Analysis 1+2

Jiaqi Wang

April 11, 2024

Contents

1 Sets, Spaces and Function

1.1 Metric Space

Definition 1.1.1 **(distance)** Let X be a set. A function $d : X \times X \to X$ is called a *distance* on X if it satisfies the following properties:

(i) Positivity: For all $a, b \in X$, it holds that $d(a, b) \geq 0$.

(ii) Non-degeneracy: For all $a, b \in X$, if $d(a, b) = 0$, then $a = b$.

(iii) Symmetry: For all $a, b \in X$, it holds that $d(a, b) = d(b, a)$.

(iv) Triangle inequality: For all $a, b, c \in X$, it holds that $d(a, c) \leq d(a, b) + d(b, c)$.

(v) Reflexivity: For all $a \in X$, it holds that $d(a, a) = 0$.

Usually conditions (ii) and (v) are combined into one condition: For all $a, b \in X, d(a, b) = 0$ if and only if $a = b$.

Definition 1.1.2 **(metric space)** A metric space is a pair $(X, dist)$, where X is a set and $dist$ is a distance function $dist : X \times X \to \mathbb{R}$ on X.

Example 1.1.3 Let $X = \{$Die Hard, Barbie, Oppenheimer$\}$

d	Die Hard	Barbie	Oppenheimer
Die Hard	0	5	2
Barbie	5	0	3
Oppenheimer	2	3	0

Then d is a distance function on X

Definition 1.1.4 **(ball in a metric space)** Let (X, d) be a metric space. Let $c \in X$ and $r \in \mathbb{R}$. The ball of radius r centered at c is the set

$$B(c, r) = \{x \in X | d(c, x) < r\}$$

Example 1.1.5 If $(X, d) = (\mathbb{R}, d_{\mathbb{R}})$, then $B(1, 3) = (-2, 4) = \{x \in \mathbb{R} \mid |x - 1| < 3\}$

Example 1.1.6 Let $X := \{$Die Hard, Barbie, Oppenheimer$\}$, with distance defined before. Then $B($Barbie, $4) = \{$ Barbie, Oppenheimer $\} = \{x \in X \mid d(x, Barbie) < 3\}$.

1.2 Normed Vector Spaces

Definition 1.2.1 **(norm)** Let V be a vector space over $\mathbb{R}$. A norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ such that

- Positivity: for all $u, v \in V$ we have $\|u\| \geq 0$ and $\|u\| = 0$ if and only if $u = 0$.
- Non-degeneracy: for all $u \in V$ if $\|u\| = 0$ then $u = 0$.
- Absolute Homogeneity: for all $u \in V$ and for all $\lambda \in \mathbb{R}$ we have $\|\lambda u\| = |\lambda| \|u\|$.
- Triangle inequality: for all $u, v \in V$ we have $\|u + v\| \leq \|u\| + \|v\|$.

Example 1.2.2 Let $V = \mathbb{R}^n$. Then $\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}$ defined by $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ is a norm on $\mathbb{R}^n$.

Proposition 1.2.3 Let $(V, \|\cdot\|)$ be a normed vector space. Then the function $d : V \times V \to \mathbb{R}$ defined by $d(u, v) = \|u - v\|$ is a distance on V. And (V, d) is a metric space.

Remark 1.2.4 (Notation for Euclidean distance on $\mathbb{R}^d$ and $\mathbb{R}$). We will usually write $\mathrm{dist}_{\mathbb{R}^d}$ instead of $\mathrm{dist}_{\|\cdot\|_2}$ for the standard (Euclidean) distance on $\mathbb{R}^d$. In particular, if $d \geq 2$, we have

$$\mathrm{dist}_{\mathbb{R}^d}(v, w) = \|v - w\|_2 = \sqrt{\sum_{i=1}^{d}(v_i - w_i)^2}$$

and if $d = 1$ we just have

$$\mathrm{dist}_{\mathbb{R}} = |v - w|$$

And if there is no room for confusion, we will just leave out the subscript altogether and write dist instead of $\mathrm{dist}_{\mathbb{R}^d}$.

1.3 The reverse triangle inequality

Lemma 1.3.1 **(Reverse triangle inequality)** Let $(V, \|\cdot\|)$ be a normed vector space. Then for all $v, w \in V$ we have,

$$|\|v\| - \|w\|| \leq \|v - w\|$$

2 Real Numbers

2.1 What are the real numbers?

Definition 2.1.1 **(Real numbers)** The real numbers are a complete totally ordered field.

2.2 The completeness axiom

Definition 2.2.1 **(Upper and Lower bound)** We say a number $M \in \mathbb{R}$ is an *upper bound* for a set $A \subseteq \mathbb{R}$ if

$$\forall a \in A[a \leq M].$$

We say a number $m \in \mathbb{R}$ is a *lower bound* for a set $A \subseteq \mathbb{R}$ if

$$\forall a \in A[a \geq m].$$

Given the definition of upper and lower bounds, we define what it means for a set to be bounded from above, bounded from below and just bounded.

Definition 2.2.2 **(bounded from above, bounded from below, bounded)** A set $A \subseteq \mathbb{R}$ is *bounded from above* if there exists an upper bound for A.
A set $A \subseteq \mathbb{R}$ is *bounded from below* if there exists a lower bound for A.
A set $A \subseteq \mathbb{R}$ is *bounded* if it is bounded from above and bounded from below.

Definition 2.2.3 **(Least upper bound (supremum))** Precisely, M is a *least upper bound* of a subset A if both

1. M is an upper bound of A.
2. For every upper bound $L \in \mathbb{R}$ of A, it holds that $M \leq L$.

Proposition 2.2.4 Suppose both M and W are a least upper bound of a subset $A \subseteq \mathbb{R}$. Then $M = W$.

Axiom 2.2.5 **(Completeness axiom)** We say that a totally ordered field $\mathbf{R}$ satisfies the *completeness axiom* if every nonempty subset of $\mathbf{R}$ that is bounded from above has a least upper bound.

Lemma 2.2.6 Every non-empty subset of the real line that is bounded from below has a *largest lower bound*.

Definition 2.2.7 **(infimum)** We usually call the largest lower bound of a non-empty set $A \subseteq \mathbb{R}$ that is bounded from below the *infimum* of A, and we denote it by $\inf A$.

2.3 Alternative characterizations of suprema and infima

Proposition 2.3.1 **(alternative characterizationa of supremum)** Let $A \subseteq \mathbb{R}$ be non-empty and bounded from above. Let $M \in \mathbb{R}$. Then M is the supremum of A if and only if

1. M is an upper bound for A,
2. and

$$\begin{aligned} &\text{for all } \varepsilon > 0, \\ &\quad \text{there exists } a \in A, \\ &\qquad a > M - \varepsilon. \end{aligned}$$

Proposition 2.3.2 (alternative characterizationa of infimum) Let $A \subseteq \mathbb{R}$ be non-empty and bounded from below. Let $m \in \mathbb{R}$. Then m is the infimum of A if and only if

1. m is a lower bound for A,
2. and

$$\begin{aligned}&\text{for all } \varepsilon > 0,\\ &\quad \text{there exists } a \in A,\\ &\quad\quad a < m + \varepsilon.\end{aligned}$$

These alternative characterizations of the supremum and infimum really provide a standard way to determining the supremum and infimum of subsets of the real line.

2.4 Maxima and minima

Definition 2.4.1 (maximum and minimum) Let $A \subseteq \mathbb{R}$ be a subset of the real numbers. We say that $y \in A$ is the *maximum* of A, and write $y = \max A$, if

$$\begin{aligned}&\text{for all } a \in A,\\ &\quad a \leq y.\end{aligned}$$

We say that $x \in A$ is the *minimum* of A, and write $x = \min A$, if

$$\begin{aligned}&\text{for all } a \in A,\\ &\quad a \geq x.\end{aligned}$$

Remark 2.4.2. Even if a set $A \subseteq \mathbb{R}$ is non-empty and bounded, it may not have a maximum or minimum. For example, the set $(0, 1)$ has no maximum or minimum.

Proposition 2.4.3 Let A be a subset of $\mathbb{R}$. If A has a maximum, then A is non-empty and bounded from above, and $\sup A = \max A$. If A has a minimum, then A is non-empty and bounded from below, and $\inf A = \min A$.

Proposition 2.4.4 Let A be a subset of $\mathbb{R}$. Assume that A is non-empty and bounded from above. If $\sup A \in A$ then A has a maximum and $\max A = \sup A$.

Proposition 2.4.5 Let A be a subset of $\mathbb{R}$. Assume that A is non-empty and bounded from below. If $\inf A \in A$ then A has a minimum and $\min A = \inf A$.

2.5 The Archimedean property

Proposition 2.5.1 (Archimedeean property) For every real number $x \in \mathbb{R}$ there exists a natural number $n \in \mathbb{N}$ such that $x < n$.

Given this proposition, we can define the ceiling function.

Definition 2.5.2 (ceiling function) The *ceiling function* $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{Z}$ is defined as follows. For $x \in \mathbb{R}$, $\lceil x \rceil$ denotes the smallest integer $z \in \mathbb{Z}$ such that $x \leq z$.

Proposition 2.5.3 For every two real numbers $a, b \in \mathbb{R}$ with $a < b$ there exists a $q \in \mathbb{Q}$ with $a < q < b$.

2.6 Computation rules for suprema

In the proposition below, we use the definitions

$$A + B = \{a + b \mid a \in A, b \in B\}$$

and

$$\lambda A = \{\lambda a \mid a \in A\}$$

for subsets $A, B \subseteq \mathbb{R}$ and a scalar $\lambda \in \mathbb{R}$.

