{ is continuous}\},\\ \mathrm{C_b}(X;Y) &:= \{f\colon X \to Y \mid f \text{ is continuous and bounded}\}. \end{aligned}$$

If $Y = \mathbb{R}$ and $d_Y = d_2$, we also write $\mathrm{C}(X) = \mathrm{C}(X;\mathbb{R})$ and $\mathrm{C_b}(X) = \mathrm{C_b}(X;\mathbb{R})$. ✧

The following proposition was discussed in the lecture for the special case $X = I \subseteq \mathbb{R}$, $d_X = d_2 = d_Y$ and $Y = \mathbb{R}$. The proof of the general case is completely analogous.

Proposition 3.20. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then the function*

$$d_\infty\colon \mathrm{C_b}(X;Y) \times \mathrm{C_b}(X;Y) \to [0,\infty), \quad (f,g) \mapsto \sup_{x\in X} d_Y(f(x), g(x))$$

defines a metric on $\mathrm{C_b}(X;Y)$.

Proof. We show the three defining properties of a metric. Since $d_\infty(f,g) = 0$ if and only if $d_Y(f(x), g(x)) = 0$ for all $x \in X$ iff and only if $f(x) = g(x)$ for all $x \in X$, the reflexivity is clear. The symmetry, $d_\infty(f,g) = d_\infty(g,f)$ is clear from the symmetry of d_Y. The triangle inequality follows from the properties of suprema of sums, and the triangle inequality for d_Y: For $f, g, h \in \mathrm{C_b}(X,Y)$,

$$\begin{aligned} d_\infty(f,h) = \sup_{x\in X} d_Y(f(x), h(x)) &\le \sup_{x\in X} \left(d_Y(f(x), g(x)) + d_Y(f(x), g(x))\right)\\ &\le \sup_{x\in X} d_Y(f(x), g(x)) + \sup_{x\in X} d_Y(g(x), h(x))\\ &= d_\infty(f,g) + d_\infty(g,h). \end{aligned}$$ ❑

Fact 3.21. *Let (X, d_X) and (Y, d_Y) be metric spaces.*

(i) *If $X \subseteq \mathbb{R}$ compact and $Y = \mathbb{R}$ equipped with the standard metric d_2, then $\mathrm{C}(X) = \mathrm{C_b}(X)$ and $d_\infty(f,g) = \max_{x\in X}|f(x) - g(x)|$.*

(ii) *Let X be compact, then $\mathrm{C}(X;Y) = \mathrm{C_b}(X;Y)$ and $d_\infty(f,g) = \max_{x\in X} d_Y(f(x), g(x))$.*

This follows from the fact that a metric d_Y is continuous from $Y \times Y$ to $[0, \infty)$ if $Y \times Y$ is equipped with the metric $d_{Y\times Y}(\mathbf{y}, \mathbf{z}) = d_Y(y_1, z_1) + d_Y(y_2, z_2)$ for $\mathbf{y}, \mathbf{z} \in Y \times Y$.

Definition 3.22 (Uniform and pointwise convergence)**.** Let (X, d_X) and (Y, d_Y) be metric spaces and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of functions $f_n \colon X \to Y$, $n \in \mathbb{N}^*$. Let $f \colon X \to Y$ be a function. We say that $(f_n)_{n\in\mathbb{N}^*}$

- *converges to f uniformly (on X)*, if

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}^*\ \forall n \geq N\ \forall x \in X\ :\ d_Y(f_n(x), f(x)) < \varepsilon.$$

- *converges pointwise to f (on X)*, if

$$\forall x \in X\ \underbrace{\forall \varepsilon > 0\ \exists N \in \mathbb{N}^*\ \forall n \geq N\ :\ d_Y(f_n(x), f(x)) < \varepsilon.}_{f_n(x) \to f(x)} \qquad \diamond$$

Fact 3.23. *Suppose we have a setting as in Definition 3.22.*

(i) *By definition of convergence it follows that "uniform convergence to f" implies "pointwise convergence to f" (the order of the quantifier $\forall x \in X$ matters!). The converse is wrong in general (as we shall see in the following examples).*

(ii) *Let $f_n \in \mathrm{C_b}(X; Y)$ for all $n \in \mathbb{N}$ and $f \in \mathrm{C_b}(X; Y)$. Then*

$$\begin{aligned} &(f_n)_{n\in\mathbb{N}^*} \text{ converges to } f \\ &\qquad\iff \quad \forall \varepsilon > 0\ \exists N \in \mathbb{N}^*\ \forall n \geq N\ :\ \sup_{x\in X} d_Y(f_n(x), f(x)) < \varepsilon \\ &\qquad\iff \quad d_\infty(f_n, f) \to 0\ (n \to \infty). \end{aligned}$$

Thus, uniform convergence means that $(f_n)_{n\in\mathbb{N}^*}$ converges in the metric space $(\mathrm{C_b}(X; Y), d_\infty)$.

(iii) *In fact, in the above point, continuity wasn't used anywhere. The corresponding statement applies to the metric space of bounded functions from X to Y with the metric d_∞.*

(iv) *In the following (also Analysis 2), we will deal with questions of the kind whether properties of the functions f_n, $n \in \mathbb{N}^*$ are carried over to the uniform limit f. For instance, we will shortly answer the question whether the uniform limit of continuous functions is continuous.*

Theorem 3.24 (Uniform convergence preserves continuity)**.** *Let X, Y be metric spaces and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of continuous functions from X to Y. If $(f_n)_n$ converges uniformly to $f \colon X \to Y$, then f is continuous.*

Proof. Let $x \in X$. We will show that f is continuous at x. Let $\varepsilon > 0$. By uniform convergence there exists $N \in \mathbb{N}^*$ such that $d_Y(f_N(y), f(y)) < \varepsilon/3$ for all $y \in X$. Since f_N is continuous at x, there exists $\delta > 0$ such that

$$\forall y \in X\ :\ d_X(x, y) < \delta \implies d_Y(f_N(x), f_N(y)) < \varepsilon/3.$$

Altogether we have for all $y \in X$ with $d_X(x,y) < \delta$ that

$$\begin{aligned} d_Y(f(x), f(y)) &\leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(y)) + d_Y(f_N(y), f(y)) \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

This means that f is continuous at x. ❑

> The above type of proof is sometimes referred to as "3ε proof".

Example 3.25. Let $q > 0$ and $X = [0, q]$, $Y = \mathbb{R}$, both with the standard metric d_2. Consider the sequence $(f_n)_{n\in\mathbb{N}^*}$ defined by $f_n(x) = x^n$, $n \in \mathbb{N}^*$, $x \in [0, q]$. By elementary properties on sequences of real numbers, we know that

$$f_n(x) \to \begin{cases} 0, & \text{if } x \in [0,1), \\ 1, & \text{if } x = 1, \\ +\infty, & \text{if } x > 1, \end{cases}$$

as $n \to \infty$. Therefore, $(f_n)_{n\in\mathbb{N}^*}$ converges pointwise on $[0, 1]$. Yet, since

$$d_\infty(f_n, f) = \sup_{t\in[0,1]} |f_n(t) - f(t)| = \sup_{t\in[0,1)} |t^n - 0| = 1,$$

for every $n \in \mathbb{N}^*$, $d_\infty(f_n, f) \not\to 1$ as $n \to \infty$. Therefore, by the above fact, $(f_n)_{n\in\mathbb{N}^*}$ does not converge uniformly on $[0, 1]$ (alternatively, this could be argued by Theorem 3.24 as f is not continuous on $[0, 1]$).

However, since for $q < 1$,

$$\sup_{t\in[0,q]} |f_n(t) - f(t)| = \sup_{t\in[0,q]} |t^n| = q^n \to 0 \ (n \to \infty),$$

we conclude that $(f_n)_{n\in\mathbb{N}^*}$ converges uniformly to the zero function on every interval $[0, q]$ with $q < 1$. ✧

Theorem 3.26. *Let $I \subseteq \mathbb{R}$ be a compact interval and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of continuous functions $f_n \colon I \to \mathbb{R}$, $n \in \mathbb{N}^*$. Then*

$$(f_n)_{n\in\mathbb{N}^*} \text{ converges uniformly} \iff (f_n)_{n\in\mathbb{N}^*} \text{ is a Cauchy sequence in } (\mathrm{C}(I), d_\infty).$$

Proof. $\boxed{\Rightarrow}$: By Fact 3.23 uniform convergence of $(f_n)_{n\in\mathbb{N}^*}$ is equivalent to $(f_n)_{n\in\mathbb{N}^*}$ converging in $(\mathrm{C}(I), d_\infty)$. Every convergent sequence is a Cauchy sequence.

$\boxed{\Leftarrow}$: Suppose that $(f_n)_{n\in\mathbb{N}^*}$ is a Cauchy sequence in $(\mathrm{C}(I), d_\infty)$. In order to show that the sequence converges, we first find a candidate for its limit. Let $x \in I$. Since for any $n, m \in \mathbb{N}^*$,

$$|f_n(x) - f_m(x)| \leq d_\infty(f_n, f_m),$$

we have that $(f_n(x))_{n\in\mathbb{N}^*}$ is a Cauchy sequence in $(\mathbb{R}, d_2)$ and thus converges by Theorem 2.9 to a limit which we denote by $f(x)$. It immediately follows that f is the pointwise limit of $(f_n)_{n\in\mathbb{N}^*}$ and it remains to show that $f_n \to f$ uniformly as $n \to \infty$. By the Cauchy sequence property and the definition of d_∞ we have: for every $\varepsilon > 0$ there exists $N \in \mathbb{N}^*$ independent of $x \in I$ such that for all $n, m \geq N$

$$|f_n(x) - f_m(x)| < \varepsilon.$$

We get, by letting $m \to \infty$ on each sides of the inequality, that

$$|f_n(x) - f(x)| \leq \varepsilon,$$

which readily implies that $f_n \to f$ uniformly as $n \to \infty$. ❑

Remark 3.27. The above theorem can immediately be generalized to sequences of functions in $\mathrm{C_b}(X) = \mathrm{C_b}(X;\mathbb{R})$ with metric spaces X. ✧

Recall from the tutorials Exercise 2.1 the following special examples of metric spaces: Let X be a vector space (over $\mathbb{R}$) and $\|\cdot\|\colon X \to [0,\infty)$ with the following properties for $f, g \in X, c \in \mathbb{R}$:

(i) $\|f\| = 0 \iff f = \mathbf{0}$,

(ii) $\|cf\| = |c|\|f\|$,

(iii) $\|f+g\| \leq \|f\| + \|g\|$.

Then $d(x,y) = \|x-y\|$, $x, y \in X$, defines a metric on X, the so-called *induced metric*. We call $\|\cdot\|$ a *norm* and $(X, \|\cdot\|)$ a *normed space*. By considering the induced metric, every normed space can thus be seen as a metric space. We note that the metric d_∞ defined on $\mathrm{C_b}(X) = \mathrm{C_b}(X;\mathbb{R})$ can be written as $d_\infty(f,g) = \sup_{x\in X}|f(t) - g(t)| = \|f-g\|_\infty$, where $\|h\|_\infty = \sup_{x\in X}|h(t)|$ defines a norm on $\mathrm{C_b}(X)$. Therefore, $\mathrm{C_b}(X)$ is a even a normed space. Recall that a series $\sum_{k=1}^\infty x_k$ in a normed space was defined by the sequence of partial sums $(\sum_{k=1}^n x_k)_{n\in\mathbb{N}^*}$.

Theorem 3.28. *Let (X, d_X) be a metric space and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence in $\mathrm{C_b}(X)$. If*

$$\sum_{k=1}^\infty \|f_k\|_\infty = \sum_{k=1}^\infty \sup_{x\in X}|f_k(x)| < +\infty,$$

then $\sum_{k=1}^\infty f_k$ converges uniformly, that is, in $(\mathrm{C_b}(X), d_\infty)$.

Proof. By what has been stated above the theorem, $(\mathrm{C_b}(X), \|\cdot\|_\infty)$ normed space with the norm inducing d_∞. Furthermore, we know by Theorem 3.26 and Remark 3.27 that every Cauchy sequence in $\mathrm{C_b}(X)$ converges. By Exercise 2.49 we can thus conclude from $\sum_{k=1}^\infty \|f_k\|_\infty < +\infty$ that the series $\sum_{k=1}^\infty f_k$ converges in $(\mathrm{C_b}(X), d_\infty)$. ❑

Remark 3.29. We point out that the statements of Theorem 3.26 and Theorem 3.28 hold in an analogous way for the metric space $\mathrm{F}_b(X;\mathbb{R})$ of *bounded functions* $f\colon X \to \mathbb{R}$ on any *set* X equipped with the metric d_∞. Analogously, here d_∞ is induced by $\|\cdot\|_\infty$. Furthermore, they also hold in analogous way for $\mathbb{C}$-valued functions $\mathrm{F}_b(X;\mathbb{C})$, where $\mathbb{C}$ is equipped with the Euclidean distance d_2. See also Remark 2.10. ✧

Following the remark above we formulate the following variant of the above results.

Theorem 3.30 (Weierstraß test)**.** *Let $(X, \|\cdot\|)$ be a metric space and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of bounded functions $f_n\colon X \to \mathbb{R}(\mathbb{C})$, $n \in \mathbb{N}^*$. Let also $(M_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers such that*

$$\sum_{n=1}^{\infty} M_n < \infty$$

and

$$\forall x \in X \ : \ |f_n(x)| \leq M_n.$$

Then $\sum_{n=0}^{\infty} f_n$ converges uniformly and absolutely, that is, $\sum_{n=0}^{\infty} |f_n(x)|$ converges for every $x \in X$.

3.4 Exercises

Exercise 3.1.

(a) Find all limit points and all *isolated points* (that is, all points of D which are not limit points of D) of $D := \left\{ \frac{\mathrm{i}^n}{n} \;\middle|\; n \in \mathbb{N}^* \right\}$ (where i is the imaginary unit) in $\mathbb{C}$ equipped with $|\cdot|$.

(b) Let $(x_n)_{n\in\mathbb{N}^*}$ be a sequence of real numbers and $D := \{x_n \mid n \in \mathbb{N}^*\}$. Are the limit superior and limit inferior always limit points of D?

(c) Let $M \neq \emptyset$ and d_0 the discrete metric on M. Show that every subset D of M has no limit points w.r.t. d_0. ✧

Exercise 3.2. Let $\mathbb{R}$ be equipped with the standard metric, $D \subseteq \mathbb{R}$ and $f\colon D \to \mathbb{R}$. Determine whether the following limits (of the form $\lim_{x\to x_0} f(x)$) exist and determine their values (if they exist).

(a) $\lim\limits_{x\to 7} x^2 - 2\sqrt{x}$ on $D := \{x \in \mathbb{R} \mid x \geq 0\}$

(b) $\lim\limits_{x\to 4} \dfrac{\sqrt{x} - 2}{x - 4}$ on $D := \{x \in \mathbb{R} \mid x \geq 0 \wedge x \neq 4\}$

(c) $\lim\limits_{x\to 0} \dfrac{|x|}{x}$ on $D := \mathbb{R} \setminus \{0\}$

(d) $\lim_{x\to 0} \frac{|x|}{x^2}$ on $D := \mathbb{R}\setminus\{0\}$ ✧

Exercise 3.3. Let $\mathbb{R}$ be equipped with the standard metric and $n \in \mathbb{N}_0$. Show that $f\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto x^n$ is continuous only using the definition of continuity. ✧

Exercise 3.4. Let $\rho > 0$ and $g\colon (0,\rho)\to\mathbb{R}$. Show that

$$\lim_{x\to 0^+} g(x) = \lim_{x\to+\infty} g(\tfrac{1}{x}),$$

where one side exists if the other exists. ✧

Exercise 3.5. Let $D\subseteq\mathbb{R}$ and $f\colon D\to\mathbb{R}$ considered with the standard metric. Let $x\in\mathbb{R}$ be a limit point of the sets D, $(x,\infty)\cap D$ and $(-\infty,x)\cap D$. Show that $\lim_{t\to x} f(t)$ exists if and only if both one-sided limits $\lim_{t\to x^+} f(t)$ and $\lim_{t\to x^-} f(t)$ exist and are equal. ✧

Exercise 3.6. Let $\mathbb{R}$ be equipped with the standard metric. Show that $f\colon\mathbb{R}\to\mathbb{R}$ given by

$$f(x) := \begin{cases} -x^3+3, & x\le 1,\\ \frac{x^2+1}{x}, & x>1,\end{cases}$$

is a continuous function. ✧

Exercise 3.7. Let $\mathbb{R}$ be equipped with the standard metric, $a,b\in\mathbb{R}$ and $f\colon\mathbb{R}\to\mathbb{R}$ be defined by

$$f(x) := \begin{cases} x^2, & x<2,\\ ax+b, & x\in[2,4],\\ \frac{x^2+2}{x-1}, & x>4.\end{cases}$$

Find values for a,b such that f is continuous. ✧

Exercise 3.8. Let $(X,\|\cdot\|)$ be a normed space. Show that $f\colon X\to\mathbb{R}$, $x\mapsto\|x\|$ is continuous, i.e., $\|\cdot\|$ is continuous. ✧

Exercise 3.9. Prove that the function $f\colon[0,\infty)\to\mathbb{R}$, $f(x):=\sqrt{x}$, is continuous. ✧

Exercise 3.10. Let $f\colon[0,\infty)\to\mathbb{R}$ be given by

$$f(x) := \begin{cases} (1+x^3)\left(\sqrt{\frac{1}{x}+1}-\sqrt{\frac{1}{x}}\right), & x>0,\\ 0, & x=0.\end{cases}$$

(a) Determine $\lim_{x\to 0^+} f(x)$.

(b) Determine at which values $x\in[0,\infty)$ f is continuous.

(c) Determine the largest set $S \subseteq [0, \infty)$ on which f (more precisely, the restriction of f to S) is uniformly continuous. ✧

Exercise 3.11. Let $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Show that f is continuous at $x = 0$.
Hint: Show that $|f(x) - 1| \leq \frac{|x|}{1-|x|}$ *for small* $|x|$. ✧

Exercise 3.12. Let (X, d_X) and (Y, d_Y) be metric spaces, $f\colon X \to Y$ continuous, and $K \subseteq X$ compact.

(a) Show that $f(K) := \{y \in Y \mid \exists x \in K \text{ with } f(x) = y\}$ is compact.

(b) Show that if X is compact and f is bijective, then $f^{-1}\colon Y \to X$ is continuous.

(c) Let $n \in \mathbb{N}^*$. Show that the n-th root $\sqrt[n]{\cdot}\colon [0, \infty) \to \mathbb{R}$ is continuous. ✧

Exercise 3.13. Let (X, d) be a metric space. Define

$$d_{X\times X}\left(\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right), \left(\begin{smallmatrix}y_1\\y_2\end{smallmatrix}\right)\right) := d_X(x_1, y_1) + d_X(x_2, y_2), \qquad \left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right), \left(\begin{smallmatrix}y_1\\y_2\end{smallmatrix}\right) \in X \times X.$$

Show that $(X \times X, d_{X\times X})$ is a metric space and that $d\colon X \times X \to [0, \infty)$ is continuous. ✧

Exercise 3.14. Let $M \neq \emptyset$ be equipped with the discrete metric d_0. Show that

(a) every function $f\colon M \to \mathbb{R}$ is continuous,

(b) $f\colon \mathbb{R} \to M$ is continuous, if and only if f is constant. ✧

Exercise 3.15. Check whether $f\colon D \to \mathbb{R}$, $x \mapsto 1/x$ is uniformly continuous.

(a) $D := (0, 1]$;

(b) $D := [1, \infty)$; ✧

Exercise 3.16. Let $D_1, D_2 \subseteq \mathbb{R}$, $f\colon D_1 \to D_2$ continuous and continuously invertible, i.e., $f^{-1}\colon D_2 \to D_1$ exists and is continuous. Show that for any $x_0 \in D_1$ and $g\colon \mathrm{B}_\rho(f(x_0)) \to \mathbb{R}$ with $\rho > 0$

$$\lim_{y \to f(x_0)} g(y) = \lim_{x \to x_0} g(f(x))$$

in the sense that one side exists, if the other side exists (and then they coincide). ✧

Exercise 3.17. Find a counterexample to: Let $D \subseteq \mathbb{R}$, $f\colon D \to \mathbb{R}$ continuous at $x_0 \in D$, $\rho > 0$ and $g\colon (f(x_0) - \rho, f(x_0) + \rho) \to \mathbb{R}$ such that $\lim_{y \to f(x_0)} g(y)$ exists. Then

$$\lim_{y \to f(x_0)} g(y) = \lim_{x \to x_0} g(f(x)).$$ ✧

Exercise 3.18. Let $\exp\colon \mathbb{C} \to \mathbb{C}$ be defined by

$$\exp(z) \coloneqq \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

which we have already seen in Exercise 2.43 and Exercise 3.11. We call it the *exponential function*. Show that

(a) The series of functions converges uniformly on bounded sets;

(b) Conclude that exp is continuous;

(c) Show that exp restricted to $\mathbb{R}$ maps to $\mathbb{R}$, i.e., $\exp\big|_{\mathbb{R}}\colon \mathbb{R} \to \mathbb{R}$; ✧

Exercise 3.19. Prove that $x\exp(x) = 2$ for some $x \in (0, 1)$. ✧

Exercise 3.20. Let $n \in \mathbb{N}^*$ and $f_n\colon (0, 1] \to \mathbb{R}$ be defined by

$$f_n(x) \coloneqq \begin{cases} n - n^2 x, & \text{if } x \in (0, \frac{1}{n}), \\ 0, & \text{if } x \in [\frac{1}{n}, 1]. \end{cases}$$

Show that f_n is continuous and show that $(f_n)_{n\in\mathbb{N}^*}$ converges pointwise to 0 (the constant 0 function), but not uniformly. ✧

Exercise 3.21. Let $f\colon [0, 2] \to \mathbb{R}$ be continuous and $f(0) = f(2)$. Prove that there exist $x, y \in [0, 2]$ such that

$$|y - x| = 1 \text{ and } f(x) = f(y).$$

Hint: Consider $g(z) \coloneqq f(z + 1) - f(z)$ for $z \in [0, 1]$. ✧

Exercise 3.22. Consider the sequence $(f_n)_{n\in\mathbb{N}^*}$ of functions $f_n\colon [0, \infty) \to \mathbb{R}$, defined by

$$f_n(x) \coloneqq \frac{x^n}{1 + x^n}, \quad x \in [0, \infty), n \in \mathbb{N}^*.$$

(a) Prove that the sequence $(f_n)_{n\in\mathbb{N}^*}$ converges pointwise to a limit $f\colon [0, \infty) \to \mathbb{R}$.

(b) Does the sequence convergence uniformly?

(c) Show that the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[0, q]$ for every $0 < q < 1$. ✧

Exercise 3.23. Show that if $f\colon [a, b] \to [a, b]$ is continuous, then f has a fixed point, i.e., there is an $\eta \in [a, b]$ such that $f(\eta) = \eta$. ✧

Exercise 3.24. Show that the following functions f are uniformly continuous.

(a) Let $f\colon [0,\infty) \to \mathbb{R}$ be continuous and $\lim_{x\to\infty} f(x) \in \mathbb{R}$.

(b) Let $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := 1/\left(x^2+1\right)$. ✧

Exercise 3.25. Let $n \in \mathbb{N}^*$ and $a_k \in \mathbb{R}$ for all $k \in \mathbb{N}_0$, $0 \le k \le n$, and $a_n \neq 0$. Given a general (non-constant) polynomial $p\colon \mathbb{R} \to \mathbb{R}$ defined by

$$p(x) := a_n x^n + \cdots + a_1 x + a_0,$$

for $x \in \mathbb{R}$. Show the following statements:

(a) If n is odd, then p has at least one zero in $\mathbb{R}$ (that is, $\exists a \in \mathbb{R}$ such that $p(a) = 0$).

(b) If n is even and $a_0 a_n < 0$, then p has at least two zeros in $\mathbb{R}$. ✧

Exercise 3.26. Show that the series of functions (in the variable x)

$$\sum_{n=1}^{\infty} \frac{n^2}{3^{2n} + nx^2}$$

converges uniformly on $[0,\infty)$. ✧

Exercise 3.27. Let $f\colon \mathbb{R} \to \mathbb{R}$ be bounded on the closed interval $[a,b]$ with $a < b$. Suppose that for every pair of points x_1, x_2 satisfying $a \le x_1 \le x_2 \le b$, it holds that

$$f\left(\frac{x_1+x_2}{2}\right) \le \frac{f(x_1)+f(x_2)}{2}.$$

Prove that f is continuous on the open interval (a,b). ✧

Chapter 4

Differentiability of functions in one variable

4.1 The derivative and basics

In the following $\mathbb{R}$ (and $\mathbb{C}$) is always understood with the standard metric d_2.

Definition 4.1. Let I be a nontrivial interval and let $f: I \to \mathbb{R}$ (or $f: I \to \mathbb{C}$). Then f is called *differentiable at $x \in I$* if the limit

$$f'(x) \coloneqq \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

exists. The value $f'(x)$ is called the *derivative of f at x* and is also denoted by $\left(\frac{\mathrm{d}}{\mathrm{d}x} f\right)(x)$ or $\frac{\mathrm{d}f}{\mathrm{d}x}(x)$. We say that f is differentiable if f is differentiable at every $x \in I$. ✧

Fact 4.2. *In the setting of the above definition, we have the following facts, most of them being consequences of what we have seen about limits in Section 3.1.*

(i) *For $t \neq x$, the term*

$$\frac{f(t) - f(x)}{t - x}$$

is called differential quotient *of f between t and x.*

(ii) *If x is not a boundary point of the interval, f is differentiable at x if and only if the right and left-sided limits (see also Definition 3.4 and Fact 3.5)*

$$f'^{+}(x) \coloneqq \lim_{t \to x^{+}} \frac{f(t) - f(x)}{t - x}, \quad f'^{-}(x) \coloneqq \lim_{t \to x^{-}} \frac{f(t) - f(x)}{t - x}$$

exist and are equal, in which case they also equal $f'(x)$.

(iii) *f is differentiable at $x \in X$ if and only if for every sequence $(t_n)_{n\in\mathbb{N}^*}$ in $I \setminus \{x\}$ converging to x, it holds that the sequence $\left(\frac{f(t_n)-f(x)}{t_n-x}\right)_{n\in\mathbb{N}^*}$ converges to a limit independent of the sequence $(t_n)_{n\in\mathbb{N}^*}$.*

(iv) *The property of f being differentiable at $x \in X$ is a "local property" of f*

Example 4.3. Let us take a look at the derivative of some common functions.

(a) Let $n \in \mathbb{N}^*$, $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$, is differentiable at every $x \in \mathbb{R}$ with $f'(x) = nx^{n-1}$. This follows from

$$\begin{aligned}\frac{f(t)-f(x)}{t-x} = \frac{t^n-x^n}{t-x} &= \frac{(t-x)(t^{n-1}+t^{n-2}x+\cdots+tx^{n-2}+x^{n-1})}{t-x}\\ &= (t^{n-1}+t^{n-2}x+\cdots+tx^{n-2}+x^{n-1}).\end{aligned}$$

for $t \neq x$ and the fact that polynomials are continuous. Alternatively, one could substitute $h = t - x$ in the above equation, and expand $f(t) = f(x+h) = (x+h)^n$ by using the binomial formula as well as continuity.

(b) For $c \in \mathbb{R}$, the constant function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = c$ for all $x \in \mathbb{R}$ is differentiable with derivative $f'(x) = 0$ for all $x \in \mathbb{R}$.

(c) The function $f(x) = |x|$ defined on $\mathbb{R}$ is not differentiable at 0. This follows from Fact 4.2 (ii).

(d) The function $\exp\colon \mathbb{R} \to \mathbb{R}$ given by

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad x \in \mathbb{R},$$

is differentiable at every point $x \in \mathbb{R}$ with $\exp'(x) = \exp(x)$, see Exercise 4.1.

(e) The functions $\sin\colon \mathbb{R} \to \mathbb{R}$ and $\cos\colon \mathbb{R} \to \mathbb{R}$ are differentiable with derivatives $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$, see Exercise 4.8. ✧

Proposition 4.4 (Differentiable $\implies$ continuous)**.** *Let I be a nontrivial interval and let $f\colon I \to \mathbb{R}$ (or $f\colon I \to \mathbb{C}$) be differentiable at $x \in I$. Then f is continuous at x.*

Proof. The statement follows by the characterization of continuity by limits Fact 3.7 and

$$\lim_{t\to x} f(t)-f(x) = \lim_{t\to x}\frac{f(t)-f(x)}{t-x}\cdot(t-x) = f'(x)\cdot\lim_{t\to x}(t-x) = f'(t)\cdot 0 = 0,$$

where we used the properties of limits Fact 3.5. ❑

Remark 4.5. The converse of Proposition 4.4 is not true, as Example 4.3 (c) shows. ✧

For the following result recall that if for a continuous function $g\colon I \to \mathbb{R}$ on some nontrivial interval we have that $g(x) \neq 0$ for some $x \in I$, then there exists $\varepsilon > 0$ such that $g(y) \neq 0$ for all $y \in (-\varepsilon + x, x + \varepsilon)$.

Proposition 4.6 (Sum, product and quotient rules)**.** *Let I be a nontrivial interval and let $f, g\colon I \to \mathbb{R}$ (or $f, g\colon I \to \mathbb{C}$) be differentiable at $x \in I$. Then*

(i) *$f + g$ is differentiable at x and $(f+g)'(x) = f'(x) + g'(x)$;*

(ii) *fg is differentiable at x and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.*

(iii) *If additionally $g(x) \neq 0$, then $\frac{f}{g}$ is differentiable at x and $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$.*

Proof. We prove the statements as listed in the proposition

(i) This is clear from properties of limits (or sequences), see Fact 3.5 (ii).

(ii) By the identity, for $t \neq x$, $t \in I$,

$$\begin{aligned}
&\frac{f(t)g(t) - f(x)g(x)}{t - x} \\
&\quad = \underbrace{\frac{f(t)g(t) - f(x)g(x)}{t-x} - f(t)\frac{g(t)-g(x)}{t-x}}_{=\frac{f(t)-f(x)}{t-x}g(x)} + f(t)\frac{g(t)-g(x)}{t-x} \\
&\quad \to f'(x)g(x) + f(x)g'(x)
\end{aligned}$$

as $t \to x$, where we used that f is continuous at x by Proposition 4.4, the assumptions on f and g as well as the properties of limits.