Proposition 2.6.1 Let A, B, C, D be non-empty subsets of $\mathbb{R}$. Assume that A and B are bounded from above and C and D are bounded from below. Then

1. $\sup(A + B) = \sup A + \sup B$.
2. $\inf(C + D) = \inf C + \inf D$.
3. For all $\lambda \geq 0$, $\sup(\lambda A) = \lambda \sup A$.
4. For all $\lambda \leq 0$, $\sup(\lambda A) = \lambda \inf A$.
5. $\sup(-C) = -\inf C$.
6. $\inf(-C) = -\sup C$.

2.7 Bernoulli's inequality

Proposition 2.7.1 **(Bernoulli's inequality)** Let $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Then

1. If $x \geq -1$, then $(1 + x)^n \geq 1 + nx$.
2. If $x \geq 0$ and $n \geq 2$, then $(1 + x)^n \geq 1 + nx$.

3 Sequences

3.1 Sequence

Definition 3.1.1 **(Sequence)** A sequence is a function for which the domain is $\mathbb{N}$.

$$a : \mathbb{N} \to Y$$

Y can be any set.

Example 3.1.2 Here are some functions that are sequences:

1. $a : \mathbb{N} \to \mathbb{Q}$
2. $b : \mathbb{N} \to (\mathbb{N} \to Y)$
3. $c : \mathbb{N} \to \mathbb{N}$

And some functions that are not sequences:

1. $d : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$
2. $e : \mathbb{Q} \to \mathbb{N}$

3.2 Terminology around sequences

3.2.1 Bounded sequences

Definition 3.2.2 **(bouneded sequence)** Let (X, dist) be a metric space. We say a sequence $a : \mathbb{N} \to X$ is bounded if

$$\text{there exists } q \in X,$$

$$\text{there exists } M > 0,$$

$$\text{for all } n \in \mathbb{N},$$

$$\text{dist}(a_n, q) \leq M.$$

In a normed linear space, we can use a simpler criterion to check whether a sequence is bounded. That is the content of the following proposition.

Proposition 3.2.3 Let $(V, \|\cdot\|)$ be a normed vector space. Let $a : \mathbb{N} \to V$ be a sequence. The sequence a is bounded if and only if

$$\text{there exists } M > 0,$$

$$\text{for all } n \in \mathbb{N},$$

$$\|a_n\| \leq M.$$

3.3 Convergence of sequences

Definition 3.3.1 **(Convergence of sequences)** Let (X, dist) be a metric space. We say that a sequence $a : \mathbb{N} \to X$ converges to a point $p \in X$ if

$$\text{for all } \epsilon > 0,$$
$$\text{there exists } N \in \mathbb{N},$$
$$\text{for all } n \geq N,$$
$$\text{dist}(a_n, p) < \epsilon.$$

We sometimes write

$$\lim_{n\to\infty} a_n = p$$

to express that the sequence (a_n) converges to p.

Definition 3.3.2 (Divergence of sequences) Let (X, dist) be a metric space. A sequence $a : \mathbb{N} \to X$ is called *divergent* is it is not convergent.

3.4 Examples and limits of simple sequences

Proposition 3.4.1 (The constant sequence) Let (X, dist) be a metric space. Let $p \in X$ and assume that the sequence (a_n) is given by $a_n = p$ for every $n \in \mathbb{N}$. We also say that (a_n) is a constant sequence. Then $\lim_{n\to\infty} = p$.

Example 3.4.2 A standard limit Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence such that $a_n = 1/n$ for $n \geq 1$. Then $a : \mathbb{N} \to \mathbb{R}$ converges to 0.

Proof. Let $\epsilon > 0$. Choose $N = \lceil 1/\epsilon \rceil + 1$. Take $n \geq N$. Then

$$\text{dist}_{\mathbb{R}}(a_n, 0) = |a_n - 0| = |1/n| = 1/n \leq 1/N < \epsilon.$$

□

3.5 Uniqueness of limits

Proposition 3.5.1 (Uniqueness of limits) Let (X, dist) be a metric space and let $a : \mathbb{N} \to \mathbb{R}$ be a sequence in X. Assume that $p, q \in X$ and assume that

$$\lim_{n\to\infty} = p \text{ and } \lim_{n\to\infty} a_n = q$$

Then $p = q$.

3.6 More properties of convergent sequences

Proposition 3.6.1 Let (X, dist) be a metric space and suppose that $a : \mathbb{N} \to X$ is a sequence. Let $p \in X$. Then the sequence $a : \mathbb{N} \to X$ converges to p if and only if the real-valued sequence

$$n \mapsto \text{dist}(a_n, p)$$

converges to 0 in $\mathbb{R}$.

Proposition 3.6.2 (Convergent sequences are bounded) Let (X, dist) be a metric space. Let $a : \mathbb{N} \to X$ be a sequence in X converging to $p \in X$. Then the sequence $a : \mathbb{N} \to X$ is bounded.

Proposition 3.6.3 Let (X, dist) be a metric space and let $a : \mathbb{N} \to X$ and $b : \mathbb{N} \to X$ be two sequences. Let $p \in X$ and suppose that $\lim_{n\to\infty} a_n = p$. Then $\lim_{n\to\infty} b_n = p$ if and only if

$$\lim_{n\to\infty} \text{dist}(a_n, b_n) = 0$$

Corollary 3.6.4 **(Eventually equal sequences have the same limit)** Let (X, dist) be a metric space and let $a : \mathbb{N} \to X$ and $b : \mathbb{N} \to X$ be two sequences such that there exists an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$a_n = b_n$$

Then the sequence $a : \mathbb{N} \to X$ converges if and only if the sequence $b : \mathbb{N} \to X$ converges. If the sequences converge, they have the same limit.

3.7 Limit theorems for sequences taking values in a normed vector space

Theorem 3.7.1 Let $(V, \|\cdot\|)$ be a normed vector space and let $a : \mathbb{N} \to V$ and $b : \mathbb{N} \to V$ be two sequences. Assume that the $\lim_{n\to\infty} a_n$ exists and is equal to $p \in V$ and that the $\lim_{n\to\infty} b_n$ exists and is equal to $q \in V$. Let $\lambda : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence. Let $\mu \in \mathbb{R}$. Assume that $\lim_{n\to\infty} \lambda_n = \mu$. Then

1. The $\lim_{n\to\infty}(a_n + b_n)$ exists and is equal to $p + q$.
2. The $\lim_{n\to\infty}(\lambda_n a_n)$ exists and is equal to μp.

3.8 Index shift

Proposition 3.8.1 **(Index shift)** Let (X, dist) be a metric space and let $a : \mathbb{N} \to X$ be a sequence. Let $k \in \mathbb{N}$ and $p \in X$. Then the sequence $a : \mathbb{N} \to X$ converges to p if and only if the sequence $(a_{n+k})_n$ (i.e. the sequence $n \mapsto a_{n+k}$) converges to p.

4 Real-valued sequences

4.1 Terminology

Definition 4.1.1 **(increasing, decreasing and monotone sequences)** We say a sequence (a_n) is

1. *increasing* if for every $n \in \mathbb{N}, a_{n+1} \geq a_n$
2. *strictly increasing* if for every $n \in \mathbb{N}, a_{n+1} > a_n$
3. *decreasing* if for every $n \in \mathbb{N}, a_{n+1} \leq a_n$
4. *strictly decreasing* if for every $n \in \mathbb{N}, a_{n+1} < a_n$
5. *monotone* if it is either increasing or decreasing
6. *strictly monotone* if it is either strictly increasing or strictly decreasing

Definition 4.1.2 **(upper bound and lower bound for a sequence)** We say that a number $M \in \mathbb{R}$ is an *upper bound* for a sequence $a : \mathbb{N} \to \mathbb{R}$ if

$$\text{for all } n \in \mathbb{N}$$

$$a_n \leq M$$

We say that a number $m \in \mathbb{R}$ is a *lower bound* for a sequence $a : \mathbb{N} \to \mathbb{R}$ if

$$\text{for all } n \in \mathbb{N}$$

$$a_n \geq m$$

Definition 4.1.3 **(bounded sequence)** We say that a sequence $a : \mathbb{N} \to \mathbb{R}$ is *bounded above* if there exists an $M \in \mathbb{R}$ such that M is an upper bound for a.
We say that a sequence $a : \mathbb{N} \to \mathbb{R}$ is *bounded below* if there exists an $m \in \mathbb{R}$ such that m is a lower bound for a.

Proposition 4.1.4 Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. Then $a : \mathbb{N} \to \mathbb{R}$ is bounded if and only if it is both bounded above and bounded below.

4.2 Monotone, bounded sequences and convergent

Theorem 4.2.1 Let (a_n) be an increasing sequence that is bounded from above. Then (a_n) convergent and

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \quad (= \sup\{a_n \mid n \in \mathbb{N}\})$$

Theorem 4.2.2 Let (a_n) be a decreasing sequence that is bounded from below. Then (a_n) is convergent and

$$\lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n \quad (= \inf\{a_n \mid n \in \mathbb{N}\})$$

4.3 Limit theorems

Theorem 4.3.1 **(Limit theorems for real-valued sequences)** Let $a : \mathbb{N} \to \mathbb{R}$ and $b : \mathbb{N} \to \mathbb{R}$ be two converging sequences, and let $c, d \in \mathbb{R}$ be real numbers such that

$$\lim_{n\to\infty} a_n = c \text{ and } \lim_{n\to\infty} b_n = d.$$

Then

1. The $\lim_{n\to\infty}(a_n + b_n)$ exists and is equal to $c + d$.
2. The $\lim_{n\to\infty}(a_n b_n)$ exists and is equal to $c \cdot d$.
3. If $d \neq 0$, then $\lim_{n\to\infty}(\frac{a_n}{b_n}$ exists and is equal to $\frac{c}{d}$.
4. For every non-negative integer $m \in \mathbb{N}$, the limit $\lim_{n\to\infty}(a_n)^m$ exists and is equal to c^m.
5. If for every $n \in \mathbb{N}$, the number a_n is non-negative, then for every positive integer $k \in \mathbb{N} \setminus \{0\}$, the limit $\lim_{n\to\infty}(a_n)^{\frac{1}{k}}$ exists and is equal to $c^{\frac{1}{k}}$.

4.4 The squeeze theorem

Theorem 4.4.1 **(The squeeze theorem)** Let $a, b, c : \mathbb{N} \to \mathbb{R}$ be three sequences. Suppose that there exists an $N \in \mathbb{N}$ such that for every $n \geq N$, we have

$$a_n \leq b_n \leq c_n$$

and assume $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$ for some $L \in \mathbb{R}$. Then $\lim_{n\to\infty} b_n$ exists and is equal to L.

4.5 Divergence to ∞ and $-\infty$

Definition 4.5.1 We say a sequence $a : \mathbb{N} \to \mathbb{R}$ *diverges to* ∞ and write

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

if

$$\text{for all } M \in \mathbb{R},$$

$$\text{there exists } N \in \mathbb{N},$$

$$\text{for all } n \geq N,$$

$$a_n > M.$$

Similarly, we say a sequence (a_n) *diverges* to $-\infty$ and write

$$\lim_{n\to\infty} a_n = -\infty$$

if

$$\text{for all } M \in \mathbb{R},$$

$$\text{there exists } N \in \mathbb{N},$$

$$\text{for all } n \geq N,$$

$$a_n < M.$$

Proposition 4.5.2 Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence such that

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

Then the sequence (a_n) is bounded from below.

Similarly, let $b : \mathbb{N} \to \mathbb{R}$ be a sequence such that

$$\lim_{n\to\infty} b_n = -\infty.$$

Then the sequence (b_n) is bounded from above.

4.6 Limit theorems for improper limits

Theorem 4.6.1 **(Limit theorems for improper limits)** Let $a, b, c, d : \mathbb{N} \to \mathbb{R}$ be four sequences such that

$$\lim_{n\to\infty} a_n = \infty \text{ and } \lim_{n\to\infty} c_n = -\infty$$

the sequence (b_n) is bounded from below and the sequence (d_n) is bounded from above. Let $\lambda : \mathbb{N} \to \mathbb{R}$ be a sequence bounded below by some $\mu > 0$. Then

i. $\lim_{n\to\infty}(a_n + b_n) = \infty$

ii. $\lim_{n\to\infty}(c_n + d_n) = -\infty$

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

iv. $\lim_{n\to\infty}(\lambda_n c_n) = -\infty$

Proposition 4.6.2 Let $a : \mathbb{N} \to \mathbb{R}$ and $b : \mathbb{N} \to (0, \infty)$ be two sequences. Then

1. $\lim_{n\to\infty} a_n = \infty$ if and only if $\lim_{n\to\infty}(-a_n) = -\infty$.

2. $\lim_{n\to\infty} b_n = \infty$ if and only if $\lim_{n\to\infty} \frac{1}{b_n} = 0$.

4.7 Standard sequences

4.7.1 Geometric sequence

Proposition 4.7.2 **(Standard limit of of geometric sequence)** Let $q \in \mathbb{R}$. The sequence (a_n) defined by $a_n := q^n$ for $n \in \mathbb{N}$

- converges to 0 if $q \in (-1, 1)$
- converges to 1 if $q = 1$
- diverges to ∞ if $q > 1$
- diverges, but not to ∞ or $-\infty$ if $q \leq -1$

4.7.3 The n^{th} root of n

Proposition 4.7.4 **(Standard limit of the n^{th} root of n)** The sequence (a_n) defined by $a_n := \sqrt[n]{n}$ for $n \in \mathbb{N}$ converges to 1.

Corollary 4.7.5 Let $a > 0$. Then the sequence (b_n) defined by $b_n := \sqrt[n]{a}$ converges to 1.

4.7.6 The number e

First let's define the sequence (a_n) by

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

We show that (a_n) is increasing and bounded from above by 3. Hence (a_n) converges to some $e \in \mathbb{R}$ by the monotone convergence theorem.

Lemma 4.7.7 The sequence (a_n) defined by $a_n := \left(1 + \frac{1}{n}\right)^n$ for $n \in \mathbb{N} \setminus \{0\}$ and $a_0 = 1$ is increasing.

Lemma 4.7.8 The sequence (a_n) defined by $a_n := \left(1 + \frac{1}{n}\right)^n$ for $n \in \mathbb{N} \setminus \{0\}$ and $a_0 = 1$ is bounded from above by 3.

By these two lemmas, the sequence

$$n \mapsto \left(1 + \frac{1}{n}\right)^n$$

converges.

Definition 4.7.9 **((Standard limit of e))** We define the number e by

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

4.7.10 Exponentials beat powers

Proposition 4.7.11 Let $a \in (1, \infty)$ and let $p \in (0, \infty)$. Then

$$\lim_{n\to\infty} \frac{n^p}{a^n} = 0.$$

4.8 Sequences with values in $\mathbb{R}^d$

Proposition 4.8.1 Consider the metric space $(\mathbb{R}^d, \|\cdot\|_2)$. Let $z \in \mathbb{R}^d$ and let $x : \mathbb{N} \to \mathbb{R}^d$ be a sequence. Denote by y_i the ith component of a vector $y \in \mathbb{R}^d$. Then the sequence $(x^{(n)})$ converges to z if and only if for all $i \in \{1, \ldots, d\}$, the sequence $(x_i^{(n)})$ converges to z_i.

5 Series

5.1 Definition

Definition 5.1.1 Let $(V, \|\cdot\|)$ be a normed vector space and let $a : \mathbb{N} \to V$ be a sequence in V. Let $K \in \mathbb{N}$. We say that a series

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

is *convergent* if the associatedd *sequence of partial sums* $S_k : \mathbb{N} \to V$, i.e. the sequence $(S_K^n)_{n\in\mathbb{N}}$ converges. The term S_K^n is, for $n \in \mathbb{N}$, defined as

$$S_K^n := \sum_{k=K}^{n} a_k$$

If $K = 0$, we usually jusst write S^n or even S_n instead of S_0^n.
If the series $\sum_{n=K}^{\infty} a_n$ is convergent, the *value* of the series is by definition equal to the limit of the sequence of partial sums, i.e.

$$\sum_{k=K}^{\infty} a_k := \lim_{n\to\infty} S_k^n = \lim_{n\to\infty} \sum_{k=K}^{n} a_k$$

5.2 Geometric series

Proposition 5.2.1 Let $a \neq 1$ and $n \in \mathbb{N}$. Then

$$\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}.$$

Proof. We consider

$$\begin{aligned}
(1-a)\sum_{k=0}^{n} a^k &= \sum_{k=0}^{n} a^k - a\sum_{k=0}^{n} a^k \\
&= \sum_{k=0}^{n} a^k - \sum_{k=0}^{n} a^{k+1} \\
&= \sum_{k=0}^{n} a^k - \sum_{k=1}^{n+1} a^k \\
&= 1 - a^{n+1}
\end{aligned}$$

□

Proposition 5.2.2 (Geometric series) Let $a \in (-1, 1)$. Then the series

$$\sum_{k=0}^{\infty} a^k$$

is convergent and has the value

$$\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}.$$

5.3 The harmonic series

Proposition 5.3.1 **(Harmonic series)** The series

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

diverges.

5.4 The hyperharmonic series

Proposition 5.4.1 **(Hyperharmonic series)** Let $p > 1$. Then the series

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

converges.

Example 5.4.2 Here is an example of a series taking values in the normed vector space $(\mathbb{R}^2, \|\cdot\|)$:

$$\sum_{k=1}^{\infty} \left(\frac{1}{k^2}, \left(\frac{1}{2}\right)^k \right)$$

5.5 Only the tail matters for convergence

Lemma 5.5.1 Let $(V, \|\cdot\|)$ be a normed vector space and let $a : \mathbb{N} \to V$ be a sequence taking values in V. Let $K, L \in \mathbb{N}$. The series

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

is conovergent is and only if the series

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

is convergent. Moreover, if either the series converges, and $K < L$, then

$$\sum_{n=K}^{\infty} a_n = \sum_{n=K}^{L-1} + \sum_{n=L}^{\infty} a_n.$$

Proposition 5.5.2 Let $a : \mathbb{N} \to V$ be a sequence, let $M \in \mathbb{N}$ and assume that the series

$$\sum_{k=M}^{\infty} a_k$$

is convergent. Then

$$\lim_{m \to \infty} \sum_{k=m}^{\infty} a_k = 0.$$

Proposition 5.5.3 **(Index shift for series)** Let $a : \mathbb{N} \to V$ be a sequence, let $M \in \mathbb{N}$ and let $\ell \in \mathbb{N}$. Then the series

$$\sum_{k=M}^{\infty} a_k$$

converges if and only if the series

$$\sum_{k=M}^{\infty} a_{k+\ell}$$

converges. Moreoever, if either series converges, then

$$\sum_{k=M}^{\infty} a_{k+\ell} = \sum_{k=M+\ell}^{\infty} a_k.$$

5.6 Divergence test

Proposition 5.6.1 Let $(V, \|\cdot\|)$ be a normed vector space, and let $a : \mathbb{N} \to V$ be a sequence in V. Suppose the series $\sum_{n=0}^{\infty} a_n$ is convergent. Then

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

Proof. Suppose the series $\sum_{n=0}^{\infty} a_n$ is convergent to $L \in V$. Then

$$a_n = S_n - S_{n-1}$$

where S_n denote the partial sum $\sum_{k=0}^{n} a_k$. Because S_n and S_{n-1} are both convergent to L, the sequence (a_n) is convergent as well and converges to $L - L = 0$. □

Theorem 5.6.2 (Divergence test) Let $(V, \|\cdot\|)$ be a normed vector space and let $a : \mathbb{N} \to V$ be a sequence in V. Suppose the limit $\lim_{n\to\infty} a_n$ does not exist or is not equal to 0. Then the series

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

is divergent.