(iii) This follows analogously to the argumentation in (ii) by the identity

$$\frac{\frac{f(t)}{g(t)} - \frac{f(x)}{g(x)}}{t-x} = \frac{1}{g(t)g(x)}\left(\frac{(f(t)-f(x)}{t-x}g(t) - f(x)\frac{g(t)-g(x)}{t-x}\right),$$

for all $t \in \{s \in I \mid g(s) \neq 0\}$. Since this set contains a subinterval in I including the point x, letting $t \to x$, shows the assertion by using continuity of g, by Proposition 4.4, and properties of limits. ❑

Example 4.7. The previous result applied on polynomials gives:

(a) For $n < 0$, the function $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = x^n$ is differentiable at every $x \neq 0$ with $f'(x) = nx^{n-1}$. This follows from the quotient rule as $f(x) = \frac{1}{x^{-n}}$.

(b) Let g and h be polynomials. Then the *rational function* $f = \frac{h}{g}$ defined on the set $\{x \in \mathbb{R} \mid g(x) \neq 0\}$ (which is a union of intervals) is differentiable. ✧

Proposition 4.8 (Chain rule)**.** *Let I, J be nontrivial intervals and let $f\colon I \to \mathbb{R}$ be differentiable at $x \in I$. Let $f(I) \subseteq J$ and let $h\colon J \to \mathbb{R}$ (or $h\colon I \to \mathbb{C}$) be differentiable at $f(x)$. Then $(h \circ f)\colon I \to \mathbb{R}$ is differentiable at x and*

$$(h \circ f)'(x) = h'(f(x))f'(x).$$

Proof. Consider the function $\psi\colon J \to \mathbb{R}$ defined by

$$\psi(s) = \begin{cases} \dfrac{h(s) - h(f(x))}{s - f(x)}, & s \neq f(x), \\ h'(f(x)), & s = f(x). \end{cases}$$

By assumption on h, ψ is continuous. By

$$\frac{h(f(t)) - h(f(x))}{t - x} = \frac{h(f(t)) - h(f(x))}{f(t) - f(x)} \cdot \frac{f(t) - f(x)}{t - x} = \psi(f(t)) \cdot \frac{f(t) - f(x)}{t - x}$$

Since f is continuous at x by Proposition 4.4, and since ψ is continuous, $\psi(f(t))$ converges to $\psi(f(x))$ as $t \to x$ as a composition of continuous functions. The quotient on the right-hand side converges to $f'(x)$ because f is differentiable at x. ❑

Example 4.9. Let $f\colon I \to J$, with nontrivial intervals I, J, be differentiable at $x \in I$ and suppose that f is bijective. If $f'(x) \neq 0$ and f and the inverse function f^{-1} are continuous on I, then $f^{-1}\colon J \to I$ is differentiable at $f(x)$ with

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}.$$ ✧

Definition 4.10 (Derivative as function and higher derivatives)**.** Let I be a nontrivial interval and $f\colon I \to \mathbb{R}$ (or $f\colon I \to \mathbb{C}$) be differentiable. Then the function

$$f'\colon I \to \mathbb{R}(\mathbb{C}), x \mapsto f'(x)$$

is called the *derivative of f*. For $n \in \mathbb{N}^*$, we can furthermore recursively define the $(n+1)$*-th derivative* or *derivative of* $(n+1)$*-th order* by

$$f^{(n+1)} := \frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} f := (f^{(n)})',$$

if $f^{(n)}$ is differentiable on I. We also say that f is $(n+1)$-times differentiable. For $n = 1, 2$ we also write f', f''. ✧

Note that in order to define $f^{(n+1)}(x)$ at some point $x \in I$, f needs to be n-times differentiable at every y in an interval containing x.

Definition 4.11. Let I be a nontrivial interval with and $f\colon I \to \mathbb{R}(\mathbb{C})$. We say that f is *continuously differentiable*, if

$$f \text{ is differentiable on } I \text{ and } f'\colon I \to \mathbb{R}(\mathbb{C}) \text{ is continuous.}$$

The set of all continuously differentiable functions $f\colon I \to \mathbb{R}(\mathbb{C})$ is denoted by $\mathrm{C}^1(I;\mathbb{R})$ ($\mathrm{C}^1(I;\mathbb{C})$). More generally, for $n \in \mathbb{N}^*$, we say that a function is n-times continuously differentiable if f is n-times differentiable and $f^{(n)}$ is continuous. We denote the set of these functions by $\mathrm{C}^n(I;\mathbb{R})$ ($\mathrm{C}^n(I;\mathbb{C})$). Finally, we set

$$\mathrm{C}^\infty(I;\mathbb{R}) := \bigcap_{n\in\mathbb{N}^*} \mathrm{C}^n(I;\mathbb{R})$$

and refer to it as the set of infinitely-often differentiable functions. ✧

Example 4.12. The function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is differentiable but not continuously differentiable. ✧

4.2 Local Extrema & the mean value theorem

Definition 4.13. Let (X, d) be a metric space and $f\colon X \to \mathbb{R}$ where $\mathbb{R}$ is equipped with the standard metric. We say that f *has a local maximum (minimum) at* $x_0 \in X$ if there exists $r > 0$ such that

$$\forall y \in \underbrace{\{y \in X \mid d(x_0, y) < r\}}_{=\mathrm{B}_r(x_0)} \;:\; f(x_0) \geq f(y) \;(f(x_0) \leq f(y)).$$

Local maxima and minima are called *local extrema* (*local extremum* for singular use). ✧

Proposition 4.14. *Let I be a nontrivial interval and $f\colon I \to \mathbb{R}$ be differentiable on I. Let x_0 be in I but not a boundary point of I. If f has a local extremum at x_0, then $f'(x_0) = 0$.*

Proof. If f has a local maximum at x_0, which is not a boundary point of the interval, then there exists $r > 0$ such that $I' := (x_0 - r, x_0 + r) \subseteq [a, b]$ and $f(x) \leq f(x_0)$ for all $x \in I'$. By definition of $f'(x_0)$ and Fact 4.2 (ii),

$$\lim_{h\to 0^-} \frac{1}{h}(f(x_0 + h) - f(x_0)) = f'(x_0) = \lim_{h\to 0^+} \frac{1}{h}(f(x_0 + h) - f(x_0)).$$

Since $\frac{1}{h}(f(x_0+h)-f(x_0)) \geq 0$ for $h \in (-r, 0)$ and $\frac{1}{h}(f(x_0+h)-f(x_0)) \leq 0$ for $h \in (0, r)$, we conclude that $f'(x_0) \geq 0$ and $f'(x_0) \leq 0$, whence $f'(x_0) = 0$. The case of f having a local minima at x_0 follows from the proven part by considering $-f$ instead of f. ❑

Corollary 4.15 (Rolle's Theorem)**.** *Let $a < b$ be real numbers and $f\colon [a, b] \to \mathbb{R}$ be continuous and $f|_{(a,b)}$ differentiable. Suppose that $f(a) = f(b)$. Then there exists $\xi \in (a, b)$ such that $f'(\xi) = 0$.*

Proof. By Corollary 3.14 and the compactness of $[a,b]$, we have that the continuous function f has a minimum and maximum, attained at $x_{\min}$ and $x_{\max}$ respectively. If $\{x_{\min}, x_{\max}\} = \{a,b\}$, then $f(a) = f(b) = f(x_{\min}) = f(x_{\max})$ implies that f must be constant, whence, $f'(x) = 0$ for all $x \in (a,b)$. Otherwise, w.l.o.g. $x_{\min} \in (a,b)$. But then, by Proposition 4.14, $f'(x_{\min}) = 0$. ❑

The following result is a special case of a more general theorem which we will prove. Because of its importance we state this special case separately.

Theorem 4.16 (Mean Value Theorem)**.** *Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be continuous functions and $f|_{(a,b)}, g|_{(a,b)}$ be differentiable. Then there exists $\xi \in (a,b)$ such that*

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

Proof. This is a special case of the following result, Theorem 4.17. ❑

Theorem 4.17 (Generalized Mean Value Theorem)**.** *Let $a < b$ be real numbers and $f, g\colon [a,b] \to \mathbb{R}$ be continuous functions and $f|_{(a,b)}, g|_{(a,b)}$ be differentiable. Suppose that $g'(t) \neq 0$ for all $t \in (a,b)$ Then $g(a) \neq g(b)$ and there exists $\xi \in (a,b)$ such that*

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

Proof. The idea is to apply Rolle's theorem, Corollary 4.15 to a suitable function $h\colon [a,b] \to \mathbb{R}$. First, we note that $g(a) \neq g(b)$ by contraposition of Corollary 4.15 since, by assumption, $g(t) \neq 0$ for all $t \in (a,b)$. Therefore, the function $h\colon [a,b] \to \mathbb{R}$, given by

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

is well-defined, continuous and $h|_{(a,b)}$ is differentiable as composition of continuous resp. differentiable functions. Furthermore, $g(a) = g(b) = 0$. Therefore, by Corollary 4.15, there exists $\xi \in (a,b)$ such that

$$f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi) = h'(\xi) = 0,$$

which shows the assertion. ❑

Corollary 4.18. *Let I be an interval with boundary points (included or not) $a < b$, where $a = -\infty$ and $b = \infty$ are allowed. Let $f\colon I \to \mathbb{R}$ be continuous and $f|_{(a,b)}$ be differentiable. The following assertions hold.*

(i) *If $f' \geq 0$ ($f' > 0$) on (a,b), that is, $f'(t) \geq 0$ ($f'(t) > 0$) for all $t \in (a,b)$, then f is increasing (increasing) on I.*

(ii) *If $f' \leq 0$ ($f' < 0$) on (a, b), that is, $f'(t) \leq 0$ ($f'(t) < 0$) for all $t \in (a, b)$, then f is decreasing (decreasing) on I.*

(iii) *If $f' = 0$ on (a, b), then f is constant on I.*

(iv) *If f is increasing (decreasing) on I, then $f' \geq 0$ ($f' \leq 0$) on (a, b).*

Proof. The proof of the first two assertions follows directly from the Mean Value Theorem, Theorem 4.16 since, e.g., $f' \leq 0$ implies that

$$\forall a < c < d < b\ \exists \xi_{c,d} \in (c, d) \ :\ \frac{f(d) - f(c)}{d - c} = f'(\xi_{c,d}) \leq 0.$$

Therefore, $f(d) - f(c) \leq f'(\xi_{c,d})(d - c) \leq 0$ for all $c, d \in (a, b)$ with $c < d$. That the inequality $f(d) \leq f(c)$ even holds for $c, d \in I$ with $c < d$ follows by continuity.

The third assertion is clear from the first two.

The last assertion is immediate from the definition of the derivative as limit of differential quotients and the properties of limits with respect to inequalities. ❑

The (strictly) increasing function $f(x) = x^3$ defined for $x \in \mathbb{R}$ shows that f increasing does in general not imply that the derivative is strictly greater than 0.

The following result can be understood as some kind of intermediate value theorem for the derivative of a differentiable function. It is surprising in the light of the already seen intermediate value theorem (Theorem 3.15) since the derivative of a differentiable function need not be continuous, Example 4.12.

Lemma 4.19. *Let I be a nontrivial interval and $f\colon I \to \mathbb{R}$ be differentiable. Then*

$$f'(I) = \{f'(x) \in \mathbb{R} \mid x \in I\}$$

is an interval.

Proof. We sketch the idea of the proof and leave it as an exercise. In order to show that $f'(I)$ is an interval, it suffices to show that for any points x_1 and x_2 in I with $x_1 < x_2$, it follows that $c \in f'(I)$ for any c between $f'(x_1)$ and $f'(x_2)$, that is any $c \in (\min\{f'(x_1), f'(x_2)\}, \max\{f'(x_1), f'(x_2)\}$. To do so, consider the function

$$g\colon [x_1, x_2] \to \mathbb{R}, x \mapsto f(x) - cx,$$

which is continuous since f is differentiable. Therefore, g has a maximum and a minimum. Note that $f'(x_1) - c = g'(x_1)$ and $f'(x_2) - c = g'(x_2)$. By considering differential quotients at x_1 and x_2, it can now be shown that g does not attain both its maximum and its minimum at the points x_1, x_2, i.e., at least one of the extrema is in the open interval (x_1, x_2). By Proposition 4.14, we conclude that $0 = g'(x) = f'(x) - c$ for some $x \in (x_1, x_2)$. ❑

Theorem 4.20 (Rule of de L'Hospital). *Let $-\infty \leq a < b \leq +\infty$ be real numbers or $\pm\infty$ and $f, g\colon (a,b) \to \mathbb{R}$ be differentiable with $g'(x) \neq 0$ for all $x \in (a,b)$. Further assume* either

(A) $\lim_{x\to a^+} f(x) = 0 = \lim_{x\to a^+} g(x)$, or

(B) $\lim_{x\to a^+} f(x) = \pm\infty = \lim_{x\to a^+} g(x)$.

The following implication holds:

$$\text{If } A := \lim_{x\to a^+} \frac{f'(x)}{g'(x)} \in \mathbb{R}\cup\{\pm\infty\} \text{ exists, then } \lim_{x\to a^+} \frac{f(x)}{g(x)} \text{ exists and equals } A.$$

The same statement holds if

$$\lim_{x\to a^+} \quad \text{is replaced by} \quad \lim_{x\to b^-} \quad \text{or} \quad \lim_{x\to x_0}$$

for some $x_0 \in (a,b)$ everywhere *above.*

Proof. See lecture. ❑

Example 4.21. We have already encountered the exponential function $\exp\colon \mathbb{R} \to \mathbb{R}$ and its inverse $\ln\colon (0,\infty) \to \mathbb{R}$ in the tutorials. If we want to compute the limit

$$\lim_{x\to\infty} \frac{x}{\exp x},$$

we can use de L'Hospital's rule because $f(x) = x$ and $g(x) = \exp(x)$ are differentiable on $\mathbb{R}$ and, by the properties of the exponential function, we have that $\lim_{x\to\infty} \exp(x) = +\infty$. Again, by using that $\lim_{x\to\infty} \exp(x) = +\infty$ as well as $\exp' = \exp$ (see tutorials), we have

$$\lim_{x\to\infty} \frac{1}{\exp'(x)} = 0.$$

Thus $\lim_{x\to\infty} \frac{x}{\exp x} = 0$ as well. Similarly, we can compute

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+} \frac{\ln(x)}{\frac{1}{x}},$$

where we can use that $\lim_{x\to 0^+} \ln(x) = -\infty$ (see below or the tutorials), and that ln is differentiable on $(0,\infty)$ with $\ln'(x) = \frac{1}{x}$ for $x > 0$, which follows from Example 4.9. Now by

$$\lim_{x\to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x\to 0^+} -x = 0,$$

we obtain the limit $\lim_{x\to 0^+} x\ln(x) = 0$ by the rule of de L'Hospital. ✧

4.3 Taylor's theorem (for functions in one variable)

As a consequence of the (generalized) Mean Value Theorem, Theorem 4.17, we derive the following result on how to approximate differentiable functions by polynomials.

Definition 4.22 (Taylor polynomial)**.** Let $k \in \mathbb{N}_0$ and I be a nontrivial interval. For a k-times differentiable function $f\colon I \to \mathbb{R}(\mathbb{C})$ and some point $y \in I$, we call the polynomial $T_k(f)$ given by

$$T_k(f)(x) = T_{k,y}(f)(x) = \sum_{n=0}^{k} \frac{f^{(n)}(y)}{n!}(x-y)^n$$

the *k-th order Taylor polynomial of f (centered) at y.* ✧

Theorem 4.23 (Taylor's theorem)**.** *Let $k \in \mathbb{N}_0$ and I be a nontrivial interval and $f\colon I \to \mathbb{R}$ k-times continuously differentiable such that $f^{(k)}|_{(a,b)}$ is differentiable, where a and b denote the boundary points of I.*
Let $y \in I$. Then for every $x \in I \setminus \{y\}$ there exists $\xi \in (\min\{x,y\}, \max\{x,y\})$ such that

$$R_{k,f,y}(x) \coloneqq f(x) - T_{k,f,y}(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-y)^{k+1}.$$

The latter term is also called the error of the k-th order Taylor polynomial (centered at y)*.*

Proof. Given $x, y \in I$ with $x \neq y$, we define $F, G\colon [\min\{x,y\}, \max\{x,y\}] \to \mathbb{R}$ by

$$F(t) \coloneqq f(x) - T_{k,f,t}(x), \quad G(t) \coloneqq (x-t)^{k+1}, \quad t \in \big[\min\{x,y\}, \max\{x,y\}\big],$$

(note that the term in the definition of F is a $T_{k,f,t}(x)$ not $T_{k,f,x}(t)$). The idea is to apply the generalized Mean Value Theorem (for the functions F and G), which is why we need to check the assumptions first. Both functions are continuous on their domain and differentiable on the open interval $(\min\{x,y\}, \max\{x,y\})$ by the assumptions on f. Therefore, by Theorem 4.17, there exists $\xi \in (\min\{x,y\}, \max\{x,y\})$ such that

$$\frac{R_{k,f,x}(y)}{(x-y)^{k+1}} = \frac{F(y)-F(x)}{G(y)-G(x)} \overset{\text{MVT}}{=} \frac{F'(\xi)}{G'(\xi)}. \tag{4.1}$$

Also note that by definition of F, G, we have by the product rule and exploiting a telescoping sum that

$$\begin{aligned} F'(\xi) &= -\sum_{n=1}^{k} \frac{f^{(n)}(\xi)}{n!} n(-1)(x-\xi)^{n-1} - \sum_{n=0}^{k} \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n \\ &= -\frac{f^{(k+1)}(\xi)}{k!}(x-\xi)^k \end{aligned}$$

and $G'(\xi) = -(k+1)(x-\xi)^k$. Plugging this in (4.1) yields the assertion. ❑

Fact 4.24. *Let the assumptions of Theorem 4.23 be satisfied.*

(i) *The Taylor polynomial $T_{k,f,y}$ is a polynomial of degree k.*

(ii) *There exist many other ways to represent the error of the Taylor polynomial. E.g., if, in the proof of Theorem 4.23, we more generally define G by*

$$G(t) = (x-t)^p,$$

for some $p \in \mathbb{N}^*$, then we derive the expression

$$R_{k,f,y}(x) = \frac{f^{(k+1)}(y)}{k!p}(x-y)^p(x-\xi)^{k-p+1}.$$

Using the substitution $\xi = \theta x + (1-\theta)y$ for $\theta \in (0,1)$, we get the Schlömilch form *of the error. The error form we have seen in Theorem 4.23 is usually called* Cauchy form.

(iii) *If additionally $f^{k+1}(x) = 0$ for all $x \in (a,b)$, then the error term vanishes and hence $f(x) = T_{k,f,y}(x)$ for all $x \in I$ (by continuity on the whole of I). Therefore, the solutions of the (differential) equation*

$$f^{(k+1)}(x) = 0 \qquad x \in (a,b), \tag{4.2}$$

that is, all the continuously differentiable functions on I, which are differentiable on (a,b) and satisfy (4.2) are given by all polynomials of degree less than $k+1$.

Example 4.25. We represent some examples for the Taylor polynomial.

(a) Let $p\colon \mathbb{R} \to \mathbb{R}$ be a polyonmial. Then $T_{k,p,0}(x) = p(x)$ for all $k \in \mathbb{N}^*$.

(b) Let $f = \exp\colon \mathbb{R} \to \mathbb{R}$. Then $T_{k,\exp,0}(x) = \sum_{n=0}^{k} \frac{x^n}{n!}$ for all $k \in \mathbb{N}^*$, where we used that $\exp^{(n)} = \exp$.

(c) Let $f\colon (0,\infty) \to \mathbb{R}, x \mapsto \sqrt{x}$ and $y > 0$. Then f is infinitely often continuously differentiable on $(0,\infty)$ and the second Taylor polynomial at $y = 1$ is given by

$$T_{2,\sqrt{\cdot},1}(x) = \sqrt{1} + \frac{1}{2\sqrt{1}}(x-1) - \frac{1}{8}(x-1)^2. \qquad ✧$$

Corollary 4.26. *Let $a < b$, $m \in \mathbb{N}^* \setminus \{1\}$ and $f\colon (a,b) \to \mathbb{R}$ be m-times differentiable such that $f^{(k)}$ is continuous at some point $y \in (c,d)$. Furthermore, suppose that*

$$f'(y) = f^{(2)}(y) = \cdots = f^{(m-1)}(y) = 0 \quad \textit{and} \quad f^{(m)}(y) \neq 0.$$

(i) *If m is even, then f has a strict local extremum at y, which is a local maximum if $f^{(m)} < 0$ and a minimum otherwise.*

(ii) *If m is odd, then there is no local extremum at y.*

Proof. We use Theorem 4.23 for $k = m - 1 \in \mathbb{N}^*$, to see that for all $x \in (c, d)$,

$$f(x) = R_{m-1,f,y}(x) + T_{m-1,f,y}(x) = R_{m-1,f,y}(x) + f(y)$$
$$= \frac{f^m(\xi)}{m!}(x-y)^m + f(y).$$

for some $\xi \in (\min\{x, y\}, \max\{x, y\})$, where we used that $T_{m-1,f,y}(x) = f(y)$ by the assumption that the derivatives of f vanish at y.

(i) If m is even, we conclude that the sign of the last term equals the one of $f^m(\xi)$, which, for x sufficiently close to y, equals the sign of $f^m(x)$ by the assumed continuity of $f^{(m)}$ at x. Therefore, $f(x) - f(y)$ is either positive or negative for such y, which proves the first assertion.

(ii) If m is odd, we see that, $(x-y)^m > 0$ if and only if $x > y$, which implies that for every $\varepsilon > 0$ there exists x_1 and x_2 such that $|x_1 - y| < \varepsilon$ and $|x_2 - y| < \varepsilon$ and

$$f(x_1) - f(y) > 0 \quad \wedge \quad f(x_2) - f(y) < 0.$$

Therefore, y cannot be a local extremum. ❑

4.4 Elementary functions: exp, sin, cos

In the following we finally formally define elementary functions such as the exponential functions and the trigonometric functions. Partially this has been executed in the tutorials already. It should be noted that all these functions are by no means "trivially" defined by what one "knows" from school, but, in the course of a proper approach to analysis, require rigorous treatment.

Definition 4.27. For $x \in \mathbb{R}$ or $x \in \mathbb{C}$ define

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

which converges (absolutely) by the ratio test for infinite series. We typically consider exp as function from $\mathbb{R}$ to $\mathbb{R}$, and write $\exp_{\mathbb{C}}$ if we refer to the function from $\mathbb{C}$ to $\mathbb{C}$ (which clearly extends the one defined on $\mathbb{R}$). We call this function the *exponential function*. We further define the *sine (function)*

$$\sin\colon \mathbb{R} \to \mathbb{R}, x \mapsto \frac{\exp_{\mathbb{C}}(\mathrm{i}x) - \exp_{\mathbb{C}}(-\mathrm{i}x)}{2\mathrm{i}} = \mathrm{Im}(\exp_{\mathbb{C}}(\mathrm{i}x)),$$

and the *cosine (function)*

$$\cos\colon \mathbb{R} \to \mathbb{R}, x \mapsto \frac{\exp_{\mathbb{C}}(\mathrm{i}x) + \exp_{\mathbb{C}}(-\mathrm{i}x)}{2} = \mathrm{Re}(\exp_{\mathbb{C}}(\mathrm{i}x))$$

where i refers to the imaginary unit and $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ are the real and imaginary part of z respectively and where we used that $\overline{\exp(\mathrm{i}x)} = \exp(-\mathrm{i}x)$ (which follows from $\overline{\mathrm{i}^n} = (-\mathrm{i})^n$ and since $z \mapsto \overline{z}$ is continuous). ✧

Let us collect some basic properties of the exponential function

Proposition 4.28 (Properties of the exponential function)**.** *The following properties hold.*

(i) $\exp_{\mathbb{C}} \colon \mathbb{C} \to \mathbb{C}$ *is an extension of* $\exp\colon \mathbb{R} \to \mathbb{R}$*. The defining series of* $\exp$ *(and* $\exp_{\mathbb{C}}$*) converges uniformly on compact sets in* $\mathbb{R}$ *(and* $\mathbb{C}$*).*

(ii) $\exp(0) = 1$ *and* $\exp(x) \geq 1 + x$ *for all* $x \geq 0$.

(iii) $\exp\colon \mathbb{R} \to \mathbb{R}$ *and* $\exp_{\mathbb{C}} \colon (\mathbb{C}, d_2) \to (\mathbb{C}, d_2)$ *are continuous. Moreover,* $\exp\colon \mathbb{R} \to \mathbb{R}$ *is infinitely often differentiable,* $\exp \in \mathrm{C}^\infty(\mathbb{R}, \mathbb{R})$ *and* $\exp = \exp' = \exp^{(k)}$ *for any* $k \in \mathbb{N}^*$.

(iv) *The function* $\mathbb{R} \to \mathbb{C}, x \mapsto \exp_{\mathbb{C}}(\mathrm{i}x)$ *is in* $\mathrm{C}^\infty(\mathbb{R}, \mathbb{C})$

(v) *For every* $a \in \mathbb{C}$, $\exp_{\mathbb{C}}(ax) \exp_{\mathbb{C}}(-ax) = 1$.

(vi) *For all* $a, f_0 \in \mathbb{C}$, $f\colon \mathbb{R} \to \mathbb{C}, x \mapsto f_0 \exp_{\mathbb{C}}(ax)$ *is the differentiable function* f *solving*

$$f'(x) = af(x) \quad \forall x \in \mathbb{R}, \qquad f(0) = f_0.$$

(vii) *For all* $a, b \in \mathbb{C}$, $\exp_{\mathbb{C}}(a + b) = \exp_{\mathbb{C}}(a) \exp_{\mathbb{C}}(b)$.

(viii) $\exp_{\mathbb{C}}(a) \neq 0$ *for all* $a \in \mathbb{C}$ *and* $\exp(x) > 1$ *for all* $x \in (0, \infty)$*. Furthermore,* $0 < \exp(x) < 1$ *for all* $x < 0$.

(ix) $\exp$ *is (strictly) increasing (on* $\mathbb{R}$*) and* $\lim_{x\to\infty} \exp(x) = \infty$ *as well as* $\lim_{x\to-\infty} \exp(x) = 0$

(x) $\exp\colon \mathbb{R} \to (0, \infty)$ *is bijective*

(xi) *For every* $m \in \mathbb{Z}$ *and* $n \in \mathbb{N}^*$, $\exp\left(\frac{m}{n}\right) = (\mathrm{e}^m)^{\frac{1}{n}}$ *(with* $\mathrm{e}^0 = 1$*). In particular,* $\exp(1) = \mathrm{e}$.

Proof. See Exercises 2.43, 4.1, 4.3, 4.4 and 4.6. ❑

Definition 4.29 (The logarithm)**.** The inverse function of the exponential function $\exp\colon \mathbb{R} \to (0, \infty)$ is called *the (natural) logarithm* and denoted by ln (or sometimes log). ✧

Proposition 4.30 (Properties of the logarithm)**.** *The following assertions hold.*

(i) $\ln\colon (0, \infty) \to \mathbb{R}$ *is strictly increasing and* $\lim_{x\to 0^+} \ln(x) = -\infty$ *and* $\lim_{x\to\infty} \ln(x) = \infty$.

(ii) $\ln$ *is in* $\mathrm{C}^\infty((0, \infty); \mathbb{R})$.

(iii) $\ln(1) = 0$ *and* $\ln(\mathrm{e}) = 1$.

Definition 4.31 (General powers of postive real numbers)**.** Let $a > 0$ and $x \in \mathbb{R}$. We define
$$a^x := \exp(x \ln(a)) = \mathrm{e}^{x \ln(a)}.$$ ✧

Proposition 4.32. *Let $a > 0$ and $x, y \in \mathbb{R}$.*

(i) $a^0 = 1$, $(a^x)^y = a^{xy}$ *and* $a^x a^y = a^{x+y}$.

(ii) *If* $x = \frac{m}{n}$ *for* $m \in \mathbb{Z}$ *and* $n \in \mathbb{N}^*$, *then* $a^x = (a^m)^{\frac{1}{n}}$

(iii) $\mathrm{e}^x = \exp(x)$ *for all* $x \in \mathbb{R}$

The last property motives to use the slightly shorter notation e^x *for the exponential function whenever convenient.*

Proposition 4.33 (Properties of sin and cos)**.** *For all $a, b, x, y \in \mathbb{R}$ we have that*

(i) $\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$ *and* $\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$ *where the convergence is uniform on compact sets in* $\mathbb{R}$.

(ii) sin *and* cos *are infinitely often differentiable with*
$$\sin' = \cos \quad \textit{and} \quad \cos' = -\sin.$$

(iii) *Euler's formula:* $\exp(\mathrm{i}x) = \cos(x) + \mathrm{i}\sin(x)$.

(iv) *De Moivre's formula:* $(\cos(x) + \mathrm{i}\sin(x))^n = \cos(nx) + \mathrm{i}\sin(nx)$ $n \in \mathbb{N}^*$.

(v) *For all* $\exp(x + \mathrm{i}y) = \exp(x)(\cos(y) + \mathrm{i}\sin(y))$.

(vi) $\cos^2(x) + \sin^2(x) = 1$. *In particular,* $\cos(x)$ *and* $\sin(x)$ *take values only within* $[-1, 1]$.

(vii) $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ *and* $\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$

(viii) cos *has a smallest positive zero* x_0, *which lies in the interval* $(0, 2)$. *We define* $\pi := 2x_0$.

(ix) $\sin(x + \frac{\pi}{2}) = \cos(x)$, $\sin(-x) = -\sin(x)$ *and* $\cos(x) = \cos(-x)$.

(x) sin *and* cos *are* 2π*-periodic, that is,* $\sin(x + 2\pi) = \sin(x)$ *and* $\cos(x + 2\pi) = \cos(x)$.

(xi) $\cos(k\pi) = (-1)^k$, $\sin(k\pi) = 0$ *for all* $k \in \mathbb{Z}$.

(xii) $\cos(x) = 0 \iff x \in \frac{\pi}{2} + \pi\mathbb{Z}$ *and* $\sin(x) = 0 \iff x \in \pi\mathbb{Z}$.