5.7 Limit laws for series

Theorem 5.7.1 (Limit laws for series) Let $(V, \|\cdot\|)$ be a normed vector space and let $a, b : \mathbb{N} \to V$ be sequences in V. Suppose the series

$$\sum_{n=0}^{\infty} a_n \quad \text{and} \quad \sum_{n=0}^{\infty} b_n$$

are convergent. Suppose $\lambda \in \mathbb{R}$. Then

1. The series
$$\sum_{n=0}^{\infty} (a_n + b_n)$$
is converget and converges to
$$\sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n.$$

2. The series
$$\sum_{n=0}^{\infty} \lambda a_n$$
is convergent and converges to
$$\lambda \sum_{n=0}^{\infty} a_n.$$

6 Series with positive terms

6.1 Comparison test

Theorem 6.1.1 (Comparison test) Let $a, b : \mathbb{N} \to [0, \infty)$ be two sequences. Assume that there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ we have $a_n \leq b_n$. Then

1. Suppose the series $\sum_{n=1}^{\infty} b_n$ converges. Then the series $\sum_{n=1}^{\infty} a_n$ converges as well.
2. Suppose the series $\sum_{n=1}^{\infty} a_n$ diverges. Then the series $\sum_{n=1}^{\infty} b_n$ diverges as well.

Example 6.1.2 Consider the series

$$\sum_{k=2}^{\infty} \frac{k}{k^2 - 1}.$$

We first observe that for all $k \geq 2$ we have

$$\frac{k}{k^2 - 1} \geq \frac{k}{k^2} = \frac{1}{k}.$$

Because the series

$$\sum_{k=2}^{\infty} \frac{1}{k}$$

diverges, the series

$$\sum_{k=2}^{\infty} \frac{k}{k^2 - 1}$$

diverges as well by the comparison test.

6.2 Limit comparison test

Theorem 6.2.1 (Limit comparison test) Let $a, b : \mathbb{N} \to [0, \infty)$ be two sequences.

1. Assume the series $\sum_{k=1}^{\infty} b_k$ converges and assume the limit

$$\lim_{n \to \infty} \frac{a_n}{b_n}$$

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

2. Assume the series $\sum_{k=1}^{\infty} b_k$ diverges and assume the limit

$$\lim_{n \to \infty} \frac{a_n}{b_n}$$

exists and is strictly larger than zero, or that the limit is infinity. Then the series $\sum_{k=1}^{\infty} a_k$ diverges as well.

Example 6.2.2 Consider the series

$$\sum_{k=1}^{\infty} \frac{1}{k^2 + 1}.$$

We use sequences $a, b : \mathbb{N} \to [0, \infty)$ defined for $k \geq 2$ by

$$a_k = \frac{k}{k^2 + 1}$$

and

$$b_k = \frac{1}{k}.$$

Then

$$\frac{a_k}{b_k} = \frac{\frac{k}{k^2+1}}{\frac{1}{k}} = \frac{1}{1 + \frac{1}{k^2}}.$$

By limit laws, we find that the limit of the denominator is 1, i.e.

$$\lim_{k\to\infty}\left(1 + \frac{1}{k^2}\right) = \lim_{k\to\infty} 1 + \lim_{k\to\infty}\frac{1}{k^2} = 1 + 0 = 1.$$

Therefore, we may apply the limit law for the quotient and conclude that

$$lim_{k\to\infty}\frac{a_k}{b_k} = \frac{1}{\lim_{k\to\infty}\left(1 + \frac{1}{k^2}\right)} = \frac{1}{1} = 1.$$

The series $\displaystyle\sum_{k=2}^{\infty}\frac{1}{k}$ diverges, and therefore it follows from the Limit Comparison Test that the series

$$\sum_{k=2}^{\infty} a_k = \sum_{k=2}^{\infty}\frac{k}{k^2+1}$$

diverges as well.

6.3 Ratio test

Theorem 6.3.1 **(Ratio Test)** Let $a : \mathbb{N} \to [0, \infty)$ be a sequence.

1. If there exists an $N \in \mathbb{N}$ and a $q \in (0, 1)$ such that for all $n \geq N$, it holds that
$$\frac{a_{n+1}}{a_n} \leq q,$$
then the series $\sum_{k=1}^{\infty} a_k$ converges.
2. If there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, it holds that
$$\frac{a_{n+1}}{a_n} \geq 1,$$
then the series $\sum_{k=1}^{\infty} a_k$ diverges.

6.4 Limit ratio test

Theorem 6.4.1 **(Limit Ratio Test)** Let $a : \mathbb{N} \to (0, \infty)$ be a sequence.

1. If $\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = q$ with $q \in [0, 1)$, then the series $\sum_{k=1}^{\infty} a_k$ converges.
2. If $\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = q$ with $q > 1$, or if $\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = \infty$, then the series $\sum_{k=1}^{\infty} a_k$ diverges.

Remark 6.4.2. We cannot conclude anything about the convergence of a series $\sum_k a_k$ when

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = 1.$$

6.5 Root test

Theorem 6.5.1 **(Root test)** Let (a_n) be a sequence of non-negative real numbers.

1. If there exists an $N \in \mathbb{N}$ and a $q \in (0, 1)$ such that for all $n \geq N$, it holds that
$$\sqrt[n]{a_n} \leq q,$$
then the series $\sum_{k=1}^{\infty} a_k$ converges.
2. If there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, it holds that
$$\sqrt[n]{a_n} \geq 1,$$
then the series $\sum_{k=1}^{\infty} a_k$ diverges.

6.6 Limit root test

Theorem 6.6.1 **(Limit Root Test)** Let (a_n) be a sequence of non-negative real numbers.

1. If $\lim_{n\to\infty} \sqrt[n]{a_n} = q$ with $q \in [0, 1)$, then the series $\sum_{k=1}^{\infty} a_k$ converges.
2. If $\lim_{n\to\infty} \sqrt[n]{a_n} = q$ with $q > 1$, or if $\lim_{n\to\infty} \sqrt[n]{a_n} = \infty$, then the series $\sum_{k=1}^{\infty} a_k$ diverges.

Remark 6.6.2. We cannot conclude anything about the convergence of a series $\sum_k a_k$ when
$$\lim_{n\to\infty} \sqrt[n]{a_n} = 1.$$

7 Series with general terms

7.1 Series with real terms: the Leibniz test

Theorem 7.1.1 (Leibniz test, a.k.a Alternating series test) Let $a, b : \mathbb{N} \to \mathbb{R}$ be two real-valued sequences such that for all $k \in \mathbb{N}$, $b_k = (-1)^k a_k$. Assume that there exists a $K \in \mathbb{N}$ such that

1. $a_k \geq 0$ for every $k \geq K$,
2. $a_k \geq a_{k+1}$ for every $k \geq K$,
3. $\lim_{k\to\infty} a_k = 0$.

Then, the series

$$\sum_{k=K}^{\infty} b_k = \sum_{k=K}^{\infty} (-1)^k a_k$$

is convergent. In addition, the following esitmate holds for every $N \geq K$,

$$\left| S_N - \sum_{k=K}^{\infty} b_k \right| \leq a_{N+1}.$$

where for all $n \in \mathbb{N}$, $S_n := \sum_{k=K}^{\infty} b_k$.

Example 7.1.2 We claim that the series

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

converges.

We would like to apply the Alternating series test. To do so, we need toe check its conditions.

We define the sequence $a : \mathbb{N} \to \mathbb{R}$ by

$$a_k := \frac{1}{k}$$

for $k \geq 1$ (and $a_0 = a_1 = 1$).

We now check the conditions for the Alternating Series Test.

1. We need to show that $a_k \geq 0$ for all $k \in \mathbb{N}$. Let $k \in \mathbb{N}$. Then,
$$a_k = \frac{1}{k} \geq 0.$$
2. We need to show that $a_k \geq a_{k+1}$ for all $k \in \mathbb{N}$. Let $k \in \mathbb{N}$. Then,
$$a_k = \frac{1}{k} \geq \frac{1}{k+1} = a_{k+1}.$$
3. We need to show that
$$\lim_{k\to\infty} a_k = 0$$
. This follow as this is a standard limit.

It follows from the Alternating Series Test that the series

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

converges.

7.2 Series charactersization of completeness in normed vector space

Definition 7.2.1 Let $(V, \|\cdot\|)$ be a normed vector space. Let $a : \mathbb{N} \to V$ be a sequence of vectors in V. We say the series

$$\sum_{k=0}^{\infty} a_k$$

converges *absolutely* if

$$\sum_{k=0}^{\infty} \|a_k\|$$

converges.

Definition 7.2.2 **(Series characterization of completeness)** We say a normed vector space $(V, \|\cdot\|)$ satisfies the *series characterization of completeness* if every series in V that is absolutely convergent is also convergent.

Proposition 7.2.3 Every finite-dimensional normed vector space satisfies the series characterization of completeness.

Example 7.2.4 Consider the series

$$\sum_{k=1}^{\infty} \frac{\sin(k)}{k^2}.$$

Since this is not an alternataing series, we cannot apply the Leibniz test.

However, for every k
$in \mathbb{N} \setminus \{0\}$, we have

$$\left|\frac{\sin(k)}{k^2}\right| \leq \frac{1}{k^2}.$$

The series

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

is a standard hyperharmonic seris, of which we know that it converges. By the Cmomparison Test, we conclude that the series

$$\sum_{k=1}^{\infty} \left|\frac{\sin(k)}{k^2}\right|$$

converges as well.

Therefore, the series

$$\sum_{k=1}^{\infty} \frac{\sin(k)}{k^2}$$

converges absolutely. Since $(\mathbb{R}, |\cdot|)$ is complete, we find that

$$\sum_{k=1}^{\infty} \frac{\sin(k)}{k^2}$$

converges.

Definition 7.2.5 Let $(V, \|\cdot\|)$ be a normed vector space. Let $a : \mathbb{N} \to V$ be a sequence. We say that a series

$$\sum_{k=0}^{\infty} a_k$$

converges *conditionally* if it converges but does not converge absolutely.

7.3 The Cauchy product

Theorem 7.3.1 (Cauchy product) Let $a, b : \mathbb{N} \to \mathbb{R}$ be two real-valued sequences. Assume that the series

$$\sum_{k=0}^{\infty} a_k$$

and

$$\sum_{k=0}^{\infty} b_k$$

converge absolutely. Then, the series

$$\sum_{k=0}^{\infty} c_k$$

converges absolutely as well, where

$$c_k := \sum_{\ell=0}^{k} a_\ell b_{k-\ell},$$

and

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right)$$

8 Subsequences, $\limsup$ and $\liminf$

8.1 Index sequences and subsequences

Definition 8.1.1 (Index sequence) We say a sequence $n : \mathbb{N} \to \mathbb{N}$ is an *index sequence* if it is strictly increasing.

Example 8.1.2 The sequence $n : \mathbb{N} \to \mathbb{N}$ defined by

$$n_k := 2k$$

is a strictly increasing sequence of natural numbers. In other words, it is an index sequence.

Definition 8.1.3 (Subsequence) Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. A sequence $b : \mathbb{N} \to \mathbb{R}$ is called a *subsequence* of a if there exists an index sequence $n : \mathbb{N} \to \mathbb{N}$ such that $b = a \circ n$

Just as we often write $(a_n)_{n\in\mathbb{N}}$ for a sequence called a, we often write $(a_{n_k})_{k\in\mathbb{N}}$ for the subsequence $a \circ n$.

8.2 (Sequential) accumulation points

Definition 8.2.1 ((Sequential) accumulation points) Let (X, dist) be a metric space. A point $p \in X$ is called an *accumulation point* of a sequence $a : \mathbb{N} \to X$ if there is a subsequence $a \circ n$ of a such that $a \circ n$ converges to p.

8.3 Subsequences of a converging sequence

Proposition 8.3.1 Let (X, dist) be a metric space. Let (a_n) be a sequence in X converging to $p \in X$. Then every subsequence of (a_n) is convergent to p.

8.4 $\limsup$

Consider a real-valued sequence (a_n) that is bounded from above and does not diverge to $-\infty$. We can then define a new sequence

$$k \mapsto \sup_{n\geq k} a_n.$$

Note that this sequence is decreasing, because for larger k the supremum is taken over a smaller set.

Lemma 8.4.1 Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence that is bounded from above and does not diverge to $-\infty$. Then, the sequence $k \mapsto \sup_{n\geq k} a_n$ is bounded from below.

Since the sequence $k \mapsto \sup_{n\geq k} a_n$ is decreasing and bounded from below, it has a limit, and the limit is in fact equal to the infumum of the sequence. This limit is called the $\limsup$

$$\begin{aligned}\limsup_{n\to\infty} a_n &:= \inf_{k\in\mathbb{N}} \sup_{n\geq k} a_n \\ &= \lim_{k\to\infty} \left(\sup_{n\geq k} a_n \right)\end{aligned}$$

Proposition 8.4.2 (Alternative characterization of $\limsup$) Let (a_n) be a real-valued sequence. Let $M \in \mathbb{R}$. Then, $M = \limsup_{n\to\infty} a_n$ if and only if

i. For every $\epsilon > 0$,
there exists $N \in \mathbb{N}$,
for all $\ell \geq N$,
$a_\ell < M + \epsilon$

ii.
For every $\epsilon > 0$,
for all $k \in \mathbb{N}$,
there exists $m \geq k$,
$a_m > M - \epsilon$

Theorem 8.4.3 Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence that is bounded from above and does not diverge to $-\infty$. Then $\limsup_{\ell\to\infty} a_\ell$ is a (sequential) accumulation point of a, i.e. there exists a subsequences of a that converges to $\limsup_{\ell\to\infty} a_\ell$.

Corollary 8.4.4 **(Bolzano-Weierstrass)** Every bounded, real-valued sequence has a subsequence that converges in $(\mathbb{R}, \text{dist}_{\mathbb{R}})$.

Theorem 8.4.5 Suppose a sequence $a : \mathbb{N} \to \mathbb{R}$ is bounded from above and does not diverge to $-\infty$. Then

$$\limsup_{\ell\to\infty} a_\ell$$

is the maximum of the set of sequential accumulation points.

8.5 lim inf

Similarly to the lim sup, we can define the lim inf. In some sense,

$$\liminf_{\ell\to\infty} a_\ell = -\limsup_{\ell\to\infty}(-a_\ell)$$

More precisely,

$$\begin{aligned}\liminf_{\ell\to\infty} a_\ell &:= \sup_{\ell\in\mathbb{N}} \inf_{k\geq\ell} a_k \\ &= \lim_{\ell\to\infty}\left(\inf_{k\geq\ell} a_k\right)\end{aligned}$$

Proposition 8.5.1 **(Alternative characterization of lim inf)** Let $a : \mathbb{N} \to \mathbb{R}$ and $M \in \mathbb{R}$. Then M equals $\liminf_{\ell\to\infty} a_\ell$ if and only if

1.
For every $\epsilon > 0$,
there exists $N \in \mathbb{N}$,
for all $\ell \geq N$,
$a_\ell > M - \epsilon$

2.
For every $\epsilon > 0$,
for all $K \in \mathbb{N}$,
there exists $m \geq K$,
$a_m < M + \epsilon$

Theorem 8.5.2 Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence that is bounded below and does not diverge to ∞. Then $\liminf_{\ell\to\infty} a_\ell$ is a sequential accumulation point of the sequence a, i.e. there is a subsequence of a that converges to $\liminf_{\ell\to\infty} a_\ell$.

Theorem 8.5.3 Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence that is bounded below and does not diverge to ∞. Then $\liminf_{\ell\to\infty} a_\ell$ is the minimum of the set of sequential accumulation points.

8.6 Relations between lim, lim sup and lim inf

Proposition 8.6.1 Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence and let $L \in \mathbb{R}$. Then $a : \mathbb{N} \to \mathbb{R}$ converges to L if and only if

$$\liminf_{\ell \to \infty} a_\ell = \limsup_{\ell \to \infty} = L$$

Proposition 8.6.2 Let $a, b : \mathbb{N} \to \mathbb{R}$ be two real-valued sequences, such that there exists an $N \in \mathbb{N}$ such that for all $\ell \geq N$, $a_\ell \leq b_\ell$. Then

$$\limsup_{\ell \to \infty} a_\ell \leq \limsup_{\ell \to \infty} b_\ell$$

and

$$\liminf_{\ell \to \infty} a_\ell \leq \liminf_{\ell \to \infty} b_\ell.$$

9 Point-set topology of metric spaces

Here we introduce three properties for subsets of a metric space: *closedness*, *completeness*, and *compactness*. For those three properties we known that every compact set is complete, and every complete set is closed. However, not every closed set is complete, and not every complete set is compact.

9.1 Open sets

Definition 9.1.1 (Open set) Let (X, dist) be a metric space. We say that a subset $O \subseteq X$ is *open* if every $x \in O$ is an interior point of O.

Now we need to say what it means to be an interior point.

Definition 9.1.2 (Interior point) Let (X, dist) be a metric space and let A be subset of X. A point $a \in A$ is called an *interior point* of A if

$$\begin{gathered}\text{there exists } r > 0 \\ B(a, r) \subseteq A\end{gathered}$$

where $B(a, r)$ is an (open) ball around point a with radius r (definition 1.1.4).

Proposition 9.1.3 Let (X, dist) be a metric space. The ball

$$B(p, r) := \{x \in X | \text{dist}(x, p) < r\}$$

is indeed open.

Proposition 9.1.4 ('Open' intervals are open) Let $a, b \in \mathbb{R}$ with $a < b$. Then the intervals $(a, b), (-\infty, b), (a, \infty)$ are all open subsets of $\mathbb{R}$.

Proposition 9.1.5 Let (X, dist) be a metric space. Then both the empty set $\emptyset$ and the set X itself (both of these are subsets of X) are open.

Proof. We first show that the empty set is open. We argue by contradiction. Suppose there exists a point $x \in \emptyset$ such that x is not an interior point of X. Then we have a contradiction, because the empty set has no elements.

We will now show that X is open. Let $x \in X$. We will show that x is an interior point, i.e. we will show that there exists an $r > 0$ such that $B(x, r) \subseteq X$.
Choose $r := 1$. Then $B(x, r) = B(x, 1) \subseteq X$. □

The set of all interior points of a subset $A \subseteq X$ is called the *interior* of the set A.