(xiii) *The restriction of* sin *to the interval* $[-\pi/2, \pi/2]$ *and the restriction of* cos *to* $[0, \pi]$ *are bijective onto the interval* $[-1, 1]$.

Proof. See tutorials and lecture. ❑

By the above properties, the following functions are well-defined.

Definition 4.34 (Tangent, and inverse of sin and cos)**.** We define the inverses of the trigonometric functions:

- The function $\tan\colon (-\pi/2, \pi, 2) \to \mathbb{R}, x \mapsto \frac{\sin(x)}{\cos(x)}$ is called the *tangent (function)*. Its inverse is called the *arcus tangent* and is denote by $\arctan\colon \mathbb{R} \to (-\pi/2, \pi/2)$.
- The inverse function of $\sin|_{[-\pi/2,\pi/2]}$ defined on $[-1, 1]$ is called *arcus sine (or arcsin)* and denoted by $\arcsin\colon [-1, 1] \to [-\pi/2, \pi/2]$.
- The inverse function of $\cos|_{[0,\pi]}$ defined on $[-1, 1]$ is called *arcus cosine (or arccos)* and denoted by $\arccos\colon [-1, 1] \to [0, \pi]$ ✧

4.5 Primitives (Anti-derivatives)

Definition 4.35 (Primitive/Antiderivative)**.** Let I be a nontrivial interval and $f\colon I \to \mathbb{R}$. A function $F\colon I \to \mathbb{R}$ is called *primitive (or antiderivative* if F is differentiable and $F' = f$. ✧

It can be useful to also define the following functions

$$\cosh\colon \mathbb{R} \to \mathbb{R}, x \mapsto \frac{\exp(x) + \exp(-x)}{2},$$

called *cosine hyperbolicus* and

$$\sinh\colon \mathbb{R} \to \mathbb{R}, x \mapsto \frac{\exp(x) - \exp(-x)}{2},$$

called *sine hyperbolicus*. It is not hard to see that they also share properties similar to the ones of sin and cos (even simpler to prove), like

$$\sinh' = -\cosh \text{ and } \cosh' = \sinh'$$

or $\cosh^2(x) - \sinh^2(x) = 1$ for all $x \in \mathbb{R}$. The sinh function is bijective from $\mathbb{R}$ to $\mathbb{R}$ and cosh is bijective when considered from $[1, \infty)$ to $[0, \infty)$.

Proposition 4.36. *Let $I, J \subseteq \mathbb{R}$ be nontrivial intervals and S_f, S_g primitves of $f\colon I \to \mathbb{R}$ and $g\colon I \to \mathbb{R}$. Then we have:*

(i) *For every $c \in \mathbb{R}$, $S_f + c$ is a primitive of f. Conversely, if F is a primitive of f, then there exists $c \in \mathbb{R}$ such that $F = S_f + c$.*

(ii) *For every $\lambda \in \mathbb{R}$, $\lambda S_f + S_g$ is a primitive of $\lambda f + g$.*

(iii) *If f, g are differentiable and if $S_{fg'}$ is a primitive of fg', then $fg - S_{fg'}$ is a primitive of $f'g$.*

(iv) *For differentiable $h\colon J \to I$, we have that $S_f \circ h\colon J \to \mathbb{R}$ is a primitive of $(f \circ h) \cdot h'$.*

Proof. That $S_f + c$ is a primitive f follows from linearity of the derivative and since the derivative of a constant function on an interval is zero. The second part of the first assertion follows from $(F - S_f)' = F' - S_f' = 0$. The other assertions follow from linearity of the derivative as well as product and chain rule. ❑

Remark 4.37. By Proposition 4.36 (i), we can rewrite items (ii), (iii) and (iv) as

(ii) $S_{\lambda f+g} = \lambda S_f + S_g + \text{constant}$

(iii) $S_{fg'} = fg - S_{fg'} + \text{constant}$

(iv) $S_{(f\circ h)h'} = S_f \circ h + \text{constant}$

We point out that for the above result (i) it is essential that the considered functions were defined on one interval (and not, e.g., on a union of disjoint intervals). ✧

Definition 4.38. Let $I \subseteq \mathbb{R}$ be a nontrivial interval and $f\colon I \to \mathbb{R}$. The set of all primitives of f is called *the indefinite integral* $\int f$ of f. ✧

Example 4.39. Let in the following I be a nontrivial interval.

(a) $I \subseteq \mathbb{R}, r \geq 0 \,:\, \int x^r = \frac{x^{r+1}}{r+1} + C$

(b) $I \subseteq (-\infty, 0) \cup (0, \infty) \,:\, \int \frac{1}{x} = \ln|x| + C$

(c) $I \subseteq \mathbb{R} \,:\, \int \mathrm{e}^x = \mathrm{e}^x + C$, $\int \sin = -\cos + C$, $\int \cos = \sin + C$, $\int \sinh = \cosh + C$, $\int \cosh = \sinh + C$

(d) $I \subseteq \mathbb{R} \,:\, \int \frac{1}{1+x^2} = \arctan(x) + C$, where $\arctan\colon \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ is the inverse function of tan.[1]

(e) $I \subseteq \mathbb{R}, r > 1 \,:\, \int \frac{x}{(1+x^2)^r} = -\frac{1}{2r(1+x^2)^{r-1}} + C$ and $\int \frac{x}{1+x^2} = \frac{1}{2}\ln(1+x^2) + C$

(f) $I \subseteq \mathbb{R}, k \in \mathbb{N}, k \geq 2 \,:\, \int \cos^k = \frac{\sin\cos^{k-1}}{k} + \frac{k-1}{k}\int \cos^{k-2} + C$, (integration by parts):

$$\begin{aligned}\int \cos^k &= \int \cos\cdot\cos^{k-1}\\ &= \sin\cdot\cos^{k-1} + (k-1)\int \sin^2\cdot\cos^{k-2}\\ &= \sin\cdot\cos^{k-1} + (k-1)\int \cos^{k-1} - (k-1)\int \cos^k\end{aligned}$$

(g) $I \subseteq \mathbb{R}, r > 0 \,:\, \int \frac{1}{(1+y^2)^r} = \frac{1}{(2r-2)}\frac{y}{(1+y^2)^{r-1}} + \frac{2r-3}{2r-2}\int \frac{1}{(1+y^2)^{r-1}}$, (substitution):

$$\begin{aligned}\int \frac{1}{(1+y^2)^r} &= \left| \begin{array}{c} y = \tan(x) = h(x) \\ h'(x) = \cos^{-2} x = 1 + \tan^2 x \end{array} \right|\\ &= \int \frac{1}{(1+\tan^2 x)^r}\cos^{-2}(x)\\ &= \int \cos^{2r-2}(x) \overset{\text{(f)}}{=} \frac{\sin\cos^{2r-3}}{2r-2} + \frac{2r-3}{2r-2}\int \cos^{2r-4}\end{aligned}$$

[1] note that $\arctan'(y) = \frac{1}{1+y^2}$ with rule for differentiating the inverse function.

(h) $I = (0, \infty) : \int \ln = x\ln(x) - x + C$

(i) $I \subseteq (-1, 1) : \int \frac{1}{\sqrt{1-y^2}} = \arcsin(y) + C$ (substitution $y = \sin x$)[2]
$\int \frac{1}{1-y^2} = \frac{1}{2}(\ln|1+y|) - \ln|1-y|) + C$

(j) $I \subseteq \mathbb{R} : \int \frac{1}{\sqrt{x^2+1}} = \operatorname{arsinh}(x) + C = \ln(x + \sqrt{x^2+1}) + C$, (substitution $y = \sinh x$)[3]

(k) $I \subseteq (1, \infty) : \int \frac{1}{\sqrt{x^2-1}} = \operatorname{arcosh}(x) + C = \ln(x + \sqrt{x^2-1}) + C$ (substitution $y = \cosh x$) ✧

4.6 Exercises

Exercise 4.1. Let $z \in \mathbb{C}$, $a \in \mathbb{C}$ and $h \in \mathbb{R}$.

(a) Let $n \in \mathbb{N}^*$. Show that $f_n(h) := \frac{1}{h}[(a(z+h))^n - (az)^n]$, $h \neq 0$, converges to $a^n n z^{n-1}$ as $h \to 0$.

(b) For $n \in \mathbb{N}^*$ let $\widetilde{f}_n(h) := f_n(h)$, $h \neq 0$, and $\widetilde{f}_n(0) := a^n n z^{n-1}$. Show that

- $\left|\widetilde{f}_n(h)\right| \leq a^n 2^n (|z|+1)^{n-1}$ for all $h \in [-1, 1]$ and
- $\sum_{n=1}^{\infty} \frac{1}{n!} a^n 2^n (|z|+1)^{n-1} < \infty$.

(c) Prove that

$$\lim_{h \to 0} \frac{\exp(a(z+h)) - \exp(az)}{h}$$

exists and equals $a\exp(az)$. ✧

Exercise 4.2. We define $\sin\colon \mathbb{C} \to \mathbb{C}$ and $\cos\colon \mathbb{C} \to \mathbb{C}$ by

$$\cos(z) := \frac{1}{2}\left(\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)\right), \quad \sin(z) := \frac{1}{2\mathrm{i}}\left(\exp(\mathrm{i}z) - \exp(-\mathrm{i}z)\right)$$

for $z \in \mathbb{C}$, where i refers to the imaginary unit.

(a) Write cos and sin as series of the form $\sum_{n=0}^{\infty} a_n z^n$ with real numbers a_n, $n \in \mathbb{N}_0$. Do these series converge uniformly on every bounded subset of $\mathbb{C}$?

(b) Show that $\sin(\mathbb{R}) \subseteq \mathbb{R}$ and $\cos(\mathbb{R}) \subseteq \mathbb{R}$.
This allows us to regard sin and cos also as function from $\mathbb{R}$ to $\mathbb{R}$.

(c) Show that $\exp(\mathrm{i}z) = \cos(z) + \mathrm{i}\sin(z)$ for every $z \in \mathbb{C}$. ✧

[2] where $\arcsin\colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\operatorname{arsinh}\colon \mathbb{R} \to \mathbb{R}$, $\operatorname{arcosh}\colon [1, \infty) \to [0, \infty)$ is the inverse function of $\sin, \sinh$ and $\cosh$.

[3] note $\cosh^2 x - \sinh^2 x = 1$

Exercise 4.3. Let us regard the following differential equation for fixed $a \in \mathbb{C}$ and $u_0 \in \mathbb{C}$:

$$\begin{aligned} u'(t) &= a \cdot u(t) \quad \text{for} \quad t \in \mathbb{R}, \\ u(0) &= u_0. \end{aligned} \tag{4.3}$$

(a) Show that $u\colon \mathbb{R} \to \mathbb{C}$, $u(t) := u_0 \cdot \exp(at)$, satisfies the differential equation (4.3).

(b) Show that $\exp(-at)\exp(at) = 1$. *Hint:* Use the product rule.

(c) Show that there is no other $\mathbb{C}$-valued function on $\mathbb{R}$ that satisfies these equations.

Hint: Assume there is another function $\widetilde{u}\colon \mathbb{R} \to \mathbb{C}$ that satisfies (4.3) and regard $t \mapsto \exp(-at)\widetilde{u}(t)$. (You need that only constant functions have derivative 0 (on intervals).) ✧

Exercise 4.4. Let $a, b \in \mathbb{C}$. Show for $t \in \mathbb{R}$

$$\exp((a+b)t) = \exp(at) \cdot \exp(bt).$$

In particular, conclude the *exponentiation identity* for arbitrary $a, b \in \mathbb{C}$

$$\exp(a+b) = \exp(a) \cdot \exp(b).$$

Hint: Use Exercise 4.3. ✧

Exercise 4.5. Let $f\colon (a,b) \to \mathbb{R}$. We say f is a *convex* function, if

$$\forall x_1, x_2 \in (a,b)\, \forall t \in [0,1] \,:\, f\big((1-t)x_1 + tx_2\big) \le (1-t)f(x_1) + tf(x_2).$$

Show

(a) that f is convex is equivalent to

$$\begin{aligned} &\forall x_1, x_2 \in (a,b)\, \forall t \in [0,1] \,: \\ &\qquad\qquad f\big(x_1 + t(x_2 - x_1)\big) \le f(x_1) + t\big(f(x_2) - f(x_1)\big). \end{aligned}$$

What is the geometric meaning of this?

(b) that f is convex is equivalent to

$$a < x_1 < x_2 < x_3 < b \quad \Longrightarrow \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \le \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Hint: Write x_2 as $(1-t)x_1 + tx_3$ for a suitable t.

(c) Conclude that f is continuous on (a,b) ✧

Exercise 4.6. Use the exponentiation identity to show that

(a) $\exp(x) > 0$ for $x \in \mathbb{R}$

(b) exp is strictly increasing on $\mathbb{R}$

(c) $\lim_{x\to+\infty} \exp(x) = +\infty$

(d) $\lim_{x\to-\infty} \exp(x) = 0$

(e) $\exp\colon \mathbb{R} \to (0,+\infty)$ is bijective

Note that this implies that there exists an inverse function of $\exp\colon \mathbb{R} \to (0,+\infty)$ denoted by $\ln\colon (0,+\infty) \to \mathbb{R}$. ✧

Exercise 4.7. Calculate the derivative of the following functions $f\colon (0,+\infty) \to \mathbb{R}$ defined by

(a) $f(x) := \sqrt{x}$,

(b) $f(x) := \sqrt[n]{x}$ for $n \in \mathbb{N}^*$,

(c) $f(x) := \ln(x)$. ✧

Exercise 4.8. Show that $\sin' = \cos$ and $\cos' = -\sin$. ✧

Exercise 4.9. Prove the following *angle sum identities* for $x, y \in \mathbb{C}$

(a) $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$,

(b) $\sin(x+y) = \cos(x)\sin(y) + \sin(x)\cos(y)$,

(c) $\cos^2(x) + \sin^2(x) = 1$. ✧

Exercise 4.10. Find $a, b \in \mathbb{R}$ such that the following function $f\colon \mathbb{R} \to \mathbb{R}$ is differentiable.

$$f(x) := \begin{cases} ax, & x \in (-\infty, 2], \\ b + 2\exp(x-2), & x \in (2, +\infty). \end{cases}$$ ✧

Exercise 4.11. Let $a, b \in \mathbb{C}$. Show for $t \in \mathbb{R}$

$$\exp((a+b)t) = \exp(at) \cdot \exp(bt).$$

In particular, conclude the *exponentiation identity* for arbitrary $a, b \in \mathbb{C}$

$$\exp(a+b) = \exp(a) \cdot \exp(b).$$

Hint: Use Exercise 4.3. ✧

Exercise 4.12. Show that $\ln(x \cdot y) = \ln(x) + \ln(y)$ for all $x, y > 0$ (ln is the inverse function of exp). ✧

Exercise 4.13. For a positive real number x and a rational number $q := \frac{n}{m}$, $n, m \in \mathbb{Z}$ ($m \neq 0$), we can define

$$x^q := \sqrt[m]{x^n}.$$

Show that

$$x^q = \exp(\ln(x)q).$$

Hence, we extend x^q for $q \in \mathbb{R}$ accordingly, i.e., $x^q := \exp(\ln(x)q)$. Solve $5^x = 10$, i.e., find a formula for x. ✧

Exercise 4.14. Show that $\exp(x) = \mathrm{e}^x$ for all $x \in \mathbb{R}$. ✧

Exercise 4.15. For $n \in \mathbb{N}^*$, calculate the limits:

(a) $\displaystyle\lim_{x\to 1-} \frac{x^2-1}{\sqrt{1-x}}$ (b) $\displaystyle\lim_{x\to+\infty} \frac{x^n}{\mathrm{e}^x}$ (c) $\displaystyle\lim_{x\to+\infty} \frac{x\ln(x)}{x^2-1}$ ✧

Exercise 4.16. Compute the n-th Taylor polynomial of $f\colon (0,+\infty) \to \mathbb{R}$, $f(x) := \ln(x)$, at $x_0 := 1$. Finally, conclude the value of $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$. ✧

Exercise 4.17. Show that

$$\frac{t^m}{m!} \geq \frac{t^{m+2}}{(m+2)!}$$

for all $m \in \mathbb{N}^*$, $m \geq 2$, and $t \in [0,2]$. Use this to show that $\cos(2) < 0$. ✧

Exercise 4.18. Calculate the derivative of $f\colon (0,+\infty) \to \mathbb{R}$, $f(x) := x^x$. ✧

Exercise 4.19. For $n \in \mathbb{N}^*$, calculate the limits:

(a) $\displaystyle\lim_{x\to 0} \frac{\sin(x)}{x}$ (b) $\displaystyle\lim_{x\to+\infty} \frac{1}{x^n}\mathrm{e}^{\sqrt{x}}$ (c) $\displaystyle\lim_{x\to 0} \frac{1}{x^n}\mathrm{e}^{-\frac{1}{x^2}}$ ✧

Exercise 4.20. Let $f\colon (a,b) \to \mathbb{R}$ be differentiable. Show that the image of f', i.e., the set

$$f'\big((a,b)\big) = \{f'(x) \mid x \in (a,b)\},$$

is an interval.

Hint: Use

(a) that $J \subseteq \mathbb{R}$ is an interval if and only if for all $y_1, y_2 \in J : y_1 < y_2 \implies [y_1, y_2] \subseteq J$, and

(b) for $x_1, x_2 \in (a,b)$ and $f'(x_1) < c < f'(x_2)$, the function $g\colon [x_1,x_2] \to \mathbb{R}, g(t) := f(t) - ct$. ✧

Exercise 4.21. Compute the n-th Taylor polynomial of $f\colon [0,+\infty) \to \mathbb{R}$, $f(x) := \sqrt{x}$, at $x_0 := 1$. ✧

Exercise 4.22. Let $f\colon (a,b) \to \mathbb{R}$ be such that there exists a $K \geq 0$ and $\alpha > 1$ such that for all $x, y \in (a,b)$

$$|f(x) - f(y)| \leq K|x-y|^{\alpha}.$$

Conclude that f is constant. ✧

Exercise 4.23. Show that the function $\rho\colon \mathbb{R} \to \mathbb{R}$ given by

$$\rho(x) := \begin{cases} \exp(-\frac{1}{x}), & x > 0, \\ 0, & x \leq 0, \end{cases}$$

is in $\mathrm{C}^\infty(\mathbb{R})$ and compute its n-th Taylor polynomial at $x_0 := 0$. ✧

Exercise 4.24. Let $x > 0$ and $\alpha > 0$ show that

$$\frac{\alpha}{(1+x)^{\alpha+1}} < \frac{1}{x^\alpha} - \frac{1}{(1+x)^\alpha}.$$

Conclude that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^{1+\alpha}}$$

converges.

Hint: Regard the function $f\colon (-1, +\infty) \to \mathbb{R}$, $f(y) := \frac{-1}{(1+y)^\alpha}$, and use the mean value theorem. ✧

Exercise 4.25. Use the rule of de l'Hôpital to calculate

$$\lim_{x\to 0} \left(\frac{\pi^2}{\sin^2(\pi x)} - \frac{1}{x^2} \right)$$

Hint: You might need to apply de l'Hôpital's rule multiple times. ✧

Exercise 4.26 (A warning from Otto Stolz)**.** Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) := x + \sin(x)\cos(x)$ and $g(x) := f(x)\exp(\sin(x))$ for $x \in \mathbb{R}$. Show that

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)}$$

does not exist, but

$$\lim_{x\to+\infty} \frac{f'(x)}{g'(x)}$$

exists. Does this contradict the rule of de l'Hôpital? ✧

Exercise 4.27. Let $f\colon (a, b) \to \mathbb{R}$ be a twice differentiable function. Show that f is convex, if and only if $f''(x) \geq 0$ for all $x \in (a, b)$. ✧

Exercise 4.28. Let $f\colon [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Show that $f'^{+}(a)$ exists, if $\lim_{x\to a+} f'(x)$ exists and that in that case $f'^{+}(a) = \lim_{x\to a+} f'(x)$ holds. ✧

Exercise 4.29. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) := x^2 \exp(-x^2).$$

Perform a *curve sketching* for f, i.e.,

(a) find all continuity points of f,

(b) find all points where f is differentiable,

(c) find the intervals, where f is monotonically increasing or decreasing,

(d) find the intervals, where f is convex or concave (f is concave, if $-f$ is convex),

(e) find all zeros of f,

(f) find all local and global extrema of f,

(g) find all inflection points of f, i.e., points where the curvature changes from convex to concave or the other way around.

(h) determine the behavior of f as $x \to \pm\infty$,

(i) draw a sketch of the function. ✧

Exercise 4.30. Let $f, g\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Moreover, let $f(a) = g(a)$ and $f'(x) < g'(x)$ for all $x \in (a,b)$. Show that $f(x) < g(x)$ for all $x \in (a,b]$.

Use this to show that $\exp(x) \geq x^2$ for $x \geq 0$. ✧

Exercise 4.31. Let $f\colon (a,b) \to \mathbb{R}$ be twice differentiable. Show that for $x \in (a,b)$

$$f''(x) = \lim_{h\to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$ ✧

Exercise 4.32. Regard the function $f\colon (0,1) \to \mathbb{R}$ given by

$$f(x) := x^2 \sin\left(\frac{1}{x}\right).$$

Decide whether

(a) f is continuous,

(b) f continuously differentiable,

(c) f can be continuously extended to $[0,1)$,

(d) f can be continuously differentiably extended to $[0,1)$,

(e) f can be differentiably extended to $[0,1)$. ✧

Exercise 4.33. Show that every convex function is continuous. ✧

Exercise 4.34. Let $f\colon (a,b) \to \mathbb{R}$ be convex and differentiable, and $x_0 \in (a,b)$ such that $f'(x_0) = 0$. Show that f attains its minimum at x_0. ✧

Exercise 4.35. Let $f\colon (a,b) \to \mathbb{R}$ be twice differentiable. Show that f has a local minimum at $x_0 \in (a,b)$, if $f'(x_0) = 0$ and $f''(x_0) > 0$.

Show on the other hand that every local minimum $x_0 \in (a,b)$ satisfies $f'(x_0) = 0$ and $f''(x_0) \geq 0$. ✧

Exercise 4.36. Let $f\colon (a,b) \to \mathbb{R}$ be n times differentiable and for $a < x_1 < x_2 < \cdots < x_n < x_{n+1} < b$ we have $f(x_j) = 0$ (for $j \in \{1,\ldots,n+1\}$). Show that there exists a $\zeta \in (a,b)$ such that $f^{(n)}(\zeta) = 0$. ✧

Exercise 4.37. Show that the sequence $(f_n)_{n\in\mathbb{N}^*}$ given by $f_n\colon \mathbb{R} \to \mathbb{R}$, $f_n(x) := \frac{\sin(nx)}{\sqrt{n}}$, converges uniformly to 0 for $n \to \infty$. Does (f_n') also converge? ✧

Exercise 4.38. Perform a curve sketching (as in Exercise 4.29) for

$$f\colon \begin{cases} \mathbb{R} & \to & \mathbb{R}, \\ x & \mapsto & (x^2-1)\exp(-|x|). \end{cases}$$ ✧

Exercise 4.39. Let $\tan\colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ be defined by $\tan(x) := \frac{\sin(x)}{\cos(x)}$. Perform a curve sketching for tan, including the behavior of tan as $x \to \pm\frac{\pi}{2}$. In particular show that tan is invertible. We denote its inverse as $\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. ✧

Exercise 4.40. Let $\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ be the inverse of tan. Calculate its derivative. ✧

Exercise 4.41. Calculate the derivatives of arcsin and arccos (these are the inverse functions of $\sin\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1,1]$ and $\cos\colon [0,\pi] \to [-1,1]$, respectively). ✧

Exercise 4.42. Calculate the anti-derivatives of ($n \in \mathbb{N}^*$)

(a) $\dfrac{1}{x(x+2)}$ (b) $\dfrac{\exp(3x)+1}{\exp(x)+1}$

(c) $x^2\sin(x)$ (d) $\sin(x)^n\cos(x)$ ✧

Exercise 4.43. Let $a, b, c > 0$. Calculate the anti-derivative of $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) := \frac{a}{b+cx^2}$. ✧

Exercise 4.44. Calculate the anti-derivative of $\frac{x+1}{x(x^4-x)}$. ✧

Chapter 5

The Riemann integral

5.1 The Darboux integral

Motivation:

- We would like to "define" the "area underneath a function".
- We would like to have the "inverse" operation to differentiation

Some intuitive examples

$$f_1(x) = 1 + x^2, \quad f_2(x) = \begin{cases} 1, & \text{if } x \in [0,1] \cap \mathbb{Q}, \\ 0, & \text{if } x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Definition 5.1 (Partition and step function)**.** Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be bounded. Then a finite subset P of $[a,b]$ is called a *partition* of $[a,b]$ if $a, b \in P$.

We make the convention: Let $N+1 \in \mathbb{N}^*$ be the number of elements in P, which we will typically denote, in increasing order, by

$$a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b.$$

Moreover, if we write "$P = \{t_i \mid i = 0, \ldots, N\}$" is a partition, we assume that the t_i's are ordered (and distinct) as above.

We call $f\colon [a,b] \to \mathbb{R}$ a *step function* if there exists a partition $P = \{t_i \mid i = 0, \ldots, N\}$ of $[a,b]$ such that the restriction $f|_{(t_{i-1},t_i)}$ of f to the (open) interval (t_{i-1}, t_i) is constant for every $i = 1, \ldots, N$. We then also say that f is a step function with respect to the partition P. ✧

Note that we don't assume anything about the values a step function takes at the points t_i, $i = 0, \ldots, N$ in the partition.

Definition 5.2 (Lower and Upper sums)**.** Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be bounded. For any partition $P = \{t_i \mid i = 0, \ldots, N\}$ of $[a,b]$, we call

$$L_P(f) := \sum_{i=1}^{N}(t_i - t_{i-1}) \inf_{t\in[t_{i-1},t_i]} f(t), \quad U_P(f) := \sum_{i=1}^{N}(t_i - t_{i-1}) \sup_{t\in[t_{i-1},t_i]} f(t).$$

the *lower sum* and the *upper sum* of f, respectively. ✧

We note that for any bounded function on $[a,b]$ and any partition P of $[a,b]$, the lower and upper sum $L_P(f)$ and $U_P(f)$ exist as real numbers.

Definition 5.3 (Darboux integral)**.** Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be bounded. We define the *lower integral*

$$\underline{\int}_{[a,b]} f = \sup\{L_P(f) \mid P \text{ is partition of } [a,b]\}$$

and the *upper integral*

$$\overline{\int}_{[a,b]} f = \inf\{U_P(f) \mid P \text{ is partition of } [a,b]\}$$

of f. Furthermore, we say that f is *(Darboux/Riemann) integrable* if

$$\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f,$$

and, which case, we call this value *the integral of f from a to b* and denote it by

$$\int_a^b f(x)\,\mathrm{d}x\,.$$ ✧

We remark that the reason for the notation "$\int_a^b f(x)\,\mathrm{d}x$" will only become apparent in due course.

Fact 5.4. *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function.*

(i) *Note that the the lower and upper integral of f are well-defined real numbers because for any partition $P = \{t_i \mid i = 0,\dots,N\}$,*

$$\begin{aligned} -\infty < (b-a) \inf_{t\in[a,b]} f(t) &\le \sum_{i=1}^{N}(t_i - t_{i-1}) \inf_{t\in[t_{i-1},t_i]} f(t) \overset{\text{Def.}}{=} L_P(f) \\ \le U_P(f) = \sum_{i=1}^{N}(t_i - t_{i-1}) &\sup_{t\in[t_{i-1},t_i]} f(t) \le (b-a) \sup_{t\in[a,b]} f(t) < \infty. \end{aligned}$$

Thus also $\underline{\int}_{[a,b]} f \le \overline{\int}_{[a,b]} f$.

(ii) *If f is a step function, then f is integrable with*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{i=1}^{N}(t_i - t_{i-1})c_i = \sum_{i=1}^{N}(t_i - t_{i-1})f\left(\frac{t_{i-1}+t_i}{2}\right),$$

where $P = \{t_i \mid i = 0, \ldots, N\}$ and $f|_{(t_{i-1},t_i)} = c_i \in \mathbb{R}$ for all $i = 1, \ldots, N$.
To see this, observe that $L_P(f) = U_P(f)$ and for this specific P and thus, by the first item, $\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f = \int_a^b f(x)\,\mathrm{d}x$.

(iii) *If P and Q are partitions of $[a,b]$, then $P \cup Q$ is a partition as well. In particular, the integrability as well as the integral of a step function does not depend on the partition chosen to represent it.*

(iv) *f is integrable if and only if for every $\varepsilon > 0$ there exists a partition P of $[a,b]$ such that*

$$U_P(f) - L_P(f) < \varepsilon.$$

(v) *f is integrable if and only if there exists two sequences of step functions $(h_n)_{n\in\mathbb{N}^*}, (g_n)_{n\in\mathbb{N}^*}$ such that*

$$g_n(x) \leq f(x) \leq h_n(x) \quad \forall x \in [a,b],$$
$$\text{and} \quad \lim_{n\to\infty} \int_a^b h_n(x) - g_n(x)\,\mathrm{d}x = 0.$$

This follows directly from the previous two items by choosing $\varepsilon = \frac{1}{n}$, with $P = P_n = \{t_i \mid i = 0, \ldots, N\}$, and

$$g_n(x) = \inf_{t\in[t_{i-1},t_i]} f(t), \quad h_n(x) = \sup_{t\in[t_{i-1},t_i]} f(t).$$

Proposition 5.5. *Let $a < b$ be real numbers and $f, g\colon [a,b] \to \mathbb{R}$ be integrable.*

(i) *For $\lambda \in \mathbb{R}$, $f + \lambda g$ is integrable and $\int_a^b f(x) + \lambda g(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \lambda \int_a^b g(x)\,\mathrm{d}x$.*

(ii) *If $f \leq g$ (that is, $f(x) \leq g(x)$ for all $x \in [a,b]$), then $\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x$.*

(iii) *The functions $f_+, f_-, |f|\colon [a,b] \to \mathbb{R}$ given by $f_+(x) = \max\{f(x), 0\}$, $f_-(x) = \min\{f(x), 0\}$, $x \in [a,b]$ and $|f(\cdot)| = f_+ + f_-$ are integrable and*

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x\,.$$

(iv) *$f \cdot g$ is integrable and if there exists $C > 0$ such that $g(x) \geq C$ for all $x \in [a,b]$, then $\frac{f}{g}$ is integrable.*

Proof. Use Fact 5.4, see lecture. ❑

Example 5.6. Note that functions can be "very discontinuous". The following examples illustrate their (non)integrability.

(a) The function $f\colon [0,1] \to \mathbb{R}$ defined by

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

for $x \in [0,1]$ is called *Dirichlet function* and is not integrable. This follows since $\underline{\int}_0^1 f = 0$ and $\overline{\int}_0^1 = 1$, which, in turn follow from the fact that for any two distinct real numbers $x < y$ (in $(0,1)$), there exists a rational and a irrational number in (x,y). This fact is a consequence of the Archimedean principle, Fact 1.28.

(b) The *Thomae's function* $f\colon [0,1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} \frac{1}{q}, & \text{if } x = \frac{p}{q} \in [0,1], \text{ where } p,q \text{ are coprime}, \\ 0, & \text{if } x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

is integrable. We shall see this shortly. ✧

Recall the notion of uniform convergence for sequences of functions on $[a,b]$ from Section 3.3.

Definition 5.7. For $a < b$ real numbers let

$$\operatorname{Reg}([a,b]) := \left\{ f\colon [a,b] \to \mathbb{R} \;\middle|\; \exists (f_n)_{n\in\mathbb{N}^*} \text{ of step functions such that } \lim_{n\to\infty} \sup_{x\in[a,b]} |f_n(x) - f(x)| = 0 \right\}$$

denote the set of *regulated functions.* ✧

It follows from the definition that elements in $\operatorname{Reg}([a,b])$ are exactly those function which are *uniform limits* of step functions.

Theorem 5.8. *Let $a < b$ be real numbers. Any function in $f \in \operatorname{Reg}([a,b])$ is integrable. In particular, any continuous function $f\colon [a,b] \to \mathbb{R}$ is integrable.*

Proof. First of all we note that any function in $\operatorname{Reg}([a,b])$ is bounded as uniform limit of bounded functions. Given $f \in \operatorname{Reg}([a,b])$, there exists a sequence $(f_n)_{n\in\mathbb{N}^*}$ of step functions which converge uniformly to f. Let $\varepsilon > 0$ and let $n \in \mathbb{N}^*$ be such that

$$\sup_{x\in[a,b]} |f(x) - f_n(x)| < \varepsilon.$$

Let $P = \{t_i \mid i = 0, \dots, N\}$ be a partition corresponding to f_n. We define step functions $g_n, h_n \colon [a, b] \to \mathbb{R}$ by

$$g_n(x) = \inf_{x \in (t_{i-1}, t_i)} f(x),\ h_n(x) = \sup_{x \in (t_{i-1}, t_i)} f(x) \quad x \in (t_{i-1}, t_i), i = 1, \dots, N,$$

and $g_n(x) = h_n(x) = f(x)$ for $x \in P$. By construction, $g_n \le f \le h_n$ and

$$\int_a^b h_n(x) - g_n(x)\,\mathrm{d}x = \int_a^b h_n(x) - f_n(x) + \int_a^b f_n(x) - g_n(x)\,\mathrm{d}x \le 2\int_a^b \varepsilon,$$

where we used that h_n and g_n are constant on each interval of the partition and

$$g_n(x) - f_n(x) \le f(x) - f_n(x) \le |f(x) - f_n(x)| < \varepsilon$$

and similarly,

$$f_n(x) - h_n(x) \le f_n(x) - f(x) \le |f_n(x) - f(x)| < \varepsilon.$$

Therefore, by Fact 5.4 (v), f is integrable.

To see that any continuous function $f \colon [a, b] \to \mathbb{R}$ is in $\mathrm{Reg}([a, b])$, recall that every continuous function on a compact interval is uniformly continuous. Therefore, given $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in [a, b]$

$$|x - y| < \delta \quad \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

Without loss of generality assume that $\delta < b - a$ and let $N = \lfloor \frac{b-a}{\delta} \rfloor$. Set $t_0 = a$, $t_i = t_{i-1} + \delta$ for $i = 1, \dots, N - 1$ and $t_N = b$. Define the step function f_n by assigning the value $f(\frac{t_i + t_{i-1}}{2})$ for each x in any interval (t_{i-1}, t_i), $i = 1, \dots, N$ and $f_n(x) = f(x)$ at all points $x = t_i$, $i = 0, \dots, N$. This shows that f is the uniform limit of $(f_n)_{n \in \mathbb{N}^*}$. ❑

Proposition 5.9. *Let $a < b$ be real numbers and $f \colon [a, b] \to \mathbb{R}$. Then $f \in \mathrm{Reg}([a, b])$ if and only if for all $x \in (a, b)$ the left and the right limit of f at x and $\lim_{t \to a^+} f(t)$ as well $\lim_{t \to b^-} f(t)$ exist.*

Proof. Skipped. ❑

Fact 5.10. *Let $f \colon [a, b] \to \mathbb{R}$ be a function.*

(i) *Integrability of f doesn't change if f is changed at finitely many points.*

(ii) *Let f be bounded and $c \in (a, b)$. Then f is integrable if and only if $f|_{[a,c]}$ and $f|_{[c,b]}$ is integrable. Also note that $f|_{[a,c]}$ is integrable if and only if $f \cdot \mathbb{1}_{[a,c]} \colon [a, b] \to \mathbb{R}$ is integrable, where $\mathbb{1}_{[a,c]} \colon [a, b] \to \mathbb{R}$ is the* indicator function *of the set $[a, c]$, that is, $\mathbb{1}_{[a,c]}(x) = 1$ if $x \in [a, c]$ and $\mathbb{1}_{[a,c]}(x) = 0$ if $x \in [a, b] \setminus [a, c]$. Moreover,*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\,.$$

5.2 The Fundamental Theorem of Calculus

The following result will be extremely useful for explicitly computing integrals.

Theorem 5.11 (Fundamental Theorem of Calculus). *Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be integrable. Then the function $F\colon [a,b] \to \mathbb{R}$ given by*

$$F(x) = \int_a^x f(y)\,\mathrm{d}y, \quad x \in [a,b],$$

is continuous. If f is continuous at $x_0 \in [a,b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof. To show that F is continuous at $x \in [a,b)$ when f is integrable, consider first $h > 0$ small enough such that $x + h \leq b$ and

$$F(x+h) - F(x) = \int_a^{x+h} f(y)\,\mathrm{d}y - \int_a^x f(y)\,\mathrm{d}y = \int_x^{x+h} f(y)\,\mathrm{d}y$$

Taking absolute values and using triangle inequality for integrals, we conclude that

$$|F(x+h) - F(x)| \leq h \sup_{y \in [x,x+h]} |f(y)|,$$

where the supremum is finite because f is bounded as an integrable function. Hence, $\lim_{h\to 0^+} F(x+h) = F(x)$. Analogously, one shows that $\lim_{h\to 0^-} F(x+h) = F(x)$ for any $x \in (a,b]$. Together this implies that $\lim_{h\to 0} F(x+h) = F(x)$, which shows continuity of F.

If f is even continuous at x_0, we have by the above reasoning—again for $h > 0$ and $x_0 + h \leq b$—that

$$\begin{aligned}\left|\frac{1}{h}(F(x+h) - F(x)) - f(x_0)\right| &\leq \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(y) - f(x_0)\,\mathrm{d}y\right| \\ &\leq \sup_{y\in[x_0,x_0+h]} |f(y) - f(x_0)|.\end{aligned}$$

Since f is continuous at x_0, it follows from the definition of continuity that

$$\lim_{h\to 0} \sup_{y\in[x_0,x_0+h]} |f(y) - f(x_0)| = 0,$$

which shows that $\lim_{h\to 0^+} \frac{1}{h}(F(x+h) - F(x)) - f(x_0) = 0$. Again, the case $h < 0$ follows analogously, showing differentiability of F. ❑

Corollary 5.12 (Computing integrals using antiderivatives). *Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be continuous. Then there exists an antiderivative(primitive) $F\colon [a,b] \to \mathbb{R}$ of f on $[a,b]$ given by*

$$F(x) = \int_a^x f(y)\,\mathrm{d}y$$

for all $x \in [a,b]$ and for any antiderivative $G\colon [a,b] \to \mathbb{R}$ of f it holds that

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a).$$

Example 5.13. We compute the integral $\int_1^3 \ln(x)\,\mathrm{d}x$ by using the fundamental theorem of calculus and the primitive we already know for $\ln\colon [1,3] \to \mathbb{R}$. ✧

Corollary 5.14. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous.*

(i) *For $h\colon [\alpha,\beta] \to [a,b]$ continuously differentiable, it holds that*

$$\int_\alpha^\beta f(h(t))h'(t)\,\mathrm{d}h = \int_{h(\alpha)}^{h(\beta)} f(x)\,\mathrm{d}x\,.$$

(ii) *If f and $g\colon [a,b] \to \mathbb{R}$ are continuously differentiable, then*

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x\,.$$

(iii) *If $f(x) \geq 0$ for all $x \in [a,b]$, then*

$$\int_a^b f(x)\,\mathrm{d}x = 0 \qquad \Longrightarrow \qquad f(x) = 0 \text{ for all } x \in [a,b].$$

Proof. The first two statements follow directly from the corresponding statements for primitives, Chapter 4, and the Fundamental Theorem of Calculus Theorem 5.11.

The last statement can be shown directly by supposing that f is not 0 at some point $x \in [a,b]$ and using continuity directly to construct a contradiction. Alternatively, we have by the Fundamental Theorem of Calculus that $F(x) = \int_a^x f(y)\,\mathrm{d}y$ is continuously differentiable on $[a,b]$ with derivative $F'(x) = f(x) \geq 0$. Therefore, F is increasing, Corollary 4.18, but at the same time $F(a) = F(b) = 0$ by assumption and definition of F. Thus, $0 = F(a) \leq F(x) \leq F(b) = 0$ for all $x \in [a,b]$, which settles the proof. ❑

5.3 Improper Riemann Integrals

Definition 5.15 (Improper and absolute integrals)**.** Let I be a nontrivial interval and denote by a and b the left and right boundary points (including $\pm\infty$) and $f\colon [a,b) \to \mathbb{R}(\mathbb{C})$. We say that f is *improperly (Riemann) integrable*, if for every $b' \in (a,b)$, $f|_{[a,b']}$ is integrable and the limit

$$\lim_{b' \to b^-} \int_a^{b'} f(x)\,\mathrm{d}x =: \int_a^{b^-} f(x)\,\mathrm{d}x$$

exists. Analogously, we say that $f\colon (a,b] \to \mathbb{R}(\mathbb{C})$ is improperly integrable if for every $a' \in (a,b)$, $f|_{[a',b]}$ is integrable and $\lim_{a'\to a^+} \int_{a'}^b f(x)\,\mathrm{d}x =: \int_{a^+}^b f(x)\,\mathrm{d}x$ exists. Likewise we say that $f\colon (a,b) \to \mathbb{R}(\mathbb{C})$ is *improperly integrable*, if there exists $c \in (a,b)$ such that $f|_{(a,c]}$ and $f|_{[c,b)}$ is improperly integrable and write

$$\int_{a^+}^{b^-} f(x)\,\mathrm{d}x := \int_{a^+}^{c} f(x)\,\mathrm{d}x + \int_{c}^{b^-} f(x)\,\mathrm{d}x\,. \qquad \diamond$$

Definition 5.16 (Absolutely integrable and absolutely improper integrable)**.** Let I be a nontrivial interval and $f\colon I \to \mathbb{R}(\mathbb{C})$ (or $\mathbb{C}$). If $I = [a,b]$ for some $a,b \in \mathbb{R}$, and $|f(\cdot)|\colon I \to \mathbb{R}(\mathbb{C})$ is integrable, then we say that f is *absolutely integrable*. For other intervals I and $|f(\cdot)|\colon I \to \mathbb{R}(\mathbb{C})$ improperly integrable, we say that f is *absolutely improper integrable.* $\diamond$

Fact 5.17. *We make the following observation about improper integrals.*

(i) *By the Fundamental Theorem of Calculus (first part), Theorem 5.11, it holds that for every integrable function $f\colon [a,b] \to \mathbb{R}$, $f|_{(a,b)}$ is improperly integrable and*

$$\int_a^b f(x)\,\mathrm{d}x = \int_{a^+}^{b^-} f(x)\,\mathrm{d}x = \int_{a^+}^{b} f(x)\,\mathrm{d}x = \int_{a}^{b^-} f(x)\,\mathrm{d}x\,.$$

Therefore, we may also simply denote improper integrals by $\int_a^b f(x)\,\mathrm{d}x$ and state explicitly that it is meant as improper integral.

(ii) *The converse implication, that an improper integrable function f is integrable, is wrong in general. This is, e.g., easily seen with the example $f\colon (0,1] \to \mathbb{R}$ with $f(x) = x^{-\frac{1}{2}}$, which is not defined on $[0,1]$ and unbounded on $(0,1]$ and therefore not integrable (even if we define $f(0)$). However,*

$$\int_\varepsilon^1 x^{-1/2}\,\mathrm{d}x = \frac{1}{2}(1-\varepsilon^{1/2}) \to \frac{1}{2} \;\; \text{as } \varepsilon \to 0^+.$$

(iii) *A function can be absolutely integrable while not being integrable. An instance is given by a variant of the Dirichlet function, Example 5.6,*

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

which is not integrable (as it can be written as two times the Dirichlet function minus the constant 1-function). The function $|f|$ equals the constant 1-function, which is clearly integrable, whence f is absolutely integrable. We shall, however see in the following result that if $f\colon [a,b] \to \mathbb{R}$ is integrable, then it f is also absolutely integrable.

(iv) *In contrast to the last sentence of the previous item, a function may be improperly integrable while not absolutely improperly integrable. This is shown in the following example. Note that this is not so surprising in the light of infinite sums which are convergent, but not absolutely convergent.*

Example 5.18. We consider

$$f\colon [0,\infty) \to \mathbb{R}, x \mapsto \begin{cases} \frac{\sin(\pi x)}{x}, & x > 0, \\ \pi, & x = 0. \end{cases}$$

We show that this function is improperly integrable on $[0,\infty)$, but not absolutely improper integrable. First, observe that f is continuous (and in particular at 0, by the fact that $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$). Therefore, $f|_{[0,1]}$ is integrable and we can focus on the interval $[1,\infty)$ for the remainder of this proof. By the properties of the sine function, the zeros of f are precisely at $x \in \mathbb{N}^*$. The idea is to split the integral (think of its interpretation as area 'underneath' the function) into the parts (that is, integrals over sets) where f is positive and the where f is negative. By the property of the sine that its zeros aren't local extrema, these sets consists of disjoint intervals. To compute the improper integral we further split in the following way.

$$\int_1^{\widetilde{b}} \frac{\sin(\pi x)}{x}\,\mathrm{d}x = \sum_{n=1}^{\lfloor \widetilde{b}\rfloor -1} \int_n^{n+1} f(x)\,\mathrm{d}x + \int_{\lfloor \widetilde{b}\rfloor}^{\widetilde{b}} f(x)\,\mathrm{d}x \tag{5.1}$$

$$= \sum_{n=1}^{\lfloor \widetilde{b}\rfloor -1} (-1)^n \underbrace{\int_n^{n+1} \left|\frac{\sin(\pi x)}{x}\right| \mathrm{d}x}_{=:c_n} + \int_{\lfloor \widetilde{b}\rfloor}^{\widetilde{b}} f(x)\,\mathrm{d}x \tag{5.2}$$

The last identity follows from the property that the sin changes its sign at every zero. The sequence $(c_n)_{n\in\mathbb{N}^*}$ is decreasing and converging to 0 as can be seen by estimating the denominator using the integral limits and the fact that sin is bounded. Therefore, by the Leibniz test, Corollary 2.37, $\sum\limits_{n=1}^{\infty} (-1)^n c_n < \infty$. For the second term in (5.2) we can proceed as follows.

$$\left|\int_{\lfloor \widetilde{b}\rfloor}^{\widetilde{b}} f(x)\,\mathrm{d}x\right| \leq \frac{1}{\lfloor \widetilde{b}\rfloor} \int_{\lfloor \widetilde{b}\rfloor}^{\lfloor \widetilde{b}\rfloor+1} |\sin(\pi x)|\,\mathrm{d}x \leq \frac{1}{\lfloor \widetilde{b}\rfloor}\frac{2}{\pi} \to 0 \quad \text{for} \quad \widetilde{b}\to\infty \tag{5.3}$$

Hence, f is improperly integrable and

$$\int_1^{\infty} \frac{\sin(\pi x)}{x}\,\mathrm{d}x = \sum_{n=1}^{\infty} (-1)^n c_n.$$

If we, instead, consider $|f(\cdot)|$, then the above series can not be dealt with with the Leibniz test. Moreover, by $\frac{1}{x} \geqslant \frac{1}{n+1}$ for $x \in [n, n+1]$, it follows that

$$c_n = \int\limits_n^{n+1} \frac{|\sin(\pi x)|}{x}\,\mathrm{d}x \geq \frac{1}{n+1} \int\limits_n^{n+1} |\sin(\pi x)|\,\mathrm{d}x = \int\limits_0^1 |\sin(\pi x)|\,\mathrm{d}x.$$

Therefore, c_n is bounded below by some positive constant independent of n, whence, by the divergence test, implies that

$$\int_1^{\widetilde{b}} \left|\frac{\sin(\pi x)}{x}\right| \mathrm{d}x = \sum_{n=1}^{\lfloor \widetilde{b} \rfloor - 1} c_n + \int_{\lfloor \widetilde{b} \rfloor}^{\widetilde{b}} |f(x)|\,\mathrm{d}x \to \infty,$$

where we used that the second term converges to 0 by (5.3). ✧

Proposition 5.19. *Let $f, g\colon [a,b) \to \mathbb{R}$ be such that*

$$f|_{[a,x]},\, g|_{[a,x]},\, |f|_{[a,x]}| \text{ and } |g|_{[a,x]}|$$

are integrable for all $x \in (a,b)$. Then the following assertions hold.

(i) *If $|g|$ is improper integrable, then so is g.*

(ii) *If g is absolutely improper integrable and there exists $c \in (a,b)$ such that $|f(x)| \leq |g(x)|$ for all $x \in [c,b)$, then f is absolutely integrable.*

The statements above also hold for the case of intervals of the form $(a,b]$ and (a,b) with the obvious modifications.

Proof. See Exercise 5.7. ❑

5.4 Interchanging Integrals and Sequences of Functions

Theorem 5.20. *Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}(\mathbb{C})$. Let also $f_n\colon [a,b] \to \mathbb{R}(\mathbb{C})$ a sequence of integrable functions which converges uniformly to f, that is, $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$. Then f is integrable and*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x\,.$$

Proof. The proof is very similar in spirit to the proof that functions in $\mathrm{Reg}([a,b])$ are integrable, Theorem 5.8. We restrict ourselves to the case where f is real-valued here. The complex-valued case follows by considering the real and imaginary part of f separately. Let $\varepsilon > 0$ and let $n \in \mathbb{N}^*$ be such that

$$\sup_{x\in[a,b]} |f(x) - f_n(x)| < \varepsilon.$$

Let $P = \{t_i \mid i = 0, \ldots, N\}$ be a partition corresponding to f_n. We define step functions $g_n, h_n\colon [a, b] \to \mathbb{R}$ by

$$g_n(x) = \inf_{x\in(t_{i-1},t_i)} f(x),\ h_n(x) = \sup_{x\in(t_{i-1},t_i)} f(x) \quad x \in (t_{i-1}, t_i), i = 1, \ldots, N,$$

and $g_n(x) = h_n(x) = f(x)$ for $x \in P$. By construction, $g_n \le f \le h_n$ and

$$\begin{aligned}\int_a^b h_n(x) - g_n(x)\,\mathrm{d}x &= \int_a^b h_n(x) - f_n(x)\,\mathrm{d}x + \int_a^b f_n(x) - g_n(x)\,\mathrm{d}x \\ &\le 2\int_a^b \varepsilon\,\mathrm{d}x,\end{aligned}$$

where we used that h_n and g_n are constant on each interval of the partition and

$$g_n(x) - f_n(x) \le f(x) - f_n(x) \le |f(x) - f_n(x)| < \varepsilon$$

and similarly,

$$f_n(x) - h_n(x) \le f_n(x) - f(x) \le |f_n(x) - f(x)| < \varepsilon.$$

Therefore, by Fact 5.4 (v), f is integrable and

$$\left|\int_a^b f(x) - f_n(x)\,\mathrm{d}x\right| \le \int_a^b |f(x) - f_n(x)|\,\mathrm{d}x \le \int_a^b \varepsilon\,\mathrm{d}x = \varepsilon(b - a). \qquad ❏$$

Theorem 5.21. *Let $a < b$ be real numbers and $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of $f_n\colon [a, b] \to \mathbb{R}(\mathbb{C})$, $n \in \mathbb{N}$ continuously differentiable functions, such that*

- *there exists $t_0 \in [a, b]$ such that $\lim_{n\to\infty} f_n(t_0) = y$ exists and*
- *$(f_n')_{n\in\mathbb{N}}$ converges uniformly to $g\colon [a, b] \to \mathbb{R}(\mathbb{C})$.*

Then $(f_n)_{n\in\mathbb{N}}$ converges uniformly to a function $f\colon [a, b] \to \mathbb{R}$, which is continuously differentiable and $f' = g$. In particular, it holds that

$$\lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\lim_{n\to\infty} f_n\Big)(x).$$

Proof. This is consequence of the Fundamental Theorem of Calculus, Theorem 5.11 and Theorem 5.20. First note that g is continuous as uniform limit of continuous functions, see Chapter 3. By Theorem 5.11, $t \mapsto \int_{t_0}^t g(x)\,\mathrm{d}x$ [1] defines a continuously differentiable function f on $[a, b]$ with $f' = g$. Furthermore, by Theorem 5.20 and again Theorem 5.11, we have for every $t \in [a, b]$,

$$f_n(t) = f_n(t_0) + \int_{t_0}^t f_n'(x)\,\mathrm{d}x \overset{n\to\infty}{\longrightarrow} \int_{t_0}^t g(x)\,\mathrm{d}x = f(t).$$

[1] with the convention $\int_\alpha^\beta = -\int_\beta^\alpha$ for $\beta < \alpha$ and $\int_\alpha^\alpha = 0$

Therefore, f_n converges pointwise to f. To show uniform convergence, we use triangle inequality for the integral, the assumption that $f_n(t_0)$ converges to $y = f(t_0)$ as well as the uniform convergence of f_n' to g and Theorem 5.20,

$$\begin{aligned}|f(t) - f_n(t)| &\leq \left| f(t_0) + \int_{t_0}^{t} g(x) - f_n'(x)\,\mathrm{d}x - f_n(t_0) \right| \\ &\leq |f(t_0 - f_n(t_0)| + \left| \int_{t_0}^{t} g(x) - f_n'(x)\,\mathrm{d}x \right| \overset{n\to\infty}{\longrightarrow} 0. \qquad \square\end{aligned}$$

Remark 5.22. We want to discuss several aspects of Theorem 5.21.

(i) In practice, Theorem 5.21 is often applied to the following setting: We are given a sequence of continuously differentiable functions $(f_n)_{n\in\mathbb{N}^*}$ on an interval $[a, b]$ which converges pointwise to a limit function f. In order to conclude that the limit function f is also (continuously) differentiable and that its derivative equals the (pointwise) limit of the derivatives of the sequence, we have (in order to apply Theorem 5.21) to show that the sequence of the derivatives $(f_n')_{n\in\mathbb{N}^*}$ converges uniformly (to some function g).

In the case where we want to interchange limits and derivatives for sequences of functions defined on unbounded intervals or union of intervals, we can instead try to apply the result to the restriction of the functions to general closed subintervals as differentiability (and continuity) is a local property.

(ii) We also note that it is possible to weaken the assumptions of Theorem 5.21 in the sense that the f_n's can be assumed to be only differentiable rather than continously differentiable. However, then we cannot use the same proof technique as above, as we cannot use the Fundamental Theorem of calculus in this more general case.

(iii) As an application of Theorem 5.21, we can conclude (what we in fact already know by other reasoning) that the exponential function, which we defined by the infinite sum

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

(which, as we have seen before, converges uniformly on any compact interval by the Weierstraß test) is differentiable and the derivative equals the exponential function. This follows by the fact that the derivative of the partial sum $S_N(x) = \sum_{n=0}^{N} \frac{x^n}{n!}$ is given by S_{N-1}, and hence the sequence of the derivatives of the partial sum converges uniformly (to $\exp(x)$) on any compact interval. Therefore, by Theorem 5.21, exp is continuously differentiable and thus arbitrarily often differentiable with higher-order derivatives being exp.

(iv) On the other hand we emphasize that in order to conclude that the limit of a convergent sequence of differentiable functions, it is not enough to argue by uniform convergence of the sequence (instead of the sequence of derivatives). For instance consider the sequence of functions given by

$$f_n(x) = \frac{1}{n}\sin(nx), \quad n \in \mathbb{N}^*$$

defined for $x \in [0,1]$. This sequence of continuously differentiable functions converges uniformly to the zero function (since $|\sin(nx)| \leq 1$), but the sequence of the derivatives $f_n'(x) = \cos(nx)$ does not converge pointwise and thus neither uniformly. ✧

5.5 Power series

Note that in the following we index the sequences with $\mathbb{N}_0$ and not $\mathbb{N}^*$.

Definition 5.23. Let $(a_k)_{k\in\mathbb{N}_0}$ be a sequence in $\mathbb{R}$ (or $\mathbb{C}$) and $z_0 \in \mathbb{R}$. Then the expression

$$\sum_{k=0}^{\infty} a_k(z-z_0)^k$$

is called *power series centered at* z_0. The sequence a_k's are called *coefficients* of the power series. For a power series, we define the *radius of convergence*

$$R = \sup\left\{ |z-z_0| \in \mathbb{R} \;\middle|\; z \in \mathbb{R}(\mathbb{C}), \sum_{k=0}^{\infty} a_k(z-z_0)^k \text{ converges} \right\},$$

which can take any value in $[0,\infty]$.[2] ✧

We have already encountered power series, e.g., in the definition of the exponential function as well as the representations of sin and cos. In all those cases, $z_0 = 0$ and the radius of convergence is $+\infty$.

Example 5.24. Consider the sequence $(a_k)_{k\in\mathbb{N}_0}$ given by $a_k = 1$ for all k and $z_0 = 0$. Then the corresponding power series

$$\sum_{k=0}^{\infty} z^k$$

converges, by what we know about geometric series, Chapter 2, for $-1 < x < 1$ (or, in the complex case, for all $x \in \mathbb{C}$ with $|x| < 1$) and diverges for

[2]Note that the supremum defining R always exists in $[0,\infty]$ as the series converges for $z = z_0$.

all other values x. Therefore, the radius of convergence is 1.
If we instead consider the example

$$\sum_{k=0}^{\infty} \frac{1}{k+1} z^k,$$

it is not hard to see by the divergence test and the Weierstrass test, that the series converges for all $x \in (-1,1)$ and diverges for all x with $|x| > 1$ or $x = 1$. However, for $x = -1$, the series converges by the Leibniz test. ✧

Proposition 5.25. *Let $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ be a power series with radius of convergence R. Then*

(i) $\forall z \in \mathbb{R} : |z - z_0| > R \implies \sum_{k=0}^{\infty} a_k(z-z_0)^k$ *diverges*

(ii) $\forall r \in (0, R)$, $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ *converges uniformly and absolutely on* $\{z \in \mathbb{R} : |z - z_0| < r\}$.

(iii) $z \mapsto \sum_{k=0}^{\infty} a_k(z-z_0)^k$ *is continuous on* $(-R + z_0, z_0 + R)$.

(iv) $R \in [0, \infty]$ *and*

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}$$

(with the convention that $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$).

Furthermore, $R = \lim_{k\to\infty} \frac{|a_k|}{|a_{k+1}|}$, if that limit exists in $\mathbb{R} \cup \{\infty\}$.

Proof. The first assertion follows directly from the definition of the radius of convergence.

We explain the proof of the second statement in the case $z_0 = 0$. The general case follows analogously. Let $r < R$ and choose $w \in \mathbb{R}$ with $r < |w| < R$ such that $\sum_{k=0}^{\infty} a_k w^k$ converges, which exists by definition of R. Clearly, the sequence defined by $a_k w^k$ (in the index k) converges (to 0), which in particular implies that the sequence is bounded, that is, there exists $C > 0$ such that $|a_k w^k| < C$ for all $k \in \mathbb{N}_0$. Let $z \in \mathbb{R}$ such that $|z| \le r < |w|$. Then

$$|a_k z^k| = |a_k w^k| \left|\frac{z^k}{w^k}\right| \le C\left|\frac{r}{w}\right|^k =: M_k.$$

Since $|\frac{r}{w}| < 1$, it follows that $\sum_{k=0}^{\infty} M_k < \infty$, whence, by Weierstrass test, Theorem 3.30, $\sum_{k=0}^{\infty} a_k z^k$ converges absolutely and uniformly on $[-r, r]$.

The third assertion follows directly from the second by Theorem 3.24.

To see the fourth assertion, let first $z \in \mathbb{R}$ with $|z| < \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}}$. Then

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1,$$

which, by the root test, Theorem 2.32, implies that $\sum_{n=0}^{\infty} a_k z^k$ converges. Thus

$$\frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}} \le R.$$

Suppose this inequality was strict. Then we could choose $z \in \mathbb{R}$ with

$$\frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}} < |z| < R.$$

By the second assertion of the proposition, we then would have that $\sum_{n=0}^{\infty} a_k z^k$ converges, while

$$\limsup_{k\to\infty} \sqrt[k]{|a_k z^k|} = |x| \limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1.$$

Again, using the root test, we would conclude that $\sum_{n=0}^{\infty} a_k z^k$ diverges, which is a contradiction. ❑

Lemma 5.26. *Let $(a_\ell)_{\ell\in\mathbb{N}^*}$ be a sequence in $\mathbb{R}$. Then*

$$\limsup_{\ell\to\infty} \sqrt[\ell]{|a_\ell|} = \limsup_{\ell\to\infty} \sqrt[\ell]{\ell|a_\ell|},$$

where we allow the value $+\infty$.