Definition 9.1.6 (The interior of a set) Let (X, dist) be a metric space and let $A \subseteq X$ be a subset of X. Then the *interior* of the set A, denoted by $\text{int}\, A$ is the set of all interior points of A, i.e $\text{int}\, A$ is defined as

$$\text{int}\, A := \{x \in A \mid x \text{ is an interior point of } A\}.$$

Example 9.1.7 The interior of the interval $[2, 5)$ (viewed as subset of $(\mathbb{R}, |\cdot|)$) is the interval $(2, 5)$.

The interior of a set is always open.

Proposition 9.1.8 Let (X, dist) be a metric space and let $A \subseteq X$. Then $\text{int}\, A$ is open.

The union of open sets is always open

Unions of open sets are always open. You may recall that if $\mathcal{I}$ is some set and if for every $\alpha \in \mathcal{I}$ we have a subset $A_\alpha \subseteq X$, then the union

$$\bigcup_{\alpha \in \mathcal{I}} A_\alpha$$

is defined as

$$\bigcup_{\alpha \in \mathcal{I}} A_\alpha := \{x \in X \mid \text{ there exists } \alpha \in \mathcal{I} \text{ such that } x \in A_\alpha\}$$

Proposition 9.1.9 Let (X, dist) be a metric space, let $\mathcal{I}$ be some set and assume that for every $\alpha \in \mathcal{I}$, we have a subset $O_\alpha \subseteq X$. Suppose that for all $\alpha \in \mathcal{I}$ the set O_α is open. Then also the union

$$\bigcup_{\alpha \in \mathcal{I}} O_\alpha$$

is open.

Example 9.1.10 We already know that for every $n \in \mathbb{N}$, the interval $(2n, 2n+1)$ is an open subset of $(\mathbb{R}, |\cdot|)$. Therefore, (choosing $\mathcal{I} = \mathbb{N}$ and $O_\alpha = (2\alpha, 2\alpha + 1)$ in the previous proposition,) we also know that the set

$$\bigcup_{n \in \mathbb{N}} (2n, 2n+1)$$

is an open set of $(\mathbb{R}, |\cdot|)$ as well.

Finite intersections of open sets are open

Proposition 9.1.11 Let (X, dist) be a metric space and let $O_1, \dots, O_N$ be open subsets of X. Then the intersection

$$O_1 \cap \dots \cap O_N$$

is also open.

Cartesian products of open sets

Proposition 9.1.12 Let $O_1, \dots, O_d$ be open subsets of $\mathbb{R}$. Then

$$O_1 \times \dots \times O_d (= \{(o_1, \dots, o_d) \mid o_i \in O_i\})$$

is an open subset of $(\mathbb{R}^d, \|\cdot\|_2)$.

9.2 Closed sets

Definition 9.2.1 Let (X, dist) be a metric space. We say that a subset $C \subseteq X$ is *closed* if its complement $X \setminus C$ is open.

Proposition 9.2.2 Let (X, dist) be a metric space. Then both the empty set $\emptyset$ and the set X itself (both of these are subsets of X) are closed.

Warning If you want to show that a set is closed *it is not enough* to show that the set is not open.

Proposition 9.2.3 (Sequence characterization of closedness) A set $C \subseteq X$ is closed if and only if for every sequence (c_n) in C converging to some $x \in X$, it holds that $x \in C$.

Example 9.2.4 Consider the subset A of the metric space $(\mathbb{R}^2, \|\cdot\|)$ defined by

$$A := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \leq (x_2)^2\}$$

Proof. By the sequence characterization of closedness, it suffices to show that for all sequences $y : \mathbb{N} \to A$, if the sequence y converges to some point $z \in \mathbb{R}^2$, then actually $z \in A$.
Let $y : \mathbb{N} \to A$ be a sequence in A.
Assume that the sequence (y) converges to some point $z \in \mathbb{R}^2$.
We need to show that $z \in A$.
Since y converges to z, we know that the components sequences y_1 and y_2 of y converge to the components z_1 and z_2 of z, namely

$$\lim_{n\to\infty} y_1^{(n)} = z_1 \quad \text{and} \quad \lim_{n\to\infty} y_2^{(n)} = z_2.$$

By limit theorems, we know that

$$\lim_{n\to\infty} \left(y_2^{(n)}\right)^2 = (z_2)^2.$$

Since for all $n \in \mathbb{N}, y^{(n)} \in A$, we also know that for all $n \in \mathbb{N}, y_1(n) \leq (y_2(n))^2$. Therefore,

$$z_1 = \lim_{n\to\infty} y_1^{(n)} \leq \lim_{n\to\infty} \left(y_2^{(n)}\right)^2 = (z_2)^2.$$

We conclude that indeed $z \in A$. ☐

Proposition 9.2.5 Let $a, b \in \mathbb{R}$ with $a < b$. Then the intervals $[a, b], (-\infty, b]$ and $[a, \infty)$ are all closed.

We now provide a few ways to create new closed sets out of sets about which you already know that they are closed.

Intersections of closed sets are always closed

Let (X, dist) be a metric space. If $\mathcal{I}$ is a set, and for every $\alpha \in \mathcal{I}$, we have a subset A_α of X, then the intersection

$$\bigcap_{\alpha\in\mathcal{I}} A_\alpha$$

is defined as

$$\bigcap_{\alpha\in\mathcal{I}} A_\alpha := \{x \in X \mid \text{ for all } \alpha \in \mathcal{I}, x \in A_\alpha\}.$$

Proposition 9.2.6 Let (X, dist) be a metric space. Let $\mathcal{I}$ be a set and suppose for every $\alpha \in \mathcal{I}$ we have a subset $C_\alpha \subseteq X$. Assume that for every $\alpha \in \mathcal{I}$ the set C_α is closed. Then the intersection

$$\bigcap_{\alpha\in\mathcal{I}} C_\alpha$$

is closed as well.

Finite unions of closed sets are closed

Proposition 9.2.7 Let (X, dist) be a metric space. Let $C_1, \dots, C_N$ be closed subsets of X. Then the finite union

$$C_1 \cup \cdots \cup C_N$$

is also closed.

Products of closed sets

Proposition 9.2.8 Let $C_1, \dots, C_d$ be closed subsets of $\mathbb{R}$. Then the Cartesian product

$$C_1 \times \cdots \times C_d (= \{(c_1, \dots, c_d) \mid c_i \in C_i\})$$

is a closed subset of $(\mathbb{R}^d, |\cdot|)$

The topological boundary of a set

Definition 9.2.9 (The topological boundary) Let (X, dist) be a metric space and let $A \subseteq X$. The *topological boundary* of a set A is denoted by ∂A and defined as

$$\partial A := X \setminus ((\text{int}A) \cup (\text{int}(X \setminus A)))$$

Example 9.2.10 The topological boundary of the interval $[2, 5)$ is the set $\{2, 5\}$ that consists of exactly the points 2 and 5.

9.3 Cauchy sequences

Definition 9.3.1 (Cauchy sequence) Let (X, dist) be a metric space. We say that a sequence $a : \mathbb{N} \to X$ is a Cauchy sequence if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\ &\quad \text{there exists } N \in \mathbb{N},\\ &\quad\quad \text{for all } m, n \geq N,\\ &\quad\quad\quad \text{dist}(a_m, a_n) < \epsilon\end{aligned}$$

Proposition 9.3.2 Every Cauchy sequence is bounded

Proposition 9.3.3 Let $a : \mathbb{N} \to X$ be a Cauchy sequence and assume that a has a subsequence converging to $p \in X$. Then the sequence a itself converge to p.

Proposition 9.3.4 Let (X, dist) be a metric space. Let (x_n) be a converging sequence in X. Then (x_n) is a Cauchy sequence.

9.4 Completeness

Definition 9.4.1 Let (X, dist) be a metric space. We say that a subset $A \subseteq X$ is *complete* (in (X, dist)) if every Cauchy sequence in A is convergent, with limit in A.
We also say the metric space (X, dist) itself is complete if X is a complete subset of X in (X, dist).

Theorem 9.4.2 The metric space $(\mathbb{R}, \text{dist}_{\mathbb{R}})$ is complete.

Proof. Let $a : \mathbb{N} \to \mathbb{R}$ be a Cauchy sequence. Because a is a Cauchy sequence, it is in particular bounded. as a consequence, by theorem 8.4.3, there is a subsequence $a \circ n$ such that $a \circ n$ converges to

$$\limsup_{k \to \infty} a_k$$

Finally, we know from proposition 9.3.3 that if a subsequence of a Cauchy sequence converges, that then the whole sequence converges. Therefore, the sequence $a : \mathbb{N} \to \mathbb{R}$ converges. □

Proposition 9.4.3 The metric space $(\mathbb{R}^d, \text{dist}_{\|\cdot\|_2})$ is complete, where $\|\cdot\|_2$ is the Euclidean norm.

Proposition 9.4.4 Let (X, dist) be a metric space. Suppose $A \subseteq X$ is complete. Then A is closed

Proposition 9.4.5 Let (X, dist) be a metric space and let $C \subseteq X$ be a complete subset. Let $A \subseteq C$ be a subset of C. Then, A is complete if and only if A is closed.

Series characterization of completeness in normed vector spaces

Theorem 9.4.6 Let $(V, \|\cdot\|)$ be a normed vector space. Then $(V, \|\cdot\|)$ is complete if and only if every absolutely converging series is convergent.

Corollary 9.4.7 Let $a : \mathbb{N} \to \mathbb{R}$ be a real-valued sequence. Suppose the series

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

converges absolutely, i.e. the series

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

converges. Then also the series

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

converges.