Proof. See Exercise 5.36. ❑

Proposition 5.27. *Let $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ be a power series with radius of convergence R. Then the power series*

$$\sum_{k=0}^{\infty} (k+1)a_{k+1}(z-z_0)^k$$

has the same radius of convergence R and the function

$$f\colon (-R+z_0, z_0+R) \to \mathbb{R}, z \mapsto \sum_{k=0}^{\infty} a_k(z-z_0)^k$$

is differentiable with

$$f'(z) = \sum_{k=0}^{\infty} (k+1)a_{k+1}(z-z_0)^k.$$

Moreover, $f \in \mathrm{C}^\infty(-R+z_0, z_0+R)$ with

$$f^{(m)}(z) = \sum_{k=0}^{\infty} \frac{(k+m)!}{k!} a_{k+m}(z-z_0)^k,$$

where the radius of convergence equals R as well. In particular, $f^{(m)}(z_0) = m!\, a_m$.

Proof. We first show that the first assertion (the differentiability), the higher-order differentiability follows by induction.

The statement is an immediate consequence of Theorem 5.21 applied to every interval $[-r+z_0, z_0+r]$, $r \in (0, R)$ to the sequence of polynomials

$$f_n(z) = \sum_{k=0}^{n} a_k(z-z_0)^k, \quad n \in \mathbb{N}^*$$

and Lemma 5.26 (with $\ell = k+1$). In fact, the derivatives f_n' converge uniformly to $f|_{[-r+z_0,z_0+r]}$ on $[-r+z_0, z_0+r]$ by Lemma 5.26 and Proposition 5.25 (iv). ❑

Remark 5.28. We give a brief outlook on the extension of the preceding considerations to complex-valued arguments and on the related notion of analytic functions.

(i) We note that power series can (as indicated above) be defined for complex values z. Indeed, the theory of power series is even more "important" in the latter case, which will be discussed in full detail in the course "Complex analysis" (also called "Complex Function Theory").

(ii) A function $f\colon (a,b) \to \mathbb{R}$ is called *(real-) analytic (on (a,b))* if

$$\forall z_0 \in (a,b)\ \exists (a_k)_{k\in\mathbb{N}_0} \in \mathbb{R}^{\mathbb{N}_0}\ \exists c, d \in \mathbb{R} \text{ with } c < z_0 < d : \\ \forall z \in (c,d) : f(z) = \sum_{k=0}^{\infty} a_k(z-z_0)^k.$$

In other words, f is called analytic if at every point in its domain, the function can be locally represented by a power series. Such a function must be arbitrarily often differentiable and for each $z_0 \in (a,b)$, the coefficients sequence $(a_k)_k$ is uniquely determined, Proposition 5.27.

The corresponding notion of "complex analytic functions" is the corner stone of complex analysis mentioned in the first item of this remark. ✧

Proposition 5.29 (Products of power series)**.** *Let $\sum_{k=0}^{\infty} a_k z^k$ and $\sum_{k=0}^{\infty} b_k z^k$ be two power series converging on some interval $(-r,r)$ with $r > 0$. Then*

$$\sum_{k=0}^{\infty} c_k(z-z_0)^k$$

with

$$c_k = \sum_{n=0}^{k} a_{n-k} b_k, \qquad k \in \mathbb{N}^*,$$

converges and equals $f(x)g(x)$ for all $x \in (-r,r)$.

Proof. Skipped. ❑

5.6 Exercises

Exercise 5.1. Show that $f\colon [0,1] \to \mathbb{R}$ piece-wise given by

$$f(x) := \begin{cases} \frac{1}{2^n}, & \text{for } x \in \left(\frac{1}{2^{n+1}}, \frac{1}{2^n}\right], n \in \mathbb{N}_0, \\ 0, & \text{for } x = 0, \end{cases}$$

is integrable on [0,1], that is, the lower and upper integral coincide. ✧

Exercise 5.2. Let $f\colon (a,b) \to \mathbb{R}$ be continuous. Show that

$$f = 0 \qquad \Longleftrightarrow \qquad \int_c^d f \, \mathrm{d}x = 0 \quad \forall c, d \in (a,b),\ c < d. \qquad ✧$$

Exercise 5.3. Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous, and 1-*periodic*, i.e., $f(x+1) = f(x)$ for all $x \in \mathbb{R}$. Show that f is bounded (on $\mathbb{R}$) and show that for every $\alpha \in \mathbb{R}$

$$\int_0^1 f(x) \, \mathrm{d}x = \int_\alpha^{\alpha+1} f(x) \, \mathrm{d}x \,. \qquad ✧$$

Exercise 5.4. Let $f\colon [-1,1] \to \mathbb{R}$ be given by $f(x) := \sqrt{1-x^2}$. Calculate the anti-derivative of f. Finally, calculate the integral

$$\int_{-1}^1 \sqrt{1-x^2} \, \mathrm{d}x \,.$$

Hint: Use the trigonometric functions and substitution. ✧

Exercise 5.5. Show that the limit

$$\gamma := \lim_{n\to\infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln(n) \right)$$

exists. This number is called *Euler–Mascheroni constant.* ✧

Exercise 5.6. Consider the function $f\colon [0,1] \to \mathbb{R}$ given by

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

(a) Show that f is integrable.

(b) Show that f is not the uniform limit of step functions (that means, $f \notin \mathrm{Reg}([0,1])$). ✧

Exercise 5.7. Let $a < b \leq \infty$ and $f, g\colon [a,b) \to \mathbb{R}$ such that f and g are integrable on every compact subinterval of $[a,b)$.

(a) Show that if g is absolutely improperly integrable, then g is improperly integrable.

(b) Show that if $|f(x)| \leq |g(x)|$ for all $x \in [a, b)$ and g is absolutely improperly integrable, then f is also absolutely improperly integrable. ✧

Exercise 5.8. Let $f, g \colon [a, b] \to \mathbb{R}$ be such that f is continuous, g is integrable and $g \geq 0$. Show that

(a) fg is integrable,

(b) there exists a $\xi \in [a, b]$ such that

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi)\int_a^b g(x)\,\mathrm{d}x\,.$$ ✧

Exercise 5.9. Calculate the anti-derivative of f for f given by:

(a) $f(x) := \dfrac{4}{x^2 + x - 2}$,

(b) $f(x) := x^2 \exp(-x)$,

(c) $f(x) := \dfrac{x - 1}{x^2 + x + 1}$,

(d) $f(x) := \dfrac{\exp(2x) - 1}{\exp(x) + 2}$. ✧

Exercise 5.10. Show that the following improper integral converges absolutely:

$$\int_0^{+\infty} \frac{\ln(x)}{1 + x^2}\,\mathrm{d}x\,.$$

Moreover, show that

$$\int_0^{+\infty} \frac{\ln(x)}{1 + x^2}\,\mathrm{d}x = 0.$$ ✧

Exercise 5.11. Calculate the improper integral

$$\int_{\mathrm{e}}^{\mathrm{e}^2} \frac{1}{x\ln(x)\sqrt{\ln(\ln(x))}}\,\mathrm{d}x\,.$$ ✧

Exercise 5.12. Let $n, m \in \mathbb{Z}$. Show

$$\frac{1}{2\pi}\int_0^{2\pi} \exp(\mathrm{i}nx)\exp(-\mathrm{i}mx)\,\mathrm{d}x = \delta_{n,m},$$

where $\delta_{n,m}$ is the *Kronecker delta*, i.e.,

$$\delta_{n,m} := \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$ ✧

Exercise 5.13. Calculate

$$\int_0^3 \frac{x^2}{\sqrt{1+x}}\,\mathrm{d}x\,. \qquad \diamond$$

Exercise 5.14. Check whether the following improper integrals converge absolutely:

(a) $\displaystyle\int_0^{+\infty} \frac{1+x}{\sqrt{x}(1+x^2)}\,\mathrm{d}x$, (b) $\displaystyle\int_0^1 \frac{x\sqrt{1-x}}{\ln(x)}\,\mathrm{d}x$. $\diamond$

Exercise 5.15. Show

$$\lim_{n\to\infty}\sum_{k=1}^{n} \ln\left(\sqrt[n]{k}\right) - \ln(n) = -1$$

Hint: Use the fundamental theorem of calculus on intervals of the form $[\frac{1}{n}, 1]$. $\diamond$

Exercise 5.16. Use the fundamental theorem of calculus to find an $\ell \in \mathbb{N}^*$ such that

$$\frac{\sum_{k=1}^{n} k^3}{n^\ell}$$

converges to a limit $x \in (0, +\infty)$ as $n \to \infty$.
Conclude that this suggests that there is a polynomial p of ℓ^{th} degree, i.e., $p(x) = \sum_{i=0}^{\ell} a_i x^i$ such that

$$\sum_{k=1}^{n} k^3 = p(n). \tag{5.4}$$

Set up a system of linear equations for $a_0, \dots, a_\ell$ to find this polynomial (you may use software to solve this system). Finally, verify (5.4) by induction. $\diamond$

Exercise 5.17. Let $f\colon (a,b] \to \mathbb{R}$ be such that f is integrable on $[\xi, b]$ for every $\xi \in (a,b]$ and $\lim_{x\to a+} f(x) = +\infty$. Show that

$$\int_a^b f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x,$$

where we extend the lower integral to this situation by

$$\underline{\int_a^b} f(x)\,\mathrm{d}x := \sup\left\{ \sum_{i=1}^{n} (t_i - t_{i-1}) \inf_{x\in(t_{i-1},t_i)} f(x) \,\middle|\, (t_i)_{i=0}^{n} \text{ partition of } [a,b] \right\}.$$

Note the only difference is that we allow functions that are unbounded from above, i.e., that value of $\underline{\int_a^b} f(x)\,\mathrm{d}x$ might be $+\infty$.
Hint: Distinguish the cases: f is improperly integrable and f is not improperly integrable. $\diamond$

Exercise 5.18. Let $f\colon [0,1] \to \mathbb{R}$ be bounded and nonnegative, P a partition of $[0,1]$ and $n \in \mathbb{N}^*$. We define

$$S_n(f) := \sum_{k=1}^{n} \frac{1}{n} f\Big(\frac{k}{n}\Big) \quad \text{and} \quad S_n^P(f) := \sum_{\substack{k=1 \\ P \cap [\frac{k-1}{n}, \frac{k}{n}] = \emptyset}}^{n} \frac{1}{n} f\Big(\frac{k}{n}\Big). \tag{5.5}$$

(a) Show that $\big|S_n(f) - S_n^P(f)\big| \le 2|P| \|f\|_\infty \frac{1}{n}$ (where $|P|$ denotes the number of elements of P).

(b) Show that $S_n^P(f) \le U_P(f)$.

(c) Conclude that for every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}^*$ such that $S_n(f) \le U_P(f) + \varepsilon$ for all $n \ge n_0$. ✧

Exercise 5.19. Let $a, b \in \mathbb{R}$ such that $a < b$. Calculate

$$\lim_{n\to\infty} \int_a^b \sin\big(\tfrac{1}{n^2} x^2 + x\big) \, \mathrm{d}x \,. \qquad ✧$$

Exercise 5.20. Calculate

$$\lim_{n\to\infty} \int_0^1 \frac{2nx^2}{(1+x^2)(nx+1)} \, \mathrm{d}x \,. \qquad ✧$$

Exercise 5.21. Let $f\colon \mathbb{R} \to \mathbb{R}$ be uniformly continuous and $(g_n)_{n\in\mathbb{N}^*}$ be a sequence of functions $g_n\colon (a,b) \to \mathbb{R}$ that converges uniformly to $g\colon (a,b) \to \mathbb{R}$ (where $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$). Show that then also $(f \circ g_n)$ converges uniformly to $f \circ g$. ✧

Exercise 5.22. Calculate the radius of convergence of the following power series

(a) $\displaystyle\sum_{k=0}^{\infty} k^{(\ln k)/k} x^k$ (b) $\displaystyle\sum_{k=0}^{\infty} \sum_{j=1}^{k+1} \frac{1}{j} x^k$ ✧

Exercise 5.23. Let $f\colon [0,+\infty) \to \mathbb{R}$ be a continuous function such that $\lim_{x\to 0+} f(x)/x$ exists. Show that the sequence $(f_n)_{n\in\mathbb{N}^*}$ given by

$$f_n(x) := \frac{nxf(x)}{nx+1}, \qquad x \ge 0,$$

converges uniformly to f on $[0,1]$. Give an additional, sufficient condition such that (f_n) converges uniformly to f on $[0,+\infty)$. ✧

Exercise 5.24. Show that the engineering substitution rule for integrals is mathematically justified.

Hint: Note that the rule is only vaguely defined. Hence, one difficulty is to find a proper formalism for the rule. ✧

Exercise 5.25. Let $f\colon [0,1] \to \mathbb{R}$ be an integrable function. Show

$$\lim_{n\to\infty} S_n(f) = \int_0^1 f(x)\,\mathrm{d}x,$$

where $S_n(f)$ is defined in (5.5).

Hint: Let $\mathfrak{P}$ denote the set of all partitions of $[0,1]$. Use Exercise 5.18 to conclude

$$\forall P \in \mathfrak{P}\ \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ :\ S_n(f) \leq U_P(f) + \varepsilon$$

also for bounded $f\colon [0,1] \to \mathbb{R}$ that are not necessarily nonnegative. Use this to show

$$\forall P \in \mathfrak{P}\ \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ :\ L_P(f) - \varepsilon \leq S_n(f),$$

i.e., there is a short argument that does not need to repeat an analogous version of Exercise 5.18. ✧

Exercise 5.26. Let $n \in \mathbb{N}^*$ and $f_n\colon [0,\infty) \to \mathbb{R}$ be given by $f_n(x) := n\sqrt{x}\exp(-nx^2)$ and $C > \varepsilon > 0$. Calculate

$$\lim_{n\to\infty} \int_\varepsilon^C f_n(x)\,\mathrm{d}x\,.$$

Moreover, calculate

$$\lim_{n\to\infty} \int_0^C f_n(x)\,\mathrm{d}x\,.$$

Hint: For the second limit find a triangle that is below the graph of f_n whose area grows for $n \to \infty$. ✧

Exercise 5.27. Calculate

$$\lim_{n\to\infty} \int_0^1 \frac{1}{\sqrt{nx}} \exp(x^2)\,\mathrm{d}x\,.$$ ✧

Exercise 5.28. Calculate

$$\lim_{n\to\infty} \int_0^\pi \cos\Big(\exp\big(-n(x-\tfrac{1}{n})^2\big)x\Big)\sin(x)\,\mathrm{d}x\,.$$ ✧

Exercise 5.29. Let $n \in \mathbb{N}^*$ and $f_n\colon [0,+\infty) \to \mathbb{R}$ be given by $f_n(x) := nx^2\exp(-x^2 n)$. Calculate

$$\lim_{n\to\infty} \int_0^\infty f_n(x)\,\mathrm{d}x\,.$$

Hint: Argue first for $\int_\varepsilon^\infty f_n(x)\,\mathrm{d}x$ for $\varepsilon > 0$. ✧

Exercise 5.30. Calculate the radius of convergence of the following power series

(a) $\displaystyle\sum_{k=0}^{\infty} \frac{\mathrm{e}^k + \mathrm{e}^{-k}}{2} x^k$ (b) $\displaystyle\sum_{k=0}^{\infty} \binom{2k}{k} (x-4)^k$ ✧

Exercise 5.31 (The integral approach to the logarithm)**.** Let $F\colon (0,+\infty) \to \mathbb{R}$ be given by

$$F(x) := \int_1^x \frac{1}{t}\,\mathrm{d}t\,.$$

Pretend to not know the logarithm. Show (without using the logarithm and its implications) that

(a) $F(xy) = F(x) + F(y)$ for $x, y \in (0,+\infty)$,

(b) $F(x^\alpha) = \alpha F(x)$ for $x \in (0,+\infty)$, $\alpha \in \mathbb{R}$,

(c) $F(\exp(x)) = x$ for $x \in \mathbb{R}$.

Note that you cannot use $\frac{\mathrm{d}}{\mathrm{d}x} x^\alpha = \alpha x^{\alpha-1}$ for general $\alpha \in \mathbb{R}$ as this was derived using the logarithm. In particular for this example we define x^α for $x > 0$ by: Let $(q_n)_{n\in\mathbb{N}}$ be a sequence of rational numbers converging to α, then we say $x^\alpha := \lim_{n\to\infty} x^{q_n}$. (You don't have to check that is a consistent definition.) ✧

Exercise 5.32. Show that

$$\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}, \quad |x| < 1,$$

by writing $1/(1+x^2)$ as a geometric series and by integrating. ✧

Exercise 5.33. Consider the power series $f(x) := \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1}$.

(a) Find the radius of convergence.

(b) Determine all values $x \in \mathbb{R}$ for which the series $f(x)$ converges.

(c) Find a closed form expression (without infinite sum) of $f(x)$.

Hint: You may want to differentiate or integrate f. ✧

Exercise 5.34. Let $f(x) := \sum_{k=0}^{\infty} a_k (x-x_0)^k$ and $g(x) := \sum_{k=0}^{\infty} b_k (x-x_0)^k$ be power series with positive radii of convergence. Show

$$\Big[\exists \varepsilon > 0\ \forall x \in \mathrm{B}_\varepsilon(x_0) \,:\, f(x) = g(x)\Big] \quad \Longleftrightarrow \quad \Big[\forall k \in \mathbb{N}_0 \,:\, a_k = b_k\Big].$$

Recall that $\mathrm{B}_r(x_0) := \big\{x \in \mathbb{R} \bigm| |x - x_0| < r\big\}$.

Hint: Consider $f^{(n)}(x_0)$ and $g^{(n)}(x_0)$. ✧

Exercise 5.35. Consider the following power series:

$$S(x) := \sum_{k=2}^{\infty} kx^{k-2}$$

(a) Find the radius of convergence and the largest set $D \subseteq \mathbb{R}$ where the series converges.

(b) Find a closed form expression for $S\colon D \to \mathbb{R}$. ✧

Exercise 5.36. Let $(a_k)_{k\in\mathbb{N}^*}$ be a sequence in $\mathbb{R}$. Show that

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = \limsup_{k\to\infty} \sqrt[k]{k|a_k|},$$

where we allow the value $+\infty$.
Hint: Use that $\lim_{k\to\infty} \sqrt[k]{k} = 1$ and the characterization that the limit superior is the largest partial limit of a sequence, see Theorem 2.25. ✧

Exercise 5.37. Let $f(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k$ be a power series with radius of convergence $R > 0$ and $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{B}_R(x_0) \setminus \{x_0\}$ that converges to x_0. Show that

$$\Big[\forall n \in \mathbb{N} \,:\, f(x_n) = 0\Big] \quad \Longrightarrow \quad \Big[\forall x \in \mathrm{B}_R(x_0) \,:\, f(x) = 0\Big].$$

Conclude that two power series coincide, if they coincide on a set that contains a limit point (accumulation point) in $\mathrm{B}_R(x_0)$. ✧

Chapter 6

Differentiability of functions in several variables

6.1 Preliminaries

We collect(recap) several facts about sequences in $(\mathbb{R}^p, d_2)$ and more general normed spaces $(X, \|\cdot\|)$, which we will need in the rest of this chapter.

Definition 6.1. Let X be a vector space (over $\mathbb{R}$ or $\mathbb{C}$) and let $\|\cdot\|\colon X \to [0, \infty)$ have following properties for all $f, g \in X, c \in \mathbb{R}$:

(i) $\|f\| = 0 \iff f = \mathbf{0}$ (positive definiteness),

(ii) $\|cf\| = |c|\|f\|$ (absolute homogeneity),

(iii) $\|f + g\| \leq \|f\| + \|g\|$ (triangle inequality). ✧

We call $\|\cdot\|$ a norm on X and $(X, \|\cdot\|)$ a *normed space*. Furthermore, $d\colon X \times X \to [0, \infty), d(x, y) = \|x - y\|$, $x, y \in X$, defines a metric on X, the so-called *induced metric* of a norm. Therefore, we will, without explicit mentioning, consider a normed space as metric space with that metric.

The most important normed space of this chapter is given by $(\mathbb{R}^p, \|\cdot\|_2)$, where

$$\|x\|_2 := \left(\sum_{i=1}^{p} |x_i|^2\right)^{\frac{1}{2}} \quad \text{for} \quad x \in \mathbb{R}^p.$$

In fact, it has to be shown that $\|\cdot\|_2$ indeed defines a norm.

Proposition 6.2 (Cauchy–Schwarz and Minkowski)**.** *Then for $x, y \in \mathbb{R}^p$ (or $\mathbb{C}^p$),*

- $|\langle x, y\rangle| = \left|\sum_{k=1}^{p} x_i y_i\right| \leq \|x\|_2 \|y\|_2$ *(Cauchy–Schwarz inequality)*
- $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$ *(Minkowski inequality)*

More generally, both inequalities hold for normed spaces arising from inner product spaces $(X, \langle\cdot,\cdot\rangle)$ *where* $\|x\| = \sqrt{\langle x, x\rangle}$ *for* $x \in X$.

Proof. The Cauchy–Schwarz inequality was shown for more general inner-product spaces (over the field $\mathbb{R}$ or $\mathbb{C}$). The Minkowski inequality is a simple consequence from Cauchy–Schwarz:

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle \\ \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|.\end{aligned}$$

Using the binomial formula on the right-hand-side and taking roots on both sides finishes the proof. ❏

Fact 6.3. *We collect some facts about convergence and continuity in* $\mathbb{R}^p$.

(i) $(\mathbb{R}^p, d_2)$ *is a normed space and a sequence* $(x_n)_{n\in\mathbb{N}^*}$ *in* $\mathbb{R}^p$ *(a sequence of vectors/points) converges to a limit* $x \in \mathbb{R}^p$ *if and only if each component sequence converges to the corresponding component of* x, *that is*

$$\forall i \in \{1, \ldots, p\} \ : \ \lim_{n\to\infty} x_{n,i} = x_i,$$

where for each i *the limit is understood in the metric space* $(\mathbb{R}, d_2)$ *(that is,* $|x_{n,i} - x_i| \to 0$ *as* $n \to \infty$*). We also say that the sequence* (x_n) *converges* component-wise *to* x. *This statement is a direct consequence of the inequalities*

$$\forall y \in \mathbb{R}^p \ : \ \max_{i=1,\ldots,p} |y_i| \leq \|y\|_2 \leq \sqrt{p} \max_{i=1,\ldots,p} |y_i|.$$

(ii) *The previous point particularly shows that a function* $f\colon D \subseteq \mathbb{R}^p \to \mathbb{R}^m$ *is continuous (with respect to the metric* d_2 *for both spaces) if and only if for every sequence* (x_n) *in* D *which converges component-wise to some limit* x *in* D, *the sequence* $(f(x_n))$ *converges component-wise to* $f(x)$.

(iii) *Let* (X, d) *be a metric space. A function* $f\colon X \to \mathbb{R}^m$ *is continuous (with respect to the standard metric on* $\mathbb{R}^m$*) if and only if the components* $f_i\colon X \to \mathbb{R}$, $i = 1, \ldots, m$, *of* f *are continuous, where, more precisely,* $f_i = \pi_i \circ f$, *with* $\pi_i\colon \mathbb{R}^m \to \mathbb{R}, x \mapsto x_i$.

Example 6.4. We give several example for continuous functions of multiple variables.

(a) The functions $\mathrm{add}_2\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + y$ and $\mathrm{mult}_2\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto xy$ are continuous and hence, by Proposition 3.8, also $\mathrm{add}_p\colon \mathbb{R}^p \to \mathbb{R}, x \mapsto \sum_{k=1}^p x_k$, $\mathrm{mult}_p\colon \mathbb{R}^p \to \mathbb{R}$, $x \mapsto \prod_{k=1}^p x_k$.

(b) $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^3 \cdot \mathrm{e}^y + \tan(xy)$ is continuous as composition of continuous functions, including the ones from the first example.

(c) The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

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

is continuous at every point $(x,y) \neq (0,0)$ as composition of continuous functions. The point $(x,y) = (0,0)$ has to be considered separately. Note that $f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2}$ for all $n \in \mathbb{N}^*$, whence f cannot be continuous at $(0,0)$.

(d) Note that if we instead of the function in (c) consider the function

$$g(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0), \end{cases}$$

then the function is continuous at $(0,0)$ since

$$|g(x,y)| \leq |y| \frac{|xy|}{x^2+y^2} \leq |y| \frac{1}{2} \to 0$$

as $(x,y) \to 0$. See the tutorials for several examples of this form.

(e) The function $f\colon \mathbb{R}^3 \to \mathbb{R}^3, (x,y,z) \to (x^2, xyz, z^3)$ is continuous since every component function is continuous as composition of continuous functions.

(f) Let (X, d_X) and (Y, d_Y) be metric spaces and let $f\colon X \to Y$. Assume that there exists a constant $C > 0$ such that for all $x, y \in X$ it holds that

$$d_Y(f(x), f(y)) \leq C d_X(x,y).$$

Then the function is continuous which follows directly from the definition of continuity. We say that f is *Lipschitz (continuous)*. In the case of normed spaces X and Y, this condition rewrites to

$$\exists C > 0 \, \forall x, y \in X \; : \; \|f(x) - f(y)\| \leq C\|x - y\|. \qquad ✧$$

We define an important class of continuous functions from $\mathbb{R}^p$ to $\mathbb{R}^m$.

Definition 6.5. Let X and Y be normed spaces over the same field ($\mathbb{R}$ or $\mathbb{C}$). A linear mapping $T\colon X \to Y$ is called *bounded* if

$$\exists C > 0 \, \forall x \in X \; : \; \|T(x)\|_Y \leq C\|x\|_X.$$

The set of bounded linear operators from X to Y is denoted by $\mathcal{L}(X,Y)$. The smallest constant C for which the above inequality holds is called the *operator norm* of T. ✧

Proposition 6.6. *Let X and Y be normed spaces and $T\colon X \to Y$ linear. Then the following assertions are equivalent.*

(i) *T is bounded;*

(ii) *T is continuous;*

(iii) *T is continuous at $\mathbf{0} \in X$.*

Furthermore, every linear mapping $T\colon \mathbb{R}^p \to Y$ is bounded and thus continuous, where $\mathbb{R}^p$ is considered with the standard norm $\|\cdot\|_2$.

Proof. If T is bounded, it is easy to see by linearity that T is Lipschitz, see Example 6.4 (f), and thus continuous. Trivially this implies that T is continuous at $\mathbf{0}$. Let us now assume that T is not bounded. Therefore, there exists a sequence $(x_n)_{n\in\mathbb{N}^*}$ in $X \setminus \{\mathbf{0}\}$ such that

$$\|T(x_n)\| > n\|x_n\| \qquad \Longleftrightarrow \qquad \|T(y_n)\| > 1$$

for all $n \in \mathbb{N}^*$, where $y_n := \frac{1}{n\|x_n\|}x_n$. Here, we used the properties of the norm and linearity of T. Now $\lim_{n\to\infty} y_n = \mathbf{0}_X$, while $T(y_n)$ cannot converge to $\mathbf{0}_Y = T(\mathbf{0}_X)$.

To show the second statement, let $(\mathrm{e}_i)_{i=1}^p$ denote the canonical basis of $\mathbb{R}^p$. Then write $x = \sum_{i=1}^p x_i \mathrm{e}_i$ for $x \in \mathbb{R}^p$. By using linearity, triangle inequality (multiple times) and Cauchy–Schwarz, we derive

$$\|T(x)\| = \left\|\left(\sum_{i=1}^p x_i T(\mathrm{e}_i)\right)\right\| \leq \sum_{i=1}^p |x_i|\|T(\mathrm{e}_i)\| \leq C\|x\|_2,$$

where $C = \left(\sum_{i=1}^p \|T(\mathrm{e}_i)\|^2\right)^{\frac{1}{2}}$, which shows that T is bounded. ❑

Remark 6.7. Later (in Analysis 3), we will show that in the above statement $(\mathbb{R}^p, \|\cdot\|)$ can be replaced by any finite-dimensional normed space $(X, \|\cdot\|_X)$. In other words, any linear mapping $T\colon X \to Y$, for a finite-dimensional normed space X and a general normed space Y is bounded and thus continuous. ✧

6.2 Directional derivatives and differentiable functions

In the following functions f will be typically defined on an open subset D of $\mathbb{R}^p$ or, more generally of some normed space X, and map to $\mathbb{R}^m$, or, more generally to some normed space Y. Note that the vector space structure (besides the metric) is crucial here to define the upcoming concepts.

We first note that it is not hard to generalize the notion of a derivative for functions $f\colon D \subseteq \mathbb{R} \to \mathbb{R}^m$, by considering the component functions, and noting what has been said in the first section of this chapter. In the following one may freely choose $X = \mathbb{R}^p$ and $Y = \mathbb{R}^m$, or even for $p, m \in \mathbb{N}^*$ or even $m = 1$ to get acquainted with the following notions.

First recall the following notion generalizing bounded, open intervals

Definition 6.8. Let (X, d) be metric space. A set $M \subseteq X$ is called *open (in the metric space* (X, d)*)* if for every $x \in M$ there exists $r > 0$ such that $\mathrm{B}_r(x) := \{y \in X \mid d(x, y) < r\} \subseteq M$. ✧

Definition 6.9. Let X and Y be normed spaces, $D \subseteq X$ be an open set and $x \in D$. Let $f\colon D \to Y$ be a function. For $v \in X \setminus \{\mathbf{0}\}$, we say that f has a *derivative in direction* v at x, if there exists an $L \in Y$ such that

$$\frac{1}{h}\big(f(x + hv) - f(x)\big) \to L \text{ in } Y \text{ as } h \to 0.$$

In other words, $L = \lim_{h\to 0} \frac{1}{h}\big(f(x+hv)-f(x)\big)$ exists as limit in the normed space Y. We call

$$\frac{\partial}{\partial v} f(x) := L = \lim_{h\to 0} \frac{1}{h}\big(f(x + hv) - f(x)\big)$$

the *directional derivative* in direction v at x. We consequently say that

- f is *Gateaux*[1] *differentiable* (or *directional differentiable*) at x if f has a directional derivative in the direction of v for all $v \in X \setminus \{\mathbf{0}\}$;

In the case that $X = \mathbb{R}^p$ (with the standard norm/metric), we say that

- *f is partially differentiable at x* if the directional derivatives in the direction $v = \mathrm{e}_i$ exist for all $i = 1, \ldots, p$. Here e_i denotes the i-th standard basis vector of $\mathbb{R}^p$. When denoting the components of x by $x_1, \ldots, x_p$. We use the common notation

  $$\frac{\partial f}{\partial x_i}(x) = \frac{\partial}{\partial x_i} f(x) = \frac{\partial}{\partial \mathrm{e}_i} f(x) \in Y, \qquad i = 1, \ldots, p$$

  and call these the *partial derivatives* of f at x.
  If additionally $Y = \mathbb{R}^m$, $m \in \mathbb{N}^*$, the matrix

  $$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_p}(x) \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_p}(x) \end{pmatrix}$$

  is called the *Jacobi matrix* or *Jacobian* of f at x. If even $m = 1$, the vector

  $$(\operatorname{grad} f)(x) = \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right)^{\mathsf{T}} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_p}(x) \end{pmatrix} \in \mathbb{R}^p$$

[1] René Eugène Gateaux 1889-1914 a french mathematician, killed on the battle field during the first world war at the age of 25.

is called the *gradient* of f at x.