Example 9.4.8 The series

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{k^2}$$

converges, because it converges absolutely.

10 Compactness

10.1 Definition of (sequential) compactness

Definition 10.1.1 **((Sequential) compactness)** Let (X, dist) be a metric space. We say a subset $K \subseteq X$ is *(sequentially) compact* if every sequence $x : \mathbb{N} \to K$ has a converging subsequence $x \circ n$, converging to a point $z \in K$.

10.2 Boundedness and total boundedness

Definition 10.2.1 **(Bounded sets)** Let (X, dist) be a metric space. We say that a subset $A \subseteq X$ is *bounded* if

$$\begin{aligned}&\text{there exists } q \in X,\\ &\quad \text{there exists } M > 0,\\ &\quad\quad \text{for all } p \in A,\\ &\quad\quad\quad \mathrm{dist}(p, q) \le M.\end{aligned}$$

Just as with the concept of boundedness for sequences, in normed vector spaces boundedness has a somewhat easier alternative characterization.

Proposition 10.2.2 Let $(V, \|\cdot\|)$ be a normed linear space. A subset $A \subseteq V$ is bounded if and only if

$$\begin{aligned}&\text{there exists } M > 0,\\ &\quad \text{for all } v \in A,\\ &\quad\quad \|v\| \le M.\end{aligned}$$

Definition 10.2.3 **(Totally bounded sets)** Let (X, dist) be a metric space. We say that a subset $A \subseteq X$ is *totally bounded* if

$$\begin{aligned}&\text{for all } r > 0,\\ &\quad \text{there exists } N \in \mathbb{N},\\ &\quad\quad \text{there exists } p_1, \dots, p_N \in X,\\ &\quad\quad\quad A \subseteq \textstyle\bigcup_{i=1}^{N} B(p_i, r).\end{aligned}$$

In the next proposition we will say that "total boundedness" is a stronger property than just "boundedness".

Proposition 10.2.4 Let (X, dist) be a metric space and let A be a subset of X. If A is totally bounded, it is bounded.

In the special case of the normed vector space $(\mathbb{R}^d, \|\cdot\|_2)$, however, a subset is totally bounded if and only if it is bounded.

Proposition 10.2.5 Consider now the normed vector space $(\mathbb{R}^d, \|\cdot\|_2)$. A subset $A \subseteq \mathbb{R}^d$ is bounded in $(\mathbb{R}^d, \|\cdot\|_2)$ if and only if it is totally bounded.

10.3 Alternative characterization of compactness

Theorem 10.3.1 A subset $K \subseteq X$ is compact if and only if it is complete and totally bounded.

In the special case of $(\mathbb{R}^d, \|\cdot\|)$ we have an easier alternative characterization of compactness.

Theorem 10.3.2 **(Heine-Borel Theorem)** A subset of $(\mathbb{R}^d, \|\cdot\|_2)$ is compact if and only if it is closed and bounded.

11 Limits and continuity

We will consider functions $f : D \to Y$ mappings from a subset $D \subseteq X$ of a metric space (X, dist_X) to a metric space (Y, dist_Y). These are quite some actors: an input metric space (X, dist_X), a subset D of the metric space, and an output metric space (Y, dist_Y). And the concept of *limits* and *continuity* depend on all these actors.

On the coarsest level, if $p \in X$ and $q \in Y$, then the statement that

$$\lim_{x \to p} f(x) = q$$

will mean that if the distance between x and p is small, but not zero, the distance between $f(x)$ and q will be small.

11.1 Accumulation points

To get a useful concept of a limit in a point $p \in X$, the point p needs to be an *accumulation point* of the domain D of the function.

Definition 11.1.1 **(Accumulation points)** Let (X, dist_X) be a metric space and let $D \subseteq X$ be a subset of X. We say a point $p \in X$ is an *accumulation point* of the set D if

$$\begin{aligned} &\text{for all } \epsilon > 0, \\ &\quad \text{there exists } x \in D. \\ &\quad\quad 0 < \mathrm{dist}_X(x, p) < \epsilon \end{aligned}$$

We denote the set of accumulation points of a set D by D'.

Note that accumulation points of a set D do not have to lie in the set D themselves. If a point odes lie in D, but is not an accumulation point, then we call it an *isolated point* of D.

Definition 11.1.2 **(Isolated points)** Let (X, dist) be a metric space and let $D \subseteq X$ be a subset of X. We say a point $a \in D$ is an *isolated point* if it is not an accumulation point, i.e. if $a \in D \setminus D'$.

11.2 Limit in an accumulation point

We can now define limits in accumulation points of D.

Definition 11.2.1 **(Limit in an accumulation point)** Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$ be a subset of X. Let $f : D \to Y$ be a function and let $q \in Y$ be a point in Y. Let $a \in D'$ be an accumulation point of D. Then we say f converges to q as x goes to a, and write

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

if

$$\begin{aligned} &\text{for all } \epsilon > 0, \\ &\quad \text{there exists } \delta > 0, \\ &\quad\quad \text{for all } x \in D, \\ &\quad\quad\quad \text{if } 0 < \mathrm{dist}_X(x, a) < \delta, \text{ then } \mathrm{dist}_Y(f(x), q) < \epsilon. \end{aligned}$$

11.3 Uniqueness of limits

Proposition 11.3.1 Let (X, dist_X) and (Y, dist_Y) be metric spaces and let $D \subseteq X$ be a subset of X. Let $f : D \to Y$ be a function on D. Let $a \in D'$ and assume

$$\lim x \to a f(x) = p \quad \text{and} \quad \lim_{x \to a} f(x) = q$$

for points $p, q \in Y$. Then $p = q$.

11.4 Sequential characterization of limits

Theorem 11.4.1 **(Sequence characterization of limits)** Let (X, dist_X) and (Y, dist_Y) be two metric spaces. Let $D \subseteq X$. Let $f : D \to Y$ and let $a \in D'$. Let $q \in Y$. Then

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

if and only if

$$\begin{aligned}&\text{for all sequences } (x^n) \text{ in } D \setminus \{a\} \text{ converging to } a,\\ &\quad \lim_{n\to\infty} f(x^n) = q\end{aligned}$$

11.5 Limit laws

Theorem 11.5.1 Let (X, dist_X) be a metric space and let $(V, \|\cdot\|)$ be a normed vector space. Let $D \subseteq X$ and let $f : D \to V$ and $g : D \to V$ be two functions. Let $a \in D'$. Moreover, assume that the limit $lim_{n\to a} f(x)$ exists and equals $p \in V$ and that limii $\lim_{n\to a} g(x)$ exists and equals $q \in V$. Let $\lambda \in \mathbb{R}$. Then

1. The limit $\lim_{x\to a}(f(x) + g(x))$ exists and equals $p + q$.
2. The limit $\lim_{x\to a}(\lambda f(x))$ exists and equals λp.

11.6 Continuity

Definition 11.6.1 **(Continuity in a point)** Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$ be a subset of X. We say a function $f : D \to Y$ is *continuous* in a point $a \in D \cap D'$ if

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

If $a \in D$ is an isolated point, i.e. if $a \in D \setminus D'$, then we also say that f is continuous in a.

We say a function is continuous if it is continuous in every point in its domain.

Definition 11.6.2 **(Continuity on the domain)** Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$ be a subset of X. We say a function $f : D \to Y$ is *continuous on* D if f is continuous in a for every $a \in D$.

Sometimes it si a bit cumbersome to make the distinction between isolated points and accumulation points. The following alternative characterization of continuity in a point circumvecnts this issue.

Proposition 11.6.3 **(Alternative $\epsilon - \delta$ characterization of continuity in a point)** Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$ be a subset of X. Let $a \in D$. Then the function f is continuous in a if and only if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\ &\quad \text{there exists } \delta > 0,\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad \text{if } 0 < \text{dist}_X(x, a) < \delta, \text{ then } \text{dist}_Y(f(x), f(a)) < \epsilon.\end{aligned}$$

11.7 Sequence characterization of continuity

As with many concepts in analysis, continuity is conveniently probed with sequences.

Theorem 11.7.1 **(Sequence characterization of continuity)** Let (X, dist_X) and (Y, dist_Y) be metric spaces. Let $D \subseteq X$ and let $f : D \to Y$ be function. Let $a \in D$. The function f is continuous in a if and only if

$$\begin{aligned}&\text{for all sequences } (x^n) \text{ in } D \text{ converging to } a,\\&\quad \lim_{n\to\infty} f(x^n) = f(a).\end{aligned}$$

11.8 Rules for continuous functions

The following proposition implies that the composition of two continuous functions is also continuous.

Proposition 11.8.1 Let (X, dist_X), (Y, dist_Y) and $Z, \text{dist}_Z)$ be metric spaces, let $D \subseteq X$ and $E \subseteq Y$. Let $f : D \to Y$ and $g : E \to Z$ be two functions, and assume that $f(D) \subseteq E$. Let $a \in D$. If f is continuous in a and g is continuous in $f(a)$, then $g \circ f$ is continuous in a.

11.9 Images of compact sets under continuous functions are compact

Proposition 11.9.1 Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $K \subseteq X$ be a a compact subset of X. Let $f : K \to Y$ be continuous on K. Then $f(K)$ is a compact subset of Y.

11.10 Uniform continuity

Definition 11.10.1 Let (X, dist_X) and (Y, dist_Y) be metric spaces and let $D \subseteq X$ be a non-empty subset. We say that $f : D \to Y$ is *uniformly continuous* on D if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\&\quad \text{there exists } \delta > 0,\\&\qquad \text{for all } p, q \in D,\\&\qquad\quad 0 < \text{dist}_X(p, q) < \delta \implies \text{dist}_Y(f(p), f(q)) < \epsilon.\end{aligned}$$

The following proposition shows that *uniform continuity* is a stronger property that continuity.