We say that f is *Gateaux differentiable* (or *directional differentiable* or *partially differentiable* if f is Gateaux differentiable or partially differentiable at every point $x \in D$, respectively. ✧

Remark 6.10. We remark that the notion of *Gateaux differentiable* is not consistently used in the literature. ✧

Fact 6.11. *In the general setting of the above definition, we make the following observations.*

(i) *Since D is open, there exists $r > 0$ such that $\mathrm{B}_r(x) \subseteq D$. Since for every $v \in X \setminus \{\mathbf{0}\}$, $\lim_{t\to 0} tv = \mathbf{0}$ (the zero element of X, as a limit in X), there exists $t_0 > 0$ such that $x + tv \in \mathrm{B}_r(x) \subseteq D$ for all $t \in (-t_0, t_0)$. Therefore, the directional derivative $\frac{\partial f}{\partial v}(x)$ exists if and only if the function*

$$f_v\colon (-t_0, t_0) \to Y, t \mapsto f(x + tv)$$

is differentiable at $t = 0$ in the analogous sense as defined in Chapter 4, that is,

$$\lim_{h\to 0} \frac{1}{h} \left(f_v(0 + h) - f_v(0)\right)$$

exists as a limit in Y.

(ii) *For $X = \mathbb{R}^p$ and $Y = \mathbb{R}^m$ equipped with the standard norms (metrics), the partial derivative $\frac{\partial f}{\partial x_i}(x)$ of f, for some $i = 1, \ldots, p$ exists if and only if the function $y \mapsto f(x_1, \ldots, y, \ldots, x_p)$ is differentiable at $y = x_i$. In other words, we "freeze" all other variables of the function and consider whether the function in the remaining variable is differentiable.*

By Fact 6.3 (and analogously to continuity), it moreover suffices to consider the component functions $f_k\colon \mathbb{R}^p \to \mathbb{R}$, $k = 1, \ldots, m$ separately, when investigating whether f is differentiable.

(iii) *If $X = \mathbb{R} = Y$ with the standard metric, then the the notion of directional derivative (in any direction) coincides with the derivative we have already seen in Chapter 4.*

Example 6.12. We give a few examples for partially differentiable functions.

(a) Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x\sin(y)$. Then the partial derivatives of f exist at any point (x, y) in $\mathbb{R}^2$ because $x \mapsto x\sin(y)$ as well as $y \mapsto x\sin(y)$ are differentiable for fixed $y \in \mathbb{R}$ and $x \in \mathbb{R}$ respectively.

(b) For the function f from Example 6.4 (c), the partial derivatives exist for every point $(x, y) \neq (0, 0)$ by the above fact since at those points the functions $f(\cdot, y)$ and $f(x, \cdot)$ are differentiable as composition of differentiable functions. As before, the point $(x, y) = (0, 0)$ has to be considered separately. These exist and equal 0 since

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{1}{h}(f(0+h,0) - f(0,0)) = \lim_{h\to 0}\frac{1}{h}(0-0) = 0,$$

and

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0}\frac{1}{h}(f(0,0+h) - f(0,0)) = \lim_{h\to 0}\frac{1}{h}(0-0) = 0.$$

Note however that for $v = (1,1)^\mathsf{T}$, the directional derivative

$$\frac{\partial f}{\partial v}(0,0) = \lim_{h\to 0}\frac{1}{h}(f(h,h) - f(0,0)) = \lim_{h\to 0}\frac{1}{h}\frac{h^2}{2h^2} = \infty$$

does not exist.

(c) For the function g from Example 6.4 (d), the directional derivatives at $(0,0)$ in any direction $v = (v_1, v_2)^\mathsf{T} \in \mathbb{R}^2 \setminus \{(0,0)\}$ indeed exist and equal

$$\frac{\partial f}{\partial v}(0,0) = \frac{v_1 v_2^2}{v_1^2 + v_2^2}.$$

(d) Looking at the previous two examples one may wonder whether continuity implies the existence of all directional derivatives (hence Gateaux differentiability). As we have already seen in Chapter 4, continuity does not imply differentiability for functions $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$. However—and in contrast to what we have seen in Chapter 4 for functions $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$—the existence of all directional derivatives at some point does not imply continuity at that point. A counterexample is given by $h\colon \mathbb{R}^2 \to \mathbb{R}$,

$$h(x,y) = \begin{cases} \dfrac{xy^3}{x^2+y^6}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0), \end{cases}$$

which is discontinuous at $(0,0)$ but its directional derivatives exist at $(0,0)$, in particular $\frac{\partial f}{\partial v}(0,0) = 0$ for all $v \in \mathbb{R}^2 \setminus \{(0,0)\}$. ✧

Definition 6.13 (Continuous partial derivatives)**.** Let $D \subseteq \mathbb{R}^p$ open and $f\colon D \to \mathbb{R}^m$ be partially differentiable. If all partial derivatives

$$\frac{\partial f}{\partial x_i}\colon D \to \mathbb{R}^m, x \mapsto \frac{\partial f}{\partial x_i}(x), \qquad i = 1,\dots,p,$$

are continuous (as functions from (D, d_2) to $(\mathbb{R}^m, d_2)$) (at some $x \in D$), we say that f has *continuous partial derivatives* (at $x \in D$). ✧

Example 6.14. The following examples demonstrate properties that arise specifically in multivariable differentiation.

(a) The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ (see Exercise 6.1)

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

is partially differentiable on $\mathbb{R}^2$ as composition of differentiable functions (in one variable) with

$$\frac{\partial f}{\partial x}(x,y) = \frac{2xy^4}{(x^2+y^2)^2}, \quad \frac{\partial f}{\partial y}(x,y) = \frac{2x^4y}{(x^2+y^2)^2}$$

for $(x,y) \neq (0,0)$ and since for $(x,y) = (0,0)$

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0.$$

In order to check for continuous partial derivatives, we have to check if the functions

$$\frac{\partial f}{\partial x} = \begin{cases} \dfrac{2xy^4}{(x^2+y^2)^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0), \end{cases}$$

and

$$\frac{\partial f}{\partial y} = \begin{cases} \dfrac{2x^4y}{(x^2+y^2)^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0), \end{cases}$$

are continuous. This is clearly true at any point different from $(0,0)$ as composition of continuous functions. To discuss $(0,0)$, we note that by $2xy \leq x^2+y^2$, we have for all $(x,y) \neq (0,0)$,

$$\left| \frac{2xy^4}{(x^2+y^2)^2} \right| \leq \frac{|y^3|}{x^2+y^2} \leq |y|,$$

where we used that $x^2+y^2 \geq y^2$ in the last step (in the denominator). This shows that $\frac{\partial f}{\partial x}$ is continuous and the argument for $\frac{\partial f}{\partial y}$ follows analogously.

(b) The function g from Example 6.4 (d), see also Example 6.12 (c) is partially differentiable. However, it can be shown that that the partial derivatives are not both continuous. Recall that in this example, as mentioned in Example 6.12 (c), all directional derivatives exist. ✧

Definition 6.15 ((Total/Fréchet) differential)**.** Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces (over $\mathbb{R}$), $D \subseteq X$ open and $x \in D$. A function $f\colon D \to Y$ is called (totally/Fréchet) differentiable, if there exist

$$r > 0, \text{a function } \phi\colon \mathrm{B}_r(\mathbf{0}) \to Y \text{ and a bdd. linear mapping } A \in \mathcal{L}(X,Y)$$

such that

- $\mathrm{B}_r(x) \subseteq D$ and $\lim_{h\to\mathbf{0},h\in X} \frac{\|\phi(h)\|_Y}{\|h\|_X} = 0$, and
- $\forall h \in \mathrm{B}_r(\mathbf{0}) \,:\, f(x+h) = f(x) + A(h) + \phi(h)$ or equivalently $\forall y \in \mathrm{B}_r(x) \,:\, f(y) = f(x) + A(y-x) + \phi(y-x)$

In this case we call $\mathrm{d}f(x) \coloneqq A$ the *differential of f at x*. If f is differentiable at every point $x \in D$, we say that f is *differentiable.* ✧

Proposition 6.16 (Properties of $\mathrm{d}f$)**.** *Under the general assumptions of Definition 6.15 let $f\colon D \to Y$ be differentiable at $x \in D$ with $\mathrm{d}f(x) = A \in \mathcal{L}(X,Y)$. Then the following assertions hold.*

(i) *f is Gateaux differentiable at $x \in D$ and*

$$\forall v \in X \setminus \{\mathbf{0}\} \,:\, \frac{\partial f}{\partial v}(x) = A(v).$$

Therefore, $\mathrm{d}f(x)$ is well-defined.

(ii) *f is continuous at x.*

(iii) *If $X = \mathbb{R}^p$, then the linear mapping $\mathrm{d}f(x)$ is uniquely determined by the partial derivatives of f and $\frac{\partial f}{\partial x_i}(x) = A(e_j)$ for all $j = 1, \ldots, p$.*
If additionally, $Y = \mathbb{R}^m$, then $\mathrm{d}f(x)$ can be represented by the Jacobian $\frac{\mathrm{d}f}{\mathrm{d}x}(x) \in \mathbb{R}^{m\times p}$, i.e.,

$$\mathrm{d}f(x)(h) = \sum_{i=1}^{m} \left(\left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix} \right)_i e_i, \qquad \forall h \in \mathbb{R}^p,$$

and

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \begin{pmatrix} ((\mathrm{d}f(x)(e_1))_1 & ((\mathrm{d}f(x)(e_2))_1 & \ldots & ((\mathrm{d}f(x)(e_p))_1 \\ \vdots & & & \\ ((\mathrm{d}f(x)(e_1))_m & ((\mathrm{d}f(x)(e_2))_m & \ldots & ((\mathrm{d}f(x)(e_p))_m \end{pmatrix}$$

and where $(h_1, \ldots, h_p)^\mathsf{T}$ is the coefficient vector of h with respect to the canonical basis $(e_i)_{i=1}^m$ of $\mathbb{R}^m$, that is, $h = \sum_{i=1}^m h_i e_i$. With some abuse of notation (by identifying vectors with their coefficient vectors with respect to the canonical basis), the former identity can be rewritten as

$$\mathrm{d}f(x)(h) = \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right) \cdot h.$$

(iv) *Let also $g\colon D \to Y$ be differentiable at x. For $\lambda \in \mathbb{R}$ the function $\lambda f + g$ is differentiable at x and*

$$\mathrm{d}(\lambda f + g)(x) = \lambda \mathrm{d}f(x) + \mathrm{d}g(x).$$

If f and g are continuously differentiable (see definition below), so is $\lambda f + g$.

Proof. To show the first assertion, let $v \in X \setminus \{\mathbf{0}\}$ and consider the mapping $f_v \colon (-r, r) \to Y$, $f(t) = f(x + tv)$, $t \in (-t_0, t_0)$ for $r > 0$ from the definition of the (total) derivative. Then $tv \in \mathrm{B}_r(\mathbf{0})$ and

$$\frac{1}{t}\left(f_v(t) - f_v(0)\right) = \frac{1}{t}\left(f(x + tv) - f(x)\right) = \frac{1}{t}\left(A(tv) + \phi(tv)\right) = A(v) + \frac{\phi(tv)}{t}$$

Since v is fixed, the last term converges to $A(v) \in Y$ as $tv \to \mathbf{0}$ (in X) by the property of ϕ. Thus, $\frac{\partial f}{\partial v}(x) = A(v)$. Hence, A is uniquely determined by the directional derivatives (as they all exist), which implies that $\mathrm{d}f(x) = A$ is well-defined (not other B can exist such that B is the derivative of f).

To see (ii), consider

$$\|f(y) - f(x)\|_Y = \|A(y - x) + \phi(y - x)\|_Y \leq \|A(y - x)\|_Y + \|\phi(y - x)\|_Y$$

for $y \in \mathrm{B}_r(x)$. Note that as $\lim_{h \to \mathbf{0}} \frac{\|\phi(h)\|}{\|h\|_X} = 0$, also $\lim_{x \to y \phi(y-x)} = 0$. Furthermore, since A is linear and bounded, thus continuous, by Proposition 6.6, we have that $\|A(x - y)\| = \|A(x) - A(y)\|_Y \to 0$ for $\|y - x\|_X \to 0$ in X. Thus f is continuous.

The third assertion is trivial from the first, by plugging in the canonical basis vectors for v and noting that a linear mapping on a finite-dimensional space is represented by its action on basis vectors.

The linearity of the differential, item (iv) is left as an exercise. ❑

Fact 6.17. *Let us collect the following observations.*

(i) *If a function $f \colon D \to Y$ is differentiable at a point x, the mapping*

$$X \mapsto Y, v \mapsto \begin{cases} \frac{\partial f}{\partial v}(x) & v \neq \mathbf{0}_X \\ \mathbf{0}_Y & v = \mathbf{0}_X \end{cases}$$

is linear and bounded. This is a direct consequence of Proposition 6.16 since the mapping coincides with $\mathrm{d}f(x)$.

(ii) *For normed spaces $(X, \|\cdot\|_X)$ (over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$) and $(Y, \|\cdot\|_Y)$, the set $\mathcal{L}(X, Y)$ defines a vector space (over $\mathbb{F}$) which is normed by the* operator norm

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{x \in X \setminus \mathbf{0}} \frac{\|T(x)\|_Y}{\|x\|_X}, \quad T \in \mathcal{L}(X, Y).$$

(it is a simple exercise to show that $\|\cdot\|_{\mathcal{L}(X,Y)}$ indeed defines a norm).

(iii) *For $X = \mathbb{R}$ (and $Y = \mathbb{R}$) (with the standard norms), the notions of "partial differentiable, Gateaux differentiable" and "(totally) differentiable)" all coincide. In the general case, we emphasize the difference in the mathematical objects: While a partial or directional derivative of a function at some point is an element of Y, the total derivative at that point is a linear, continuous mapping. This "distinction" will again be prominently present when considering "higher-order derivatives" soon.*

By Fact 6.17, $\mathcal{L}(X, Y)$ is a normed space and thus has an induced metric. Therefore, for a differentiable function $f\colon D \to Y$, we can ask whether $x \mapsto \mathrm{d}f(x)$ is continuous.

Definition 6.18. Let X, Y be normed spaces (over $\mathbb{R}$), $D \subseteq X$, $x \in D$. A differentiable function $f\colon D \to Y$ is called *continuously differentiable* at $x \in D$ if the function

$$\mathrm{d}f\colon D \to \mathcal{L}(X, Y), x \mapsto \mathrm{d}f(x)$$

is continuous at x, where $\mathcal{L}(X, Y)$ is equipped with the operator norm

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{x \in X \setminus \mathbf{0}} \frac{\|T(x)\|_Y}{\|x\|_X}, \quad T \in \mathcal{L}(X, Y).$$

We say that f is continuous differentiable, if it is continuously differentiable at every $x \in D$. ✧

The following result is very useful in order to check differentiability by investigating partial derivatives only.

Theorem 6.19. *Let $D \subseteq \mathbb{R}^p$ open, and $f\colon D \to \mathbb{R}^m$. Then f is continuously partially differentiable if and only if f is continuously differentiable.*

Proof. We write down most of the proof for the general case, where $f\colon D \to Y$ and specify $Y = \mathbb{R}^m$ at some point.

$\boxed{\Rightarrow}$: We first assume that f has continuous partial derivatives and show that f is differentiable. Let $x \in D$. We define, for $v = (v_1, \ldots, v_p)^\mathsf{T} \in \mathbb{R}^p$,

$$A(v) = \sum_{i=1}^{p} v_i A\mathrm{e}_i = \sum_{i=1}^{p} v_i \frac{\partial f}{\partial x_i}(x) \in Y, \tag{6.1}$$

which clearly defines a linear mapping from $\mathbb{R}^p$ to Y, and hence $A \in \mathcal{L}(\mathbb{R}^p, Y)$, by Proposition 6.6. Furthermore, set

$$\phi\colon \mathrm{B}_r(\mathbf{0}) \to Y, h \mapsto f(x+h) - f(x) - A(h), \quad h \in \mathrm{B}_r(\mathbf{0}) \setminus \{\mathbf{0}\}$$

where r is chosen such that $\mathrm{B}_r(x) \subseteq D$. We aim to show that $\lim_{h \to \mathbf{0}} \frac{\|\phi(h)\|_Y}{\|h\|_2} = 0$. For $j \in \{1, \ldots, p\}$ define $\tilde{x}_j \coloneqq x + \sum_{i=1}^{j} h_i \mathrm{e}_i \in \mathbb{R}^p$. Then, by a telescopic sum, and (6.1), we have

$$\phi(h) = \sum_{j=1}^{p} \left(f(\tilde{x}_j) - f(\tilde{x}_{j-1}) - h_j \frac{\partial f}{\partial x_j}(x) \right),$$

which, by triangle inequality, yields

$$\frac{\|\phi(h)\|_Y}{\|h\|_2} \leq \frac{1}{\|h\|_2} \sum_{j=1}^{p} \left\| f(\tilde{x}_j) - f(\tilde{x}_{j-1}) - h_j \frac{\partial f}{\partial x_j}(x) \right\|_Y \tag{6.2}$$

Because of $\tilde{x}_j = \tilde{x}_{j-1} + h_j \mathrm{e}_j$, we can rewrite the terms in the sum using the Fundamental Theorem of Calculus applied to the continuously differentiable functions

$$[-r/2, r/2] \to Y, s \mapsto f(\tilde{x}_j + s\mathrm{e}_j),$$

so that

$$f(\tilde{x}_{j-1} + h_j \mathrm{e}_j) - f(\tilde{x}_{j-1}) - h_j \frac{\partial f}{\partial x_j}(x) = \int_0^{h_j} \frac{\partial f}{\partial x_j}(\tilde{x}_j + s\mathrm{e}_j) - \frac{\partial f}{\partial x_j}(x)\,\mathrm{d}s\,.$$

In fact, we have only proved the Fundamental Theorem of Calculus Theorem 5.11 for scalar-valued functions in Chapter 4. However, it is easy to see that this readily generalizes to functions $f\colon [a,b] \to \mathbb{R}^m$ (or, even more general, $f\colon [a,b] \to Y$). Therefore, using triangle inequality of the integral,

$$\begin{aligned}\left\| f(\tilde{x}_{j-1} + h_j \mathbf{e}_j) - f(\tilde{x}_{j-1}) - h_j \frac{\partial f}{\partial x_j}(x) \right\| & \\ \leq \int_0^{h_j} \left\| \frac{\partial f}{\partial x_j}(\tilde{x}_j + s\mathrm{e}_j) - \frac{\partial f}{\partial x_j}(x) \right\| \mathrm{d}s \leq |h_j| M_{h_j},\end{aligned}$$

where $M_{h_j} = \sup_{s\in[0,h_j]} \left\| \frac{\partial f}{\partial x_j}(\tilde{x}_j + s\mathrm{e}_j) - \frac{\partial f}{\partial x_j}(x) \right\|_Y$. Note that $\lim_{h_j \to 0} M_{h_j} = 0$ for every $j = 1, \ldots, p$ since the partial derivatives are continuous. Therefore, with (6.2),

$$\frac{\|\phi(h)\|_Y}{\|h\|_2} \leq \sum_{j=1}^{p} \frac{|h_j|}{\|h\|_2} M_{h_j} \leq \sum_{j=1}^{p} M_{h_j} \to 0 \qquad \text{for } \|h\|_2 \to 0.$$

This shows that f is differentiable at x for every $x \in D$. To see that $x \mapsto \mathrm{d}f(x)$ is continuous, let $v \in X \setminus \{\mathbf{0}\}$, $x, y \in D$ and consider

$$\begin{aligned}\|\mathrm{d}f(x)(v) - \mathrm{d}f(y)(v)\| &= \left\| \sum_{j=1}^{p} v_i \left(\frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y) \right) \right\| \\ &\leq \|v\|_2 \left(\sum_{j=1}^{p} \left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y) \right\|_Y^2 \right)^{\frac{1}{2}},\end{aligned}$$

where we used (6.1), triangle inequality for normed spaces and Cauchy–Schwarz. Since f has continuous partial derivatives, each term in the sum goes to zero and thus the sum. This shows that

$$\|\mathrm{d}f(x) - \mathrm{d}f(y)\|_{\mathcal{L}(X,Y)} \stackrel{\text{def}}{=} \sup_{v \in X\setminus\{\mathbf{0}\}} \frac{\|\mathrm{d}f(x)(v) - \mathrm{d}f(y)(v)\|}{\|v\|_2} \to 0 \text{ as } y \to x,$$

which shows that $\mathrm{d}f$ is continuous.

$\boxed{\Leftarrow}$: The converse direction is left as an exercise. ❑

Theorem 6.20 (Chain rule)**.** *Let X, Y, Z be normed spaces, $D_f \subseteq X, D_g \subseteq Y$ open and $f\colon D_f \subseteq X \to Y$, $g\colon D_g \subseteq Y \to Z$ und $x \in D_f$ such that*

- *$f(D_f) \subseteq D_g$,*
- *f is differentiable at x and*
- *g is differentiable at $f(x)$.*

Then $g \circ f$ is differentiable at x and

$$\mathrm{d}(g \circ f)(x) = \mathrm{d}g(f(x)) \circ (\mathrm{d}f(x)).$$

Proof. By definition of differentiability, there exists balls $\mathrm{B}_r(x) \subseteq X$ and $\mathrm{B}_{r'}(f(x)) \subseteq Y$ and error functions

$$\phi_f(h) = f(x+h) - f(x) - \mathrm{d}f(x)(h) \tag{6.3}$$

$$\phi_g(\tilde{h}) = g(f(x)+\tilde{h}) - g(f(x)) - \mathrm{d}g(f(x))(\tilde{h}) \tag{6.4}$$

defined for $h \in \mathrm{B}_r(\mathbf{0}_X) \subseteq X$ and $\tilde{h} \in \mathrm{B}_{r'}(\mathbf{0}_Y) \subseteq Y$ respectively. Since for $h \to \mathbf{0}_X$ (in X), we have that $\mathrm{d}f(x)(h) + \phi_f(h) \to \mathbf{0}_Y$ in Y, it follows that (set $\tilde{h} = \mathrm{d}f(x)(h) + \phi_f(h)$ in the second identity above)

$$\begin{aligned} g\big(f(x+h)\big) - g\big(f(x)\big) & \\ &\overset{(6.3)}{=} g\Big(f(x) + \underbrace{\mathrm{d}f(x)(h) + \phi_f(h)}_{=\tilde{h}}\Big) - g\big(f(x)\big) \\ &\overset{(6.4)}{=} \mathrm{d}g(f(x))\big(\mathrm{d}f(x)(h) + \phi_f(h)\big) + \phi_g\big(\mathrm{d}f(x)(h) + \phi_f(h)\big) \\ &= \Big(\mathrm{d}g\big(f(x)\big) \circ \mathrm{d}f(x)\Big)(h) + \phi_{g\circ f}(h), \end{aligned}$$

where

$$\phi_{g\circ f}(h) := \mathrm{d}g(f(x))(\phi_f(h)) + \phi_g(\mathrm{d}f(x)(h) + \phi_f(h)).$$

It remains to show that $\frac{\|\phi_{g\circ f}(h)\|}{\|h\|_2} \to 0$ as $h \to \mathbf{0}_X$. For the first term, we have by the definition of the operator norm,

$$\|\mathrm{d}g(f(x))(\phi_f(h))\| \leq \|\mathrm{d}g(f(x))\|_{\mathcal{L}(X,Y)} \|\phi_f(h)\|_X.$$

For the second, we note that

$$\|\phi_g(\mathrm{d}f(x)(h) + \phi_f(h))\|_Z = \frac{\|\phi_g(\mathrm{d}f(x)(h) + \phi_f(h))\|_Z}{\|\mathrm{d}f(x)(h) + \phi_f(h)\|_Y} \|\mathrm{d}f(x)(h) + \phi_f(h)\|_Y$$

and that $\|\mathrm{d}f(x)(h) + \phi_f(h)\|_Y \leq \|\mathrm{d}f(x)\|_{\mathcal{L}(X,Y)} \|h\|_X + \|\phi_f(h)\|_Y$. Altogether, using the defining property of ϕ_f and ϕ_g, this shows that $\frac{\|\phi_{g\circ f}(h)\|}{\|h\|_X} \to 0$ as $h \to \mathbf{0}_X$. ❑

6.3 Higher derivatives and Taylor's theorem

Note that as stated in Fact 6.17 (ii), the set $\mathcal{L}(X,Y)$ is a vector space (with addition and scalar multiplication of linear mappings) which is even normed with the operator norm. This thus allows for the following definition

Definition 6.21. Let X,Y be normed spaces (over $\mathbb{R}$ or $\mathbb{C}$), $D \subseteq X$ open, $x \in D$ and $f\colon D \to Y$. We say that

- f is *twice differentiable* (or *two-times differentiable*) at x if f is differentiable on D and
$$\mathrm{d}f\colon D \to \mathcal{L}(X,Y), x \mapsto \mathrm{d}f(x)$$
is differentiable at x. The (total) derivative of $\mathrm{d}f$ at x is called *second(-order) derivative of f at x* and denoted by
$$\mathrm{d}^2 f(x) \coloneqq \mathrm{d}(\mathrm{d}f)(x) \in \mathcal{L}(X, \mathcal{L}(X,Y)).$$
- f is *twice differentiable* if f is twice differentiable at every $x \in D$.
- f is *twice continuously differentiable* (at $x \in D$). if f is twice differentiable and $\mathrm{d}^2 f\colon D \to \mathcal{L}(X,\mathcal{L}(X,Y))$ is continuous (at $x \in D$). The twice continuously differentiable functions from D to Y are also denoted by $\mathrm{C}^2(D,Y)$.

Analogously, we define $(k+1)$-th order derivatives, $k \in \mathbb{N}^*$.

- f is $(k+1)$*-times differentiable* at x if f is k-differentiable on D and
$$\mathrm{d}^k f\colon D \to \mathcal{L}(X,Y), x \mapsto \mathrm{d}^k f(x)$$
is differentiable at x. The derivative of $\mathrm{d}^k f$ at x is called $(k+1)$*-th order derivative of f at x* and denoted by $\mathrm{d}^{k+1} f(x) = \mathrm{d}(\mathrm{d}^k f)(x) \in \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, \dots)))$.
- f is $(k+1)$*-times differentiable* if f is $(k+1)$-times differentiable at every $x \in D$.
- f is $(k+1)$*-times continuously differentiable* (at $x \in D$). if f is $(k+1)$-times differentiable and $\mathrm{d}^{k+1} f\colon D \to \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, \dots)))$ is continuous (at $x \in D$). The $(k+1)$-times continuously differentiable functions from D to Y are also denoted by $\mathrm{C}^{k+1}(D,Y)$.

Likewise, we can define *higher-order partial derivatives*: We start by saying that partial derivatives we have already encountered are called first-order partial derivatives. Then for $k \in \mathbb{N}^*$,

- f is called $(k+1)$*-times partially differentiable* at $x \in D$ if f is k-times partially differentiable (on D) and the partial k-th order partial derivative
$$D \to Y, y \mapsto \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(y)$$

is partially differentiable at x for every $i_1, i_2, \ldots, i_k \in \{1, \ldots, p\}$. These partial derivatives are denoted by

$$\frac{\partial^{k+1} f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k} \partial x_{i_{k+1}}}(x).$$

- In the case that $Y = \mathbb{R}$, we also define the *Hesse matrix* or *Hessian* of f at x by

$$\operatorname{Hess}(f)(x) = \left(\tfrac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1,\ldots,p} \in \mathbb{R}^{p \times p} \qquad \diamondsuit$$

Fact 6.22. *We discuss some conceptual and notational aspects of higher derivatives.*

(i) *We emphasize that for f twice differentiable at $x \in D$, the total second-order derivative $\mathrm{d}^2 f(x)$ is a linear mapping from X to $\mathcal{L}(X, \mathcal{L}(X, Y))$, in particular, for every $v \in X$, $\mathrm{d}^2 f(x)(v)$ is a bounded linear mapping from X to Y, i.e., $\mathrm{d}^2 f(x)(v) \in \mathcal{L}(X, Y)$. Analogous properties hold for higher-order derivatives.*

(ii) *As seen in Linear Structures (Algebra) for finite-dimensional X and $Y = \mathbb{R}$, one can identify the vector space $\mathcal{L}(X, \mathcal{L}(X, Y))$ with the the bilinear forms from $X \times X$ to Y. More generally, this identification can be done for $\underbrace{\mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, \ldots)))}_{k\text{-times}}$ with the multilinear forms from $\underbrace{X \times X \times \cdots \times X}_{k\text{-times}}$. This also justifies the more convenient notation*

$$\left(\cdots \left((((\mathrm{d}^k f(x))(v_1))(v_2))(v_3) \right) \cdots \right)(v_k) = \mathrm{d}^k f(x)(v_1, v_2, \ldots, v_k)$$

which we will use in the following. In the case where $v_1 = v_2 = \cdots = v_k = v$, we may also simply

$$\mathrm{d}^k f(x)(v_1, v_2, \ldots, v_k) = \mathrm{d}^k f(x) v^k$$

(iii) *By Proposition 6.16, applied iteratively, and using that $\frac{\partial}{\partial v} \mathrm{d} f(x)(\tilde{v}) = \frac{\partial}{\partial v} \frac{\partial}{\partial \tilde{v}} f$, one deduces that if $f \colon D \to Y$ is k-times differentiable at $x \in D$ (and $X = \mathbb{R}^p$), then f is k-times Gateaux differentiable (k-times partially differentiable) at x and*

$$\mathrm{d}^k f(x)(v_1, v_2, \ldots, v_k) = \frac{\partial}{\partial v_1} \frac{\partial}{\partial v_2} \frac{\partial}{\partial v_3} \cdots \frac{\partial f}{\partial v_k}(x) \in Y,$$

for $v_1, v_2, v_3, \ldots, v_k \in X \setminus \{\mathbf{0}_X\}$.

Theorem 6.23 (Schwarz). *Let X, Y be normed spaces, $D \subseteq X$ open and $x \in D$. If $f\colon D \subseteq X \to Y$ is k-times totally differentiable at x, then $\mathrm{d}^k f(x)$ is symmetric, i.e.,*

$$\mathrm{d}^k f(x)(v_1, v_2, v_3, \ldots, v_k) = \mathrm{d}^k f(x)(v_{\pi(1)}, v_{\pi(2)}, v_{\pi(3)}, \ldots, v_{\pi(k)}),$$

for any bijective mapping $\pi\colon \{1, \ldots, k\} \to \{1, \ldots, k\}$. In other words, for totally differentiable functions (for $X = \mathbb{R}^p$), the order of the directional derivatives (the partial derivatives) is irrelevant.

Proof. See Theorem A.3. ❑

Corollary 6.24. *Let $D \subseteq \mathbb{R}^p$ open and $f\colon D \to \mathbb{R}^m$ twice continuously partially differentiable at $x \in D$. Then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x), \qquad i, j = 1, \ldots, p.$$

Proof. In the lecture, we provided a separate proof of this special case of Theorem 6.23, which is shorter than the one of the general statement, but requires the stronger assumption of continuous differentiability. ❑

To further ease notation, it is convenient to use *multi-index* notation for partial derivatives whenever their order does not matter: If $X = \mathbb{R}^p$ f is k-times totally differentiable (or, even k-times continuously partially differentiable) we may write for the k-th order partial derivative (if it exits) and the p-tuple $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_p) \in \mathbb{N}_0^p$ with $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_p = k$,

$$\mathrm{d}^\alpha f(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \frac{\partial^{\alpha_3}}{\partial x_3^{\alpha_3}} \cdots \frac{\partial^{\alpha_k} f}{\partial x_p^{\alpha_p}}(x),$$

where we have the convention that $\frac{\partial^0}{\partial x_j^0} f(x) = f(x)$.

Note that by Schwarz' theorem, we have that the Hesse matrix of a twice continuously differentiable function is symmetric.

Theorem 6.25 (Taylor). *Let X be a normed space, $k \in \mathbb{N}^*$, $D \subseteq X$ open and $f\colon D \to \mathbb{R}$. If $x \in D$ and $h \in X$ such that $x + th \in D$ for all $t \in [0, 1]$ and f is k-times continuously differentiable at x, then there exists $\xi \in (0, 1)$ such that*

$$f(x+h) = \sum_{j=0}^{k-1} \frac{1}{j!} \mathrm{d}^j f(x) \underbrace{(h, \ldots, h)}_{j\text{-times}} + \frac{1}{k!} \mathrm{d}^k f(x + \xi h) \underbrace{(h, \ldots, h)}_{k\text{-times}}$$

and

$$\lim_{h \to \mathbf{0}_X} \frac{1}{\|h\|_X^k} \left| f(x+h) - \sum_{j=0}^{k} \frac{1}{j!} \mathrm{d}^j f(x) \underbrace{(h, \ldots, h)}_{j\text{-times}} \right| = 0$$

with $\mathrm{d}^0 f(x) = f(x)$.

6.4 Extreme values

Recall that we have already defined, Definition 4.13 local extrema for functions $f\colon (M,d) \to \mathbb{R}$, where (M,d) is a general metric space. Here, we will study the relation of differentiability and (local) extrema for functions defined on subsets D of a normed space X.

Note that while D is not a normed space, D with (the restriction to D of) the metric induced by the norm of X is a metric space. In complete analogy to the situation in Chapter 5 for functions in one variable, we have the following result.

Proposition 6.26. *Let X be a normed space, $D \subseteq X$ and let $f\colon D \to \mathbb{R}$ be Gateaux differentiable (partially differentiable) at $x \in D$. If f has a local extremum at $x \in D$, that is, either a local maximum, i.e.,*

$$\exists \varepsilon > 0 \ \forall y \in \mathrm{B}_r(x) \ : \ f(x) \geq f(y),$$

or a local minimum, i.e.,

$$\exists \varepsilon > 0 \ \forall y \in \mathrm{B}_r(x) \ : \ f(x) \leq f(y),$$

then all directional derivatives(all partial derivatives) equal 0.

If f is even differentiable (and has a local extremum at $x \in D$), then the total derivative at x equals $\mathbf{0}_{\mathcal{L}(X,\mathbb{R})}$.

Proof. Self-Exercise ❑

Definition 6.27. Let X be a normed space and $D \subseteq X$ open, $f\colon D \to \mathbb{R}$ k-times continuously differentiable and $x \in D$. We say that the total derivative $\mathrm{d}^k f(x)$ is

- *positive definite*, if $\exists \kappa > 0 \ \forall h \in X \ : \ \mathrm{d}^k f(x)(\underbrace{h, \dots, h}_{k\text{-times}}) \geq \kappa \|h\|^k$,
- *negative definite*, if $\exists \kappa > 0 \ \forall h \in X \ : \ \mathrm{d}^k f(x)(\underbrace{h, \dots, h}_{k\text{-times}}) \leq \kappa \|h\|^k$,
- *positive semidefinite*, if $\forall h \in X \ : \ \mathrm{d}^k f(x)(\underbrace{h, \dots, h}_{k\text{-times}}) \geq 0$,
- *negative semidefinite*, if $\forall h \in X \ : \ \mathrm{d}^k f(x)(\underbrace{h, \dots, h}_{k\text{-times}}) \geq 0$,
- *indefinite*, if $\mathrm{d}^k f(x)$ is neither negative semidefinite nor positive semidefinite. ✧

We note that for $X = \mathbb{R}$, the differential $\mathrm{d}^k f(x)$ is $\begin{matrix}\text{positive}\\ \text{negative}\end{matrix}$ definite if and only if $f^{(k)}(x) \gtrless 0$.

Theorem 6.28 (Sufficient condition for local extrema)**.** *Let $D \subseteq X$ open and $f\colon D \to \mathbb{R}$ k-times continuously differentiable, $k \in \mathbb{N}$, $x \in D$ and*

$$\mathrm{d}f(x) = \mathbf{0}, \mathrm{d}^2 f(x) = \mathbf{0}, \dots, \mathrm{d}^{(k-1)} f(x) = \mathbf{0} \quad \textit{and} \quad \mathrm{d}^k f(x) \neq \mathbf{0}.$$

Then

(i) *If k is even and $\mathrm{d}^k f(x)$ is $\begin{smallmatrix}\textit{positive}\\ \textit{negative}\end{smallmatrix}$ definite, then $f(x)$ is a local $\begin{smallmatrix}\textit{minimum}\\ \textit{maximum}\end{smallmatrix}$ f at x.*

(ii) *If k is odd or $\mathrm{d}^k f(x)$ is indefinite, then x is not an extremum.*

(iii) *If f has a $\begin{smallmatrix}\textit{minimum}\\ \textit{maximum}\end{smallmatrix}$ at x, then k is even and $\mathrm{d}^k f(x)$ is $\begin{smallmatrix}\textit{positive}\\ \textit{negative}\end{smallmatrix}$ semidefinite.*

Proof. We prove (i) and refer to the literature for the other assertions. Suppose $k = 2m$ for some $m \in \mathbb{N}^*$ and that $\mathrm{d}^k f(x)$ is positive definite. Let $\delta > 0$ such that $\mathrm{B}_\delta(x) \subseteq D$. Then by Theorem 6.25, there exists $\xi \in (0, \delta)$ such that for every $h \in \mathrm{B}_\delta(\mathbf{0})$

$$f(x+h) - f(x) = \frac{1}{k!}\mathrm{d}^k f(x) h^k + R_{k,x,f}(h),$$

with $R_{k,x,f} = \frac{1}{k!}\left(\mathrm{d}^k f(x+\xi h)h^k - \mathrm{d}^k f(x)h^k\right)$. Furthermore, by Theorem 6.25

$$|R_{k,x,f}(h)| \leq \frac{\kappa}{2k!}\|h\|_X^k$$

for all $h \in \mathrm{B}_{\delta'}(\mathbf{0})$, where $\delta' \in (0, \delta)$ is chosen sufficiently small. Therefore, by assumption that $\mathrm{d}^k f(x)$ positive definite,

$$f(x+h) - f(x) \geq \kappa\|h\|_X^k - \frac{\kappa}{2}\|h\|_X^k = \frac{\kappa}{2}\|h\|_X^k.$$

Therefore, f has a local minimum at x, which is even strict since the last term is (strictly) positive.

The case that $\mathrm{d}^k f(x)$ negative definite follows readily from what has been shown by considering $-f$ instead of f. ❑

The following result helps to check the conditions posed in the previous results regarding definiteness in the case that $X = \mathbb{R}^p$ and $k = 2$.

Proposition 6.29 (Definiteness of Hessian)**.** *Let $X = \mathbb{R}^p$ and $k = 2$ in the situation of 6.28. Then:*

$$\mathrm{d}^2 f(x) \textit{ is } \begin{matrix}\textit{positive}\\ \textit{negative}\\ \textit{positive semi-}\\ \textit{negative semi-}\end{matrix} \textit{ definite} \iff \operatorname{Hess} f(x) \textit{ is } \begin{matrix}\textit{positive}\\ \textit{negative}\\ \textit{positive semi-}\\ \textit{negative semi-}\end{matrix} \textit{ definite.}$$

Therefore, $\mathrm{d}^2 f(x)$ is indefinite if and only if $\operatorname{Hess} f(x)$ is indefinite.

Proof. It remains to show that $\operatorname{Hess} f(\mathbf{x}) > 0$ in the sense of bilinear forms, implies that there exists $\kappa > 0$ such that

$$h^\mathsf{T} \operatorname{Hess} f(x) h \geq \kappa \|h\|^2 \tag{6.5}$$

for all $h \in \mathbb{R}^p$. To see this note first that the set $K = \{h \in \mathbb{R}^p : \|h\|_2 = 1\}$ is compact by Bolzane–Weierstraß. Further note that the mapping

$$g \colon h \mapsto h^\mathsf{T} \operatorname{Hess} f(x) h$$

is continuous from $\mathbb{R}^p$ to $\mathbb{R}$. Therefore g attains a minimum κ on the set K. This minimum must be greater than 0 since, by assumption $g(h) > 0$ for all $h \neq \mathbf{0}$. Therefore, $g(h) \geq \kappa$ for all $h \in K$ and (6.5) follows for general h by dividing h by its norm (if nonzero). ❑

Reminder on Linear Algebra. The symmetric bilinear form represented by a matrix $B \in \mathbb{R}^{p \times p}$ (with respect to the canonical basis)

$$\text{is} \quad \begin{matrix} \text{positive} \\ \text{negative} \\ \text{positive semi-} \\ \text{negative semi-} \end{matrix} \quad \text{definite, if } h^\mathsf{T} \cdot B \cdot h \begin{matrix} > \\ < \\ \geq \\ \leq \end{matrix} 0 \ \forall h \in \mathbb{R}^p \setminus \{\mathbf{0}\}.$$

This property is equivalent to stating that the eigenvalues of (the linear mapping given by) B are $\begin{matrix} \text{positive} \\ \text{negative} \\ \text{greater or equal } 0 \\ \text{less or equal } 0 \end{matrix}$.

Corollary 6.30. *Let $D \subseteq \mathbb{R}^p$ be open and $f \colon D \to \mathbb{R}$ twice continuously differentiable and $x \in D$.*

(i) *Then* $\operatorname{grad} f(x) = \mathbf{0}$*, if f has a local extremum at x.*

(ii) *If* $\operatorname{grad} f(x) = \mathbf{0}$ *and* $\operatorname{Hess} f(x)$ *is* $\begin{matrix} \textit{positive} \\ \textit{negative} \end{matrix}$ *definite, then f has a strict local* $\begin{matrix} \textit{Minimum} \\ \textit{Maximum} \end{matrix}$ *at x.*

6.5 Exercises

Exercise 6.1. Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be given by

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

(a) Determine all points in $\mathbb{R}^2$ at which the function is continuous.

(b) Show that the partial derivatives $\frac{\partial}{\partial x} f(x,y)$ and $\frac{\partial}{\partial y} f(x,y)$ exist at every point $(x,y) \in \mathbb{R}^2$.

(c) Show that the partial derivatives $\frac{\partial}{\partial x} f$, $\frac{\partial}{\partial y} f$ are continuous as functions from $\mathbb{R}^2 \to \mathbb{R}$. ✧

Exercise 6.2. Let $g\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$g(x,y) := \begin{cases} \dfrac{xy^2}{x^2+y^4}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Determine all points in $\mathbb{R}^2$ at which the function is continuous.

(b) Argue that all directional derivatives exist for all $(x,y) \neq (0,0)$.

(c) Calculate the directional derivatives in $(0,0)$. ✧

Exercise 6.3. Let f and g be defined as in Exercise 6.1 and Exercise 6.2. Find the Jacobian of the function $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$h(x,y) := (f(x,y), g(x,y)), \quad (x,y) \in \mathbb{R}^2,$$

and the directional derivative $\frac{\partial h}{\partial v}(0,0)$ for $v := (1,1)$. ✧

Exercise 6.4. Let $A \in \mathbb{R}^{d\times d}$ be a square matrix such that $\|A\| < 1$ show that

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k.$$ ✧

Exercise 6.5. Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) := \begin{cases} x^2 y \sin\left(\dfrac{1}{x^2+y^2}\right), & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Prove that f is continuous.

(b) Compute the partial derivatives $\frac{\partial f}{\partial x}(x,y)$ and $\frac{\partial f}{\partial y}(x,y)$.

(c) Is f differentiable at $(0,0)$? ✧

Exercise 6.6. Let $I(x) := \int_{-\exp(x)}^{x^2} \cos(xt^2)\,\mathrm{d}t$ for $x \in \mathbb{R}$. Calculate $I'(x)$

Hint: You can regard the function $J(x,y,z) := \int_y^z \cos(xt^2)\,\mathrm{d}t$ and write I as a composition of J and other mappings. ✧

Exercise 6.7. Let $A \in \mathbb{R}^{d\times d}$ and $x \in \mathbb{R}^d$. We define $f\colon \mathbb{R}^d \to \mathbb{R}^d$ by $f(x) := Ax$ and $g\colon \mathbb{R}^d \to \mathbb{R}$ by $g(x) := x^\top Ax$. Calculate their derivatives $\mathrm{d}f$ and $\mathrm{d}g$. ✧

Exercise 6.8. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) := \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Show that f is continuous.

(b) Show that f is partially differentiable and find the gradient $\operatorname{grad} f(x,y)$ for all $(x,y) \in \mathbb{R}^2$.

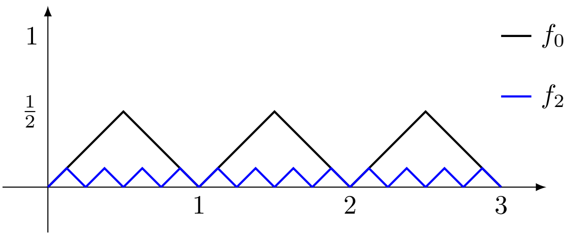
(c) Show that the directional derivative $\frac{\partial}{\partial v} f(x,y)$ exists for every $(x,y) \in \mathbb{R}^2$ and $v \in \mathbb{R}^2 \setminus \{(0,0)\}$.

(d) Show that f is not (totally) differentiable in $(0,0)$. ✧

Exercise 6.9. Let $f_0\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f_0(x) := \begin{cases} x - \lfloor x \rfloor, & \text{if } x - \lfloor x \rfloor \in [0, \frac{1}{2}), \\ 1 - (x - \lfloor x \rfloor), & \text{if } x - \lfloor x \rfloor \in [\frac{1}{2}, 1). \end{cases}$$

Here $\lfloor \cdot \rfloor$ denotes the floor function.



Also define $f_k\colon \mathbb{R} \to \mathbb{R}$ by $f_k(x) := f_0(2^k x)/2^k$ for $x \in \mathbb{R}$ and $k \in \mathbb{N}^*$.

(a) Show that f_0 is periodic with period 1.

(b) Show that $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) := \sum_{k=0}^{\infty} f_k(x), \quad x \in \mathbb{R},$$

is continuous.

(c) Let $k \in \mathbb{N}^*$. Show that for every $x \in \mathbb{R}$ with $2^{k+1}x \notin \mathbb{Z}$ it holds that $f_k'(x) \in \{-1, 1\}$.

(d) Let $x \in [0,1)$. For every $n \in \mathbb{N}^*$ let $p = p_{x,n} \in \mathbb{Z}$ be such that $x \in [\frac{p}{2^n}, \frac{p+1}{2^n}] =: [a_n, b_n]$. Show that for every $k \in \mathbb{N}^*$ and $n > k$ the value

$$\frac{f_k(b_n) - f_k(a_n)}{b_n - a_n}$$

only depends on k (and not on n) and is either -1 or $+1$. Denote this value by c_k.

(e) With the assumptions in (d), show that for every $n \in \mathbb{N}^*$,

$$\sum_{k=0}^{n-1} c_k = \frac{f(b_n) - f(a_n)}{b_n - a_n}$$

and conclude that f is not differentiable at x. ✧

Exercise 6.10. Calculate the Jacobi matrix of $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$f(x) := \begin{pmatrix} x_1^2 x_2 - x_3 \\ \cos(x_2 - x_1) x_3^3 \end{pmatrix}.$$

Argue whether the Frechet derivative $\mathrm{d}f$ exists. If it exists calculate $\mathrm{d}f$. Moreover, calculate the directional derivative $\frac{\partial f}{\partial v}(0,0,1)$ for $v := \left(\frac{1}{2}\ \frac{1}{2}\ 1 \right)^\mathsf{T}$. ✧

Exercise 6.11. Calculate the Jacobi matrix of $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$f(x) := x_1 \begin{pmatrix} \cos(x_2)\cos(x_3) \\ \sin(x_2)\cos(x_3) \\ \sin(x_3) \end{pmatrix}.$$

Moreover, calculate $\det\left(\frac{\mathrm{d}f}{\mathrm{d}x}(x)\right)$. ✧

Exercise 6.12. Let $c \in \mathbb{R}$. Determine the Hesse matrix of $f\colon \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto c - x^2 - y^2$, and determine the local and global extrema. ✧

Exercise 6.13. Let $f\colon [0,1] \times [0,1] \to \mathbb{R}$ be given by

$$f(x,y) := 3x^2 - 2(y+1)x + 3y - 1.$$

Find all local extrema in $(0,1) \times (0,1)$ and all global extrema in $[0,1] \times [0,1]$. ✧

Exercise 6.14. Let X, Y be normed vector spaces, $D \subseteq X$ open and $f\colon D \to Y$. If the total differential $\mathrm{d}f$ is continuous, then all directional derivatives of f are continuous. ✧

Exercise 6.15. Regard the function $f\colon \mathbb{C} \setminus \{0\} \to \mathbb{C}$ given by $f(z) := \frac{1}{z}$ as function $\vec{f}\colon \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$ by decomposing z and $f(z)$ into their real- and imaginary part, i.e., a complex number z can be identified with the real vector $\vec{z} := \left(\begin{smallmatrix} \operatorname{Re} z \\ \operatorname{Im} z \end{smallmatrix} \right)$.

(a) Calculate $\mathrm{d}\vec{f}$.

(b) Show that for $z \in \mathbb{C} \setminus \{0\}$ and $h \in \mathbb{C}$

$$\mathrm{d}\vec{f}(\vec{z})\vec{h} \cong -\frac{h}{z^2}$$

in the sense that we either understand both sides as real vectors or complex numbers.

(c) Show that also in the complex number sense the derivative in $z \neq 0$ exists, i.e., the limit

$$\lim_{\substack{h \in \mathbb{C} \setminus \{0\} \\ h \to 0}} \frac{1}{h}\big(f(z+h) - f(z)\big)$$

exists. Note that the main difference is that $\frac{1}{h}$ is a complex number and the limit is taken in $\mathbb{C}$. ✧

Exercise 6.16. Let X, Y be normed vector spaces and $T\colon X \to Y$ a linear and bounded mapping, i.e., $T \in \mathcal{L}(X,Y)$. Show that the operatornorm

$$\|T\| := \sup_{x \in X \setminus \{\mathbf{0}\}} \frac{\|Tx\|_Y}{\|x\|_X}$$

is indeed a norm on $\mathcal{L}(X,Y)$. ✧

Exercise 6.17. Let X, Y and Z be normed vector spaces. Show that the operatornorm is *sub-multiplicative*, i.e.,

$$\|T_2 T_1\|_{\mathcal{L}(X,Z)} \leq \|T_2\|_{\mathcal{L}(Y,Z)} \|T_1\|_{\mathcal{L}(X,Y)}$$

for arbitrary $T_1 \in \mathcal{L}(X,Y)$ and $T_2 \in \mathcal{L}(Y,Z)$.

Use the sub-multiplicativity to show that for a Banach space X and $T \in \mathcal{L}(X)$ such that $\|T\| < 1$ the series

$$\sum_{k=0}^{\infty} T^k$$

is convergent and equals $(I-T)^{-1}$. ✧

Exercise 6.18. Let X be a Banach space[2] and $\operatorname{Inv} X$ denote the set of all invertible linear bounded operators, i.e., $\operatorname{Inv} X := \{T \in \mathcal{L}(X) \mid T \text{ is invertible}\}$.

(a) Show that $\operatorname{Inv} X$ is open in $\mathcal{L}(X)$ by showing that for every invertible T and every $S \in \mathrm{B}_{\frac{1}{\|T^{-1}\|}}(0)$ the operator $T + S$ is invertible.

(b) Calculate the derivative of $f\colon \operatorname{Inv} X \to \mathcal{L}(X)$, $T \mapsto T^{-1}$.

[2] A Banach space is a normed vector space that is complete, i.e., every Cauchy sequence converges.

Hint: You may use Exercise 6.17. ✧

Exercise 6.19. Let $f\colon \overline{\mathrm{B}}_1(0) \subseteq \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) \coloneqq 2x^3 - 3y^2,$$

where $\overline{\mathrm{B}}_1(0)$ is the closed unit ball, i.e., $\overline{\mathrm{B}}_1(0) = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. Find all local extrema in the interior (i.e., in $\mathrm{B}_1(0)$) and all global extrema (i.e., in $\overline{\mathrm{B}}_1(0)$). ✧

Exercise 6.20. Let $f\colon \overline{\mathrm{B}}_1(0) \subseteq \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) \coloneqq 4x^2 - 3xy - 1.$$

Find all local extrema in the interior and all global extrema. ✧

Exercise 6.21. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) \coloneqq \begin{cases} \dfrac{x^2y^3}{x^2+y^4}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Investigate whether f is (continuously) differentiable.

(b) Find all local extreme values.

(c) Find all global extreme values. ✧

Chapter 7

The Riemann integral for functions in several variables

7.1 Multidimensional integration

Definition 7.1 (Rectangles and grids)**.** Let Q be a subset of $\mathbb{R}^p$.

(i) We say Q is a p-dimensional *rectangle*, if it can be written as

$$Q = [a_1, b_1] \times \ldots \times [a_p, b_p] = \big\{\mathbf{x} \in \mathbb{R}^p \;\big|\; x_i \in [a_i, b_i] \;\forall i \in \{1, \ldots, p\}\big\},$$

and we denote its volume by $|Q| = (b_1 - a_1) \cdot \ldots \cdot (b_p - a_p)$.

(ii) A *grid* G on a p-dimensional rectangle Q is the collection of rectangles obtained by a subdivision of the sides of Q, that is, for each dimension $i = 1, \ldots, p$ there is a partition $\mathcal{P}_i(G) = \{x_1^i, \ldots, x_{\nu_i}^i\}$ of $[a_i, b_i]$, and the elements of G are all rectangles Q_j that are of the form

$$Q_j = I_1 \times \ldots \times I_p, \text{ with } I_i = \big[x_{k_i - 1}^i, x_{k_i}^i\big].$$

For two grids G_1, G_2 on the same Q, we say that G_2 is *finer* than G_1 if $\mathcal{P}_i(G_1) \subseteq \mathcal{P}_i(G_2)$ for all $i = 1, \ldots, p$. ✧

Definition 7.2. For a subset $E \subseteq \mathbb{R}^p$ contained in a rectangle Q and a grid G on Q, we call

$$V(E, G) := \sum_{\substack{Q_j \in G \\ Q_j \cap \overline{E} \neq \emptyset}} |Q_j|. \tag{7.1}$$

the *outer sum* of E with respect to G. ✧

Lemma 7.3. *For any subset $E \subseteq \mathbb{R}^p$, if G_2 is finer than G_1, then $V(E, G_2) \leq V(E, G_1)$.*

Proof. By G_2 being finer than G_1, each rectangle of G_1 is a finite union of rectangles of G_2, so we can write $V(E, G_1)$ as a sum of volumes of rectangles of G_2. Let us group them as

$$\mathcal{I}_1 = \left\{ Q \in G_2 \,\middle|\, Q \cap \overline{E} \neq \emptyset \right\},$$
$$\mathcal{I}_2 = \left\{ Q \in G_2 \setminus \mathcal{I}_1 \,\middle|\, Q \subseteq \widetilde{Q} \text{ where } \widetilde{Q} \in G_1 \text{ and } \widetilde{Q} \cap \overline{E} \neq \emptyset \right\}.$$

Then

$$V(E, G_1) = \sum_{Q \in \mathcal{I}_1} |Q| + \sum_{Q \in \mathcal{I}_2} |Q| \geq \sum_{Q \in \mathcal{I}_1} |Q| = V(E, G_2). \qquad ❑$$

Definition 7.4. A subset $E \subseteq \mathbb{R}^p$ is of *volume zero*, whenever given $\varepsilon > 0$ there is a rectangle $Q \supseteq E$ and a grid G on Q such that $V(E, G) \leq \varepsilon$. ✧

Definition 7.5. A subset $E \subseteq \mathbb{R}^p$ is called a *Jordan region*, whenever it is bounded and its boundary ∂E is of volume zero. In this case, we define

$$\mathrm{Vol}(E) \coloneqq \inf\{V(E, G) \mid G \text{ is a grid on } Q\}. \tag{7.2}$$

and call it the *Jordan content* or *volume* of E. ✧

Example 7.6. (a) Let $E = \mathbb{Q}^2 \cap [0, 1]^2$ be the points of the unit square with rational coordinates. Then since E is dense, $\partial E \cap [a_1, b_2] \times [a_2, b_2] \neq \emptyset$ for all $a_1 < b_1$ and $a_2 < b_2$, meaning that $V(\partial E, G) = 1$ for any grid G on $Q = [0, 1]^2$, and E is not a Jordan region.

(b) For $E = [\mathbf{x}, \mathbf{y}] = \{(1-t)\mathbf{x} + t\mathbf{y} \mid t \in [0, 1]\}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, it is reasonably easy to check that it is of volume zero and $\partial[\mathbf{x}, \mathbf{y}] = [\mathbf{x}, \mathbf{y}]$. This does not prevent it from (trivially) fitting in Definition 7.5, we just have $\mathrm{Vol}([\mathbf{x}, \mathbf{y}]) = 0$. ✧

Remark 7.7. In the definition above we make use of a rectangle Q containing E. However, the definition does not depend on which rectangle we choose. To see this, assume that f satisfies the definition with $Q = [a_1, b_1] \times \ldots \times [a_p, b_p]$ and let $\widetilde{Q} = [\tilde{a}_1, \tilde{b}_1] \times \ldots \times [\tilde{a}_p, \tilde{b}_p] \supseteq E$. If G is a grid on Q made up of partitions $\mathcal{P}_i(G)$, we can refine it to a new grid G' (still on Q) with

$$\mathcal{P}_i(G') = \mathcal{P}_i(G) \cup \{\max(a_i, \tilde{a}_i), \min(b_i, \tilde{b}_i)\},$$

and define $\widetilde{G}$ on $\widetilde{Q}$ by

$$\mathcal{P}_i(\widetilde{G}) = \mathcal{P}_i(G') \cap [\tilde{a}_i, \tilde{b}_i],$$

for which $V(E, \tilde{G}) = V(E, G')$ since the rectangles of either not in common with the other cannot intersect E. Since the definition of Vol involves an infimum over all such partitions and we can exchange the roles of Q, and $\tilde{Q}$, the result must be the same in both cases. ✧

Definition 7.8. Let $E \subseteq \mathbb{R}^p$ be bounded, $f\colon E \to \mathbb{R}$ be a bounded function which we extend to all of $\mathbb{R}^p$ by setting $f(\mathbf{x}) = 0$ for $\mathbf{x} \notin E$, Q an p-dimensional rectangle with $E \subseteq Q$, and G a grid on Q. We call

$$L(f,G) := \sum_{\substack{Q_j \in G \\ Q_j \cap \overline{E} \neq \emptyset}} |Q_j| \inf_{\mathbf{x} \in Q_j} f(\mathbf{x}), \quad \text{and} \quad U(f,G) := \sum_{\substack{Q_j \in G \\ Q_j \cap \overline{E} \neq \emptyset}} |Q_j| \sup_{\mathbf{x} \in Q_j} f(\mathbf{x})$$

the *lower sum* and *upper sum* of f on E with respect to G. ✧

Remark 7.9. If G_2 is finer than G_1, then

$$L(f,G_1) \leq L(f,G_2) \leq U(f,G_2) \leq U(f,G_1).$$

To see this, we can argue that as in Lemma 7.3 any rectangle $Q_j^1 \in G_1$ is a finite union of disjoint rectangles of G_2, that is,

$$Q_j^1 = \bigcup_{i=1}^{N_j} Q_i^2, \quad |Q_i^1| = \sum_{i=1}^{N_j} |Q_i^2|, \quad |Q_i^1| \sup_{\mathbf{x} \in Q_j^1} f(\mathbf{x}) \geq \sum_{i=1}^{N_j} |Q_i^2| \sup_{\mathbf{x} \in Q_i^2} f(\mathbf{x}),$$

so that summing over all Q_j^1 we get $U(f,G_1) \geq U(f,G_2)$, and similarly $L(f,G_1) \leq L(f,G_2)$. ✧

Definition 7.10. For $E \subseteq \mathbb{R}^p$ bounded, a bounded function $f\colon E \to \mathbb{R}$, we define the *lower integral* and *upper integral* of f on E as

$$\underline{\int_E} f(\mathbf{x})\,\mathrm{d}\mathbf{x} := \sup\{L(f,G) \mid G \text{ is a grid on } Q\}, \text{ and}$$
$$\overline{\int_E} f(\mathbf{x})\,\mathrm{d}\mathbf{x} := \inf\{U(f,G) \mid G \text{ is a grid on } Q\}.$$

The function f is said to be *Riemann integrable* whenever for every $\varepsilon > 0$ there is a grid G on a rectangle $Q \supseteq E$ for which

$$U(f,G) - L(f,G) < \varepsilon.$$

In this case we have

$$\underline{\int_E} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \overline{\int_E} f(\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{7.3}$$

and we call this common number the *integral* of f on E, which we denote as

$$\int_E f(\mathbf{x})\,\mathrm{d}\mathbf{x}\,.$$ ✧

Remark 7.11. As in Remark 7.7, this definition does not depend on Q, because we can transfer grids between rectangles containing E without affecting the value of the supremum and the infimum in (7.3). ✧

For the next theorem, we will need a lemma (which also contains the definition of a connected metric space):

Lemma 7.12. *Let $E \subseteq \mathbb{R}^p$ be convex. Then it is connected, that is,*

$$\Big(E = E_1 \cup E_2 \textit{ for } E_i = (\mathcal{O}_i \cap E),\ \mathcal{O}_i \subseteq \mathbb{R}^p \textit{ open}, E_1 \cap E_2 = \emptyset\Big) \implies \Big(E_1 = \emptyset \textit{ or } E_2 = \emptyset\Big).$$

Proof. Assume that E is not connected, that is, there are $\mathcal{O}_1, \mathcal{O}_2 \subseteq \mathbb{R}^p$ as in the statement with $\mathcal{O}_i \cap E \neq \emptyset$. Pick arbitrary $\mathbf{x}_i \in \mathcal{O}_i \cap E$ for $i = 1, 2$. Defining the continuous function $\mathbf{g}\colon [0,1] \to E$ with $\mathbf{g}(t) = (1-t)\mathbf{x}_1 + t\mathbf{x}_2$ (note that $\mathbf{g}(t) \in E$ because E is convex), the sets $\mathbf{g}^{-1}(\mathcal{O}_1 \cap E)$ and $\mathbf{g}^{-1}(\mathcal{O}_2 \cap E)$ are open in $[0,1]$, disjoint and nonempty, and we must have

$$[0,1] = \mathbf{g}^{-1}(\mathcal{O}_1 \cap E) \cup \mathbf{g}^{-1}(\mathcal{O}_2 \cap E),$$

but $[0,1]$ is connected. To see this, take U_1, U_2 disjoint, open in $[0,1]$ and with $U_1 \cup U_2 = [0,1]$, choose arbitrary $t_1 \in U_1$, $t_2 \in U_2$ with $t_1 \leq t_2$ (which is always possible, switching U_1 and U_2 if needed), and consider $c := \sup\{t \in U_1 \mid t \leq t_2\}$. Then $c \in [0,1]$ since it is an interval, and since $U_1 \cap [0,1]$ is closed in $[0,1]$, also $c \in U_1$. But U_1 is open in $[0,1]$, so if $c < 1$ we get a contradiction, and if $c = 1$ then we must have $t_2 = 1$ and $U_1 \cap U_2 \neq \emptyset$, again a contradiction. ❑

Theorem 7.13. *If $E \subseteq \mathbb{R}^p$ is a closed Jordan region and $f\colon E \to \mathbb{R}$ is continuous, then f is integrable.*

Proof. Since E is bounded and closed, by the Heine-Borel theorem it is compact and hence f attains its maximum and minimum, so it is in particular bounded and there is $M > 0$ for which $|f(\mathbf{x})| \leq M$ for all $\mathbf{x} \in M$. Now, fix $\varepsilon > 0$ and let Q be a rectangle such that $E \subseteq Q$. Since E is a Jordan region, there is a grid G_0 on Q for which

$$V(\partial E, G_0) < \frac{\varepsilon}{4M}.$$

Moreover, since f is continuous on a compact set it is uniformly continuous on E by Theorem 3.18, that is, there is $\delta > 0$ such that

$$\forall \mathbf{x}, \mathbf{y} \in E \text{ with } \|\mathbf{x} - \mathbf{y}\| < \delta \ :\ |f(\mathbf{x}) - f(\mathbf{y})| < \frac{\varepsilon}{2|Q|}.$$

Now, construct G_1 by refining G_0 in a way that $\|\mathbf{x} - \mathbf{y}\| < \delta$ if $\mathbf{x}, \mathbf{y} \in Q_j \in G_1$, which we can do by, for example, making sure that the coordinates $x_j^i \in \mathcal{P}_i(G_1)$ satisfy $|x_k^i - x_{k-1}^i| < \delta/\sqrt{p}$. Further, for any $Q_j \in G_1$ we denote

$$M_j = \sup_{\mathbf{x} \in Q_j} f(\mathbf{x}), \text{ and } M_j = \inf_{\mathbf{x} \in Q_j} f(\mathbf{x}).$$

Now we notice that any rectangle Q_j with $Q_j \cap \overline{E} \neq \emptyset$, since these are connected by Lemma 7.12, must either be completely contained in E° or intersect ∂E, because otherwise we would have $Q_j = (Q_j \cap E^\circ) \cup (Q_j \cap \mathbb{R}^p \setminus E)$. With this and Lemma 7.3, we can estimate:

$$\begin{aligned} U(f, G_1) - L(f, G_1) &= \sum_{Q_j \cap E \neq \emptyset} |Q_j|(M_j - m_j) \\ &= \sum_{Q_j \subseteq E^\circ} |Q_j|(M_j - m_j) + \sum_{Q_j \cap \partial E \neq \emptyset} |Q_j|(M_j - m_j) \\ &< \frac{\varepsilon}{2|Q|} \sum_{Q_j \subseteq E^\circ} |Q_j| + 2M \sum_{Q_j \cap \partial E \neq \emptyset} |Q_j| \\ &\leq \frac{\varepsilon}{2|Q|} |Q| + 2M\, V(\partial E, G_1) \\ &\leq \frac{\varepsilon}{2} + 2M\, V(\partial E, G_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

❑

Proposition 7.14. *If $f, g\colon E \to \mathbb{R}$ are integrable functions on a Jordan region $E \subseteq \mathbb{R}^p$ and $c_1, c_2 \in \mathbb{R}$, we have*

$$\int_E c_1 f(\mathbf{x}) + c_2 g(\mathbf{x})\, \mathrm{d}\mathbf{x} = c_1 \int_E f(\mathbf{x})\, \mathrm{d}\mathbf{x} + c_2 \int_E g(\mathbf{x})\, \mathrm{d}\mathbf{x},$$

and if $f(\mathbf{x}) \leq g(\mathbf{x})$ for all $\mathbf{x} \in E$, also

$$\int_E f(\mathbf{x})\, \mathrm{d}\mathbf{x} \leq \int_E g(\mathbf{x})\, \mathrm{d}\mathbf{x}\,.$$

Moreover, if $E_1, E_2 \subseteq E$ are Jordan regions with $E_1 \cap E_2 = \emptyset$ and f is integrable on both E_1 and E_2, then it is integrable on $E_1 \cup E_2$ and

$$\int_{E_1 \cup E_2} f(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{E_1} f(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_{E_2} f(\mathbf{x})\, \mathrm{d}\mathbf{x}\,.$$

Proof sketch/ideas. The first and second are immediate from the definition in terms of lower and upper sums. For the third, one can create a new grid by taking the unions of the partitions coming from E_1 and E_2. ❑

Theorem 7.15 (≈Fubini)**.** *Let $Q = [a_1, b_1] \times [a_2, b_2]$ and $f\colon Q \to \mathbb{R}$ be such that*

(i) *f integrable on Q.*

(ii) *$f(x_1, \cdot)$ is integrable on $[a_2, b_2]$ for each $x_1 \in [a_1, b_1]$.*

(iii) *$f(\cdot, x_2)$ is integrable on $[a_1, b_1]$ for each $x_2 \in [a_2, b_2]$.*

Then

$$\int_{[a_1,b_2]\times[a_2,b_2]} f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$
$$= \int_{a_2}^{b_2}\left[\int_{a_1}^{b_1} f(x_1,x_2)\,\mathrm{d}x_1\right]\mathrm{d}x_2 = \int_{a_1}^{b_1}\left[\int_{a_2}^{b_2} f(x_1,x_2)\,\mathrm{d}x_2\right]\mathrm{d}x_1\,.$$

Proof. We have defined grids G in terms of partitions for each coordinate, so it's convenient to relate two-dimensional lower and upper sums to the iterated integrals. We start by taking a grid G constructed from partitions $\mathcal{P}_i(G) = \{x_1^i, \ldots, x_{\nu_i}^i\}$ such that

$$U(f,G) - \varepsilon < \int_{[a_1,b_2]\times[a_2,b_2]} f(\mathbf{x})\,\mathrm{d}\mathbf{x} < L(f,G) + \varepsilon.$$

On it, we can estimate

$$\begin{aligned}
\overline{\int_{a_1}^{b_1}}\left(\int_{a_2}^{b_2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 &= \sum_{j=1}^{\nu_1}\overline{\int_{x_{j-1}^1}^{x_j^1}}\left(\sum_{k=1}^{\nu_2}\int_{x_{k-1}^2}^{x_k^2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1\\
&\le \sum_{j=1}^{\nu_1}\sum_{k=1}^{\nu_2}\overline{\int_{x_{j-1}^1}^{x_j^1}}\left(\int_{x_{k-1}^2}^{x_k^2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1\\
&\le \sum_{j=1}^{\nu_1}\sum_{k=1}^{\nu_2}(x_j^1 - x_{j-1}^1)(x_k^2 - x_{k-1}^2)\sup_{\mathbf{x}\in Q_{jk}} f(\mathbf{x})\\
&= U(f,G),
\end{aligned}$$

where we have denoted $Q_{jk} = [x_{j-1}^1, x_j^1]\times[x_{k-1}^2, x_k^2]$. Since the left hand side does not depend on ε and arguing similarly for the lower integrals and lower sums, we end up with

$$\begin{aligned}
\int_{[a_1,b_2]\times[a_2,b_2]} f(\mathbf{x})\,\mathrm{d}\mathbf{x} &\le \underline{\int_{a_1}^{b_1}}\left(\int_{a_2}^{b_2} f(x_1,x_2)\mathrm{d}x_2\right)\mathrm{d}x_1\\
&\le \overline{\int_{a_1}^{b_1}}\left(\int_{a_2}^{b_2} f(x_1,x_2)\mathrm{d}x_2\right)\mathrm{d}x_1\\
&\le \int_{[a_1,b_2]\times[a_2,b_2]} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\,.
\end{aligned}$$ ❑

Definition 7.16. For $\mathbf{x}\in\mathbb{R}^p$ and $j\in\{1,\ldots,p\}$ we define

$$\mathbf{x}_{(j)} \coloneqq (x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_p)^\mathsf{T},$$

i.e., $\mathbf{x}_{(j)}\in\mathbb{R}^{p-1}$ denotes the vector $\mathbf{x}$ without x_j. Furthermore, we call $E\subseteq\mathbb{R}^p$ a *projectable region*, if there is a closed Jordan region $H\subseteq\mathbb{R}^{p-1}$, an index $j\in\{1,\ldots,p\}$ and two continuous functions $\theta_\ell,\theta_u\colon H\to\mathbb{R}$ for which

$$E = \left\{\mathbf{x}\in\mathbb{R}^p \;\middle|\; \mathbf{x}_{(j)}\in H \text{ and } \theta_\ell(\mathbf{x}_{(j)}) \le x_j \le \theta_u(\mathbf{x}_{(j)})\right\}$$ ✧

Proposition 7.17. *If $E \subseteq \mathbb{R}^p$ is a projectable region arising in direction j over the Jordan region H and with the bounds θ_ℓ, θ_u. Then E is a Jordan region, and if $f\colon E \to \mathbb{R}$ is continuous, then*

$$\int_E f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_H \left[\int_{\theta_\ell(\mathbf{x}_{(j)})}^{\theta_u(\mathbf{x}_{(j)})} f(x_1,\dots,x_p)\,\mathrm{d}x_j\right]\mathrm{d}\mathbf{x}_{(j)}\,.$$

Proof sketch/ideas. We will see in the exercises that sets defined like E are Jordan regions. The formula for the integral is from Fubini, after extending f by zero to a rectangle containing E. ❏

Example 7.18. It is useful to remember that Proposition 7.17 can be used iteratively. For example, we can compute the integral of $f\colon E \to \mathbb{R}$ where E is the "half-ball"

$$E = \overline{\mathrm{B}}_1(0) \cap \{(x_1,x_2,x_3)^\mathsf{T} \mid x_3 \geq 0\} \subseteq \mathbb{R}^3$$

with an iterated integral (notice that there are different options!) as

$$\int_E f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{-1}^{1}\left[\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}}\left[\int_0^{\sqrt{1-(x_1^2+x_2^2)}} f(x_1,x_2,x_3)\,\mathrm{d}x_3\right]\mathrm{d}x_2\right]\mathrm{d}x_1\,.$$

✧

Theorem 7.19 (Change of variables)**.** *Let $E \subseteq \mathbb{R}^p$ be a Jordan region, and $\Phi\colon E \to \mathbb{R}^p$ be continuous on E, injective and C^1 on E°, such that $\Phi(E)$ is also a Jordan region, and with $\det \mathrm{d}\Phi(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in E^\circ$. Then, if $f \circ \Phi$ is integrable on E and f is integrable on $\Phi(E)$, we have*

$$\int_{\Phi(E)} f(\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_E f\big(\Phi(\mathbf{x})\big)\cdot|\det \mathrm{d}\Phi(\mathbf{x})|\,\mathrm{d}\mathbf{x}\,. \tag{7.4}$$

Proof sketch/ideas. The infimum and supremum of an integrable function are close to each other in most rectangles of finer and finer grids, and Jordan regions are well approximated by unions of rectangles. On the other hand, $|\det \mathrm{d}\Phi(\mathbf{x})|$ is the factor of change of volume of a rectangle under the linear transformation $\mathrm{d}\Phi(\mathbf{x})$, which is itself a local approximation of Φ near $\mathbf{x}$. The fact that we have to deal with the deformation by Φ of rectangles in grids on $Q \supseteq E$ and their relation with grids on $Q' \supseteq \Phi(E)$ makes the proof quite involved. ❏

Chapter 8

Conclusion and Outlook

In this lecture, we have established the foundational concepts of analysis by introducing:

- The real numbers $\mathbb{R}$;
- Sequences, series, and their limits;
- Continuity of functions;
- Derivatives in one and multiple variables;
- Integrals, and their connection to derivatives via the Fundamental Theorem of Calculus (Theorem 5.11).

These concepts form the basis for numerous applications in fields such as engineering, data science, and economics, where they provide powerful tools for modeling and solving real-world problems.

As you progress in your studies, you will build upon these ideas in areas such as differential equations, Fourier analysis, and optimization, which extend the power of analysis to even more complex problems. The typical "next step" in analysis covers more involved integrals (such as over curves and surfaces), topology and manifolds and, in modern presentations, hinges on a good portion of measure theory. See for instance, [6].

Furthermore, these concepts are integral to other branches of mathematics, such as topology and functional analysis, and have deep connections to real-world applications, ranging from machine learning to financial modeling.

Finally, we hope you have gained an appreciation for the unparalleled certainty and rigor of mathematical reasoning—something unmatched by other disciplines. To conclude with a touch of humor—after all, a picture says more than a thousand words—the `xkcd` comic Figure 8.1 captures the spirit perfectly.

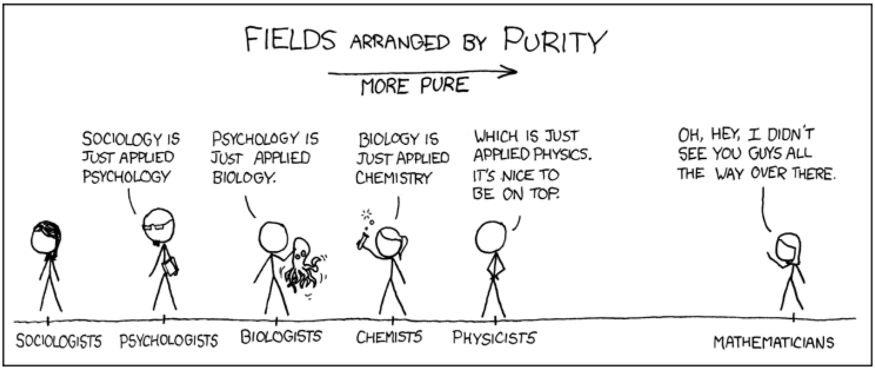


Figure 8.1: Purity: https://xkcd.com/435/

Appendix A

Schwarz's theorem

In this chapter we will give a more general version of Schwarz's theorem that does not rely on the continuity of the second derivative. Nevertheless we will restrict ourselves to the case where $f\colon D \subseteq \mathbb{R}^p \to Y$, i.e., the domain of f is a subset of $X = \mathbb{R}^p$ and the codomain Y is a (real) Hilbert space.

We follow [1] for the proof of Schwarz's theorem. However, in [1] the authors deal with an even more general case, namely $f\colon D \subseteq X \to Y$, where X, Y are real or complex Banach spaces. We present a less general version of Schwarz's theorem so that we avoid Banach space theory like the theorem of Hahn–Banach.

Lemma A.1 (Mean value inequality)**.** *Let Y be a real Hilbert space, $D \subseteq \mathbb{R}^p$ open, $f\colon D \subseteq \mathbb{R}^p \to Y$ be differentiable and $x, y \in D$ such that also the connecting line $[x, y]$ is contained in D, i.e., $[x, y] \subseteq D$. Then there exists a $\xi \in [x, y]$ such that*

$$\|f(y) - f(x)\|_Y \leq \|\mathrm{d}f(\xi)(y - x)\|_Y.$$

We will boil it down to the mean value theorem Theorem 4.16.

Proof. We define $g\colon [0, 1] \to \mathbb{R}$ by

$$g(t) \coloneqq \left\langle \frac{f(y) - f(x)}{\|f(y) - f(x)\|}, f(x + t(y - x)) - f(x) \right\rangle_Y.$$

Note that by construction we have g is continuous on $[0, 1]$ and differentiable on $(0, 1)$. Moreover,

$$g(0) = 0 \quad \text{and} \quad g(1) = \|f(y) - f(x)\|_Y.$$

Hence, by the mean value theorem Theorem 4.16 we obtain an $\eta \in (0, 1)$ such that

$$\|f(y) - f(x)\|_Y = g(1) - g(0) = \frac{g(1) - g(0)}{1 - 0} = g'(\eta).$$

By the chain rule and Cauchy–Schwarz we have

$$\begin{aligned} g'(\eta) &= \left\langle \frac{f(y)-f(x)}{\|f(y)-f(x)\|}, \mathrm{d}f(x+\eta(y-x))(y-x) \right\rangle_Y \\ &\leq \|\mathrm{d}f(x+\eta(y-x))(y-x)\|_Y. \end{aligned}$$

Defining $\xi = x+\eta(y-x)$ finishes the proof. ❑

Before we come to Schwarz's theorem we want to give the following characterization of differentiable.

Proposition A.2. *Let $D \subseteq \mathbb{R}^p$ be open, Y a Banach space and $f\colon D \subseteq \mathbb{R}^p \to Y$. Then the following are equivalent:*

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

(ii) *There exists a continuous function $\Phi_{x_0}\colon D \to \mathcal{L}(\mathbb{R}^p; Y)$ such that*

$$\forall y \in D \,:\, f(y) - f(x_0) = \Phi_{x_0}(y)(y-x_0).$$

Moreover, if (ii) holds then $\mathrm{d}f(x_0) = \Phi_{x_0}(x_0)$.

The property (ii) in Proposition A.2 is called *Carathéodory differentiable* and is—as shown in the following—equivalent to (Fréchet) differentiable. However, as this characterization does not involve error functions, it is convenient for proofs, especially for Schwarz's theorem.

Proof. $\boxed{\Leftarrow}$: We use as candidate for the derivative $A \coloneqq \Phi_{x_0}(x_0)$.

$$\begin{aligned} \|f(x+h) - f(x) - A(h)\| &= \|\Phi_{x_0}(x_0+h)(h) - \Phi_{x_0}(x_0)(h)\| \\ &\leq \|\Phi_{x_0}(x_0+h) - \Phi_{x_0}(x_0)\| \|h\| \end{aligned}$$

Hence, the error function $\phi(h) \coloneqq f(x+h) - f(x) - A(h)$ satisfies $\frac{\|\phi(h)\|}{\|h\|} \to 0$ for $h \to 0$. This shows that f is differentiable in x_0 with derivative $\mathrm{d}f(x_0) = A = \Phi_{x_0}(x_0)$.

$\boxed{\Rightarrow}$: If f is differentiable in x_0, then we define

$$\Phi_{x_0}(y) \coloneqq \begin{cases} \mathrm{d}f(x_0), & y = x_0, \\ \dfrac{f(y)-f(x_0)-\mathrm{d}f(x_0)(y-x_0)}{\|y-x_0\|} \left\langle \dfrac{y-x_0}{\|y-x_0\|}, \cdot \right\rangle + \mathrm{d}f(x_0), & y \neq x_0. \end{cases}$$

By construction we have for $y \in D$:

$$\begin{aligned} &\Phi_{x_0}(y)(y-x_0) \\ &\quad = \frac{f(y)-f(x_0)-\mathrm{d}f(x_0)(y-x_0)}{\|y-x_0\|} \underbrace{\left\langle \frac{y-x_0}{\|y-x_0\|}, y-x_0 \right\rangle}_{=\|y-x_0\|} + \mathrm{d}f(x_0)(y-x) \\ &\quad = f(y) - f(x_0). \end{aligned}$$

The continuity of Φ_{x_0} for $y \neq x_0$ is straightforward. Therefore, we will check continuity in x_0:

$$\begin{aligned}\|\Phi_{x_0}(y) - \Phi_{x_0}(x_0)\| &= \left\| \frac{f(y) - f(x_0) - \mathrm{d}f(x_0)(y - x_0)}{\|y - x_0\|} \left\langle \frac{y - x_0}{\|y - x_0\|}, \cdot \right\rangle \right\| \\ &\leq \left\| \left\langle \frac{y - x_0}{\|y - x_0\|}, \cdot \right\rangle \right\| \frac{\phi(y - x)}{\|y - x\|} = \frac{\phi(y - x)}{\|y - x\|} \to 0\end{aligned}$$

for $y \to x_0$, which finishes the proof. ❏

Note the function Φ_{x_0} mimics the one-dimensional secand: For $f \colon D \subseteq \mathbb{R} \to Y$

$$\frac{f(y) - f(x_0)}{y - x_0}.$$

Contrary to the one-dimensional case, we need to define Φ_{x_0} also for other directions, which is done in the proof by tangential directions. However, there might be multiple Φ_{x_0} that satisfy (ii), but $\Phi_{x_0}(x_0)$ will always be the same as the Frechet derivative is uniquely determined.

Theorem A.3 (Schwarz)**.** *Let Y be a Hilbert space and $f \colon D \subseteq \mathbb{R}^p \to Y$ differentiable such that $\mathrm{d}f$ is differentiable in $x_0 \in D$. Then for $u, v \in \mathbb{R}^p$ we have*

$$\mathrm{d}^2 f(x_0)(u, v) = \mathrm{d}^2 f(x_0)(v, u).$$

Proof. For fixed $u, v \in \mathbb{R}^p$ there exists a small enough $s > 0$ such that $x_0 + ru + tv \in D$ for $r, t < 2s$. We fix this s and define $h_s \colon [0, s] \to Y$ and $g_s \colon [0, s] \to Y$ by

$$h_s(t) \coloneqq f(x_0 + su + tv) - f(x_0 + tv) \quad \text{and} \quad g_s(t) \coloneqq h_s(t) - st\mathrm{d}^2 f(x_0)(v, u).$$

By the mean value inequality Lemma A.1 there exists a $\xi_s \in (0, s)$ such that

$$\|g_s(s) - g_s(0)\| \leq \|g'(\xi_s)(s - 0)\|$$

or equivalently ($g_s(0) = h_s(0)$)

$$\|h_s(s) - h_s(0) - s^2\mathrm{d}^2 f(x_0)(v, u)\| \leq \|h_s'(\xi_s)s - s^2\mathrm{d}^2 f(x_0)(v, u)\|. \quad \text{(A.1)}$$

by the chain rule we have

$$\begin{aligned}h_s'(t) &= \mathrm{d}f(x_0 + su + tv)(v) - \mathrm{d}f(x_0 + tv)(v) \\ &= \Big(\mathrm{d}f(x_0 + su + tv) - \mathrm{d}f(x_0)\Big)(v) - \Big(\mathrm{d}f(x_0 + tv) - \mathrm{d}f(x_0)\Big)(v)\end{aligned}$$

Proposition A.2 applied on $\mathrm{d}f(\cdot)(v)$ gives a Φ_{x_0} such that

$$\begin{aligned}h_s'(t) &= \Phi_{x_0}(x_0 + su + tv)(su + tv) - \Phi_{x_0}(x_0 + tv)(tv) \\ &= s\Phi_{x_0}(x_0 + su + tv)(u) + t\Phi_{x_0}(x_0 + su + tv)(v) - t\Phi_{x_0}(x_0 + tv)(v)\end{aligned}$$

In particular for $t = \xi_s$ we have

$$h_s'(\xi_s) = s\Phi_{x_0}(x_0 + su + \xi_s v)(u) \\ + \xi_s \Phi_{x_0}(x_0 + su + \xi_s v)(v) - \xi_s \Phi_{x_0}(x_0 + \xi_s v)(v) \quad \text{(A.2)}$$

Finally,

$$\begin{aligned} &\left\| \frac{h_s(s) - h_s(0)}{s^2} - \mathrm{d}^2 f(x_0)(v,u) \right\| \\ &\overset{\text{(A.1)}}{\leq} \left\| \frac{h_s'(\xi_s)}{s} - \mathrm{d}^2 f(x_0)(v,u) \right\| \\ &\overset{\text{(A.2)}}{\leq} \left\| \Phi_{x_0}(x_0 + su + \xi_s v)(u) - \mathrm{d}^2 f(x_0)(v,u) \right\| \\ &\quad + \frac{\xi_s}{s} \left\| \Phi_{x_0}(x_0 + su + \xi_s v)(v) - \Phi_{x_0}(x_0 + \xi_s v)(v) \right\| \end{aligned}$$

Since $\xi_s \in (0, s)$, we have $\frac{\xi_s}{s} \in (0,1)$ and $\xi_s \to 0$ for $s \to 0$. By the continuity of Φ_{x_0} we conclude

$$\lim_{s \to 0} \left\| \frac{h_s(s) - h_s(0)}{s^2} - \mathrm{d}^2 f(x_0)(v,u) \right\| = 0.$$

Note that $h_s(s) - h_s(0)$ is symmetric in $u, v \in \mathbb{R}^p$. Hence, if we repeat the previous steps with u, v swaped we obtain

$$\mathrm{d}^2 f(x_0)(u,v) = \lim_{s \to 0} \frac{h_s(s) - h_s(0)}{s^2} = \mathrm{d}^2 f(x_0)(v,u). \qquad ❑$$

Bibliography

[1] S. Arora, H. Browne, and D. Daners. An alternative approach to Fréchet derivatives. *J. Aust. Math. Soc.*, 111(2):202–220, 2021. doi:10.1017/S1446788720000166.

[2] O. Forster. *Analysis 1. Differential- und Integralrechnung einer Veränderlichen.* Grundkurs Math. Heidelberg: Springer Spektrum, 12th edition edition, 2016. doi:10.1007/978-3-658-11545-6.

[3] O. Forster. *Analysis 2. Differentialrechnung im* $\mathbb{R}^n$*, gewöhnliche Differentialgleichungen.* Grundkurs Math. Wiesbaden: Springer Spektrum, 11th edition edition, 2017. doi:10.1007/978-3-658-19411-6.

[4] H. Heuser. *Lehrbuch der Analysis. Teil 1.* Wiesbaden: Vieweg+Teubner, 17th revised ed. edition, 2009.

[5] M. Kaltenbäck. *Fundament Analysis*, volume 26 of *Berl. Studienr. Math.* Lemgo: Heldermann Verlag, 2014.

[6] M. Kaltenbäck. *Aufbau Analysis*, volume 27 of *Berl. Studienr. Math.* Lemgo: Heldermann Verlag, 2021.

[7] K. Königsberger. *Analysis 1.* Springer-Lehrb. Berlin: Springer, 5., neu bearbeitete Aufl. edition, 2001.

[8] K. Königsberger. *Analysis 2.* Springer-Lehrb. Berlin: Springer, 4., überarb. Aufl. edition, 2002.

[9] W. Rudin. *Principles of mathematical analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976.

[10] T. Tao. *An epsilon of room. I: Real analysis. Pages from year three of a mathematical blog*, volume 117 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2010.

[11] T. Tao. *Analysis I*, volume 37 of *Texts Read. Math.* New Delhi: Hindustan Book Agency, 4th edition edition, 2022. doi:10.1007/978-981-19-7261-4.

[12] T. Tao. *Analysis II*, volume 38 of *Texts Read. Math.* New Delhi: Hindustan Book Agency; Singapore: Springer, 4th edition edition, 2022. doi:10.1007/978-981-19-7284-3.

[13] V. A. Zorich. *Mathematical analysis I. Translated from the 4th and 6th Russian editions by Roger Cooke and Octavio T. Paniagua.* Universitext. Berlin: Springer, 2nd edition edition, 2015. doi:10.1007/978-3-662-48792-1.