Proposition 11.10.2 Let (X, dist_X) and (Y, dist_Y) be metric spaces and let $D \subseteq X$ be a non-empty subset. Let $f : D \to Y$ be uniformly continuous on D. Then f is continuous on D.

Although uniform continuity is a stronger property than continuity, it is not as strong as continuity on compact sets.

Theorem 11.10.3 Let (X, dist_X) and (Y, dist_Y) be metric spaces, let $K \subseteq X$ be compact and let $f : K \to Y$ be continuous on K. Then f is uniformly continuous on K.

12 Real-valued functions

12.1 More limit laws

Theorem 12.1.1 **(Limit laws for real-valued functinos)** Let (X, dist) be a metric space, let D be a subset of X and assume that $a \in D'$. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be two real-valued functions and assume that $\lim_{x\to a} f(x)$ exists and equals $M \in \mathbb{R}$ and that $\lim_{x\to a} g(x)$ exists and equals $L \in \mathbb{R}$. Then

1. For every $m \in \mathbb{N}$, the limit $\lim_{x\to a}(f(x))^m$ exists and equals M^m.
2. The limit $\lim_{x\to a}(f(x)g(x))$ exists and equals ML.
3. If $L \neq 0$, then the limit $\lim_{x\to a} \frac{f(x)}{g(x)}$ exists and equals $\frac{M}{L}$.
4. If for all $x \in D$, $f(x) \geq 0$, then for every $k \in \mathbb{N} \setminus \{0\}$,
$$\lim_{x\to a} \sqrt[k]{f(x)} = \sqrt[k]{M}$$

12.2 Building new continuous functions

The following theorem translates the limit laws from the previous section into statements about continuity.

Theorem 12.2.1 Let (X, dist) be a metric space, let D be a subset of X and assume $a \in D$. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be two real-valued functions that are continuous in a. Then

1. For every $m \in \mathbb{N}$, the function f^m is continuous in a.
2. The function $f + g$ is continuous in a.
3. The function fg is continuous in a.
4. If $g(a) \neq 0$, then the function $\frac{f}{g}$ is continuous in a.
5. If for all $x \in D$, $f(x) \geq 0$, then for every $k \in \mathbb{N} \setminus \{0\}$, the function $\sqrt[k]{f}$ is continuous in a.

12.3 Continuity of standard functions

Proposition 12.3.1 **(Polynomials are continuous)** Every (possibly multivariate) polynomial is continuous as a function from $(\mathbb{R}^d, \|\cdot\|_2)$ to $(\mathbb{R}, |\cdot|)$.

Proposition 12.3.2 **(Rational functions are continuous)** Every (possibly multivariate) rational function is continuous as a function from $(\mathbb{R}^d, \|\cdot\|_2)$ to $(\mathbb{R}, |\cdot|)$.

Proposition 12.3.3 **(Continuity of some standard functions)** The functions

$$\begin{aligned} \exp &: \mathbb{R} \to \mathbb{R} & \ln &: (0, \infty) \to \mathbb{R} \\ \sin &: \mathbb{R} \to \mathbb{R} & \arcsin &: [-1, 1] \to \mathbb{R} \\ \cos &: \mathbb{R} \to \mathbb{R} & \arccos &: [-1, 1] \to \mathbb{R} \\ \tan &: (-\pi/2, \pi/2) \to \mathbb{R} & \arctan &: \mathbb{R} \to \mathbb{R} \end{aligned}$$

are all continuous.

12.4 Limits from the left and from the right

Definition 12.4.1 (Limit from the left) Let (Y, dist_Y) be a metric space, and let $D \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$. Let $f : D \to Y$ be a function. Let $a \in \mathbb{R}$ be such that $a \in ((-\infty, a) \cap D)'$, i.e. such that a is an accumulation point in the set $(-\infty, a) \cap D$ in the metric space $(\mathbb{R}, \mathrm{dist}_\mathbb{R})$. Let $q \in Y$. We say that $f(x)$ *converges to q as x approaches a from the left (or from below)*, and write

$$\lim_{x \uparrow a} f(x) = q (\lim_{x \to a^-} f(x) = q)$$

if

$$\begin{aligned}
&\text{for all } \varepsilon > 0, \\
&\quad \text{there exists } \delta > 0, \\
&\quad\quad \text{for all } x \in D \cap (-\infty, a), \\
&\quad\quad\quad 0 < \mathrm{dist}_\mathbb{R}(x, a) < \delta \implies \mathrm{dist}_Y(f(x), q) < \varepsilon
\end{aligned}$$

Definition 12.4.2 (Limit from the right) Let (Y, dist_Y) be a metric space, and let $D \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$. Let $f : D \to Y$ be a function. Let $a \in \mathbb{R}$ be such that $a \in ((a, \infty) \cap D)'$, i.e. such that a is an accumulation point in the set $(a, \infty) \cap D$ in the metric space $(\mathbb{R}, \mathrm{dist}_\mathbb{R})$. Let $q \in Y$. We say that $f(x)$ *converges to q as x approaches a from the right (or from above)*, and write

$$\lim_{x \downarrow a} f(x) = q (\lim_{x \to a^+} f(x) = q)$$

if

$$\begin{aligned}
&\text{for all } \varepsilon > 0, \\
&\quad \text{there exists } \delta > 0, \\
&\quad\quad \text{for all } x \in D \cap (a, \infty), \\
&\quad\quad\quad 0 < \mathrm{dist}_\mathbb{R}(x, a) < \delta \implies \mathrm{dist}_Y(f(x), q) < \varepsilon
\end{aligned}$$

12.5 The extended real line

Definition 12.5.1 (The extended real line) The extended real line $\mathbb{R}_{\mathrm{ext}}$ is the union -f the set $\mathbb{R}$ and two symbols, "∞" and "$-\infty$". That is $\mathbb{R}_{\mathrm{ext}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$.

To turn $\mathbb{R}_{\mathrm{ext}}$ into a metric space, we need to define a distance function. First, we define a map $\iota : \mathbb{R}_{\mathrm{ext}} \to [-1, 1]$ by

$$\iota(x) = \begin{cases} -1 & \text{if } x = -\infty \\ \frac{x}{1+x} & \text{if } x \in \mathbb{R} \wedge x \geq 0 \\ \frac{x}{1-x} & \text{if } x \in \mathbb{R} \wedge x < 0 \\ 1 & \text{if } x = \infty \end{cases}$$

Because this function is injective, we can now build a distance on $\mathbb{R}_{\mathrm{ext}}$.

Definition 12.5.2 (Distance on extended real line) Given the definition of the injective function $\iota : \mathbb{R}_{\mathrm{ext}} \to [-1, 1]$ above, we define the distance on $\mathbb{R}_{\mathrm{ext}}$ by

$$\mathrm{dist}_{\mathbb{R}_{\mathrm{ext}}}(x, y) := dist_\mathbb{R}(\iota(x), \iota(y)) \qquad \text{for } x, y \in \mathbb{R}_{\mathrm{ext}}$$

12.6 Limits to ∞ or $-\infty$

Definition 12.6.1 (Divergence to ∞) Let (X, dist_X) be a metric space and $D \subseteq X$ and assume $a \in D'$. Let $f : D \to \mathbb{R}$. We say that f diverges to ∞ in a if

$$\begin{aligned}&\text{for all } M \in \mathbb{R},\\ &\quad \text{there exists } \delta > 0,\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad 0 < \mathrm{dist}_X(x,a) < \delta \implies f(x) > M\end{aligned}$$

Definition 12.6.2 **(Divergence to $-\infty$)** Let (X, dist_X) be a metric space and $D \subseteq X$ and assume $a \in D'$. Let $f : D \to \mathbb{R}$. We say that f diverges to $-\infty$ in a if

$$\begin{aligned}&\text{for all } M \in \mathbb{R},\\ &\quad \text{there exists } \delta > 0,\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad 0 < \mathrm{dist}_X(x,a) < \delta \implies f(x) < M\end{aligned}$$

Proposition 12.6.3 **(Alternative characterization of divergence to ∞)** Let (X, dist_X) be a metric space and $D \subseteq X$ and assume $a \in D'$. Let $f : D \to \mathbb{R}$. Then f diverges to ∞ in a if and only if f *converges* in a to the element $\infty \in \mathbb{R}_{\mathrm{ext}}$ when viewed as a function mapping from D as a subset of (X, dist_X) to the extended real line $(\mathbb{R}_{\mathrm{ext}}, \mathrm{dist}_{\mathbb{R}_{\mathrm{ext}}})$.

12.7 Limits at ∞ and $-\infty$

Definition 12.7.1 **(Limit at ∞)** Let (Y, dist_Y) be a metric space and let D be a subset of $\mathbb{R}$ that is unbounded from above. Let $q \in Y$ and $f : D \to Y$ be a function. We say that $f(x)$ converges to q as $x \to \infty$, and write

$$\lim_{x\to\infty} f(x) = q$$

if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\ &\quad \text{there exists } z \in \mathbb{R},\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad x > z \implies \mathrm{dist}_Y(f(x), q) < \epsilon\end{aligned}$$

Definition 12.7.2 **(Limit at $-\infty$)** Let (Y, dist_Y) be a metric space and let D be a subset of $\mathbb{R}$ that is unbounded from below. Let $q \in Y$ and $f : D \to Y$ be a function. We say that $f(x)$ converges to q as $x \to -\infty$, and write

$$\lim_{x\to-\infty} f(x) = q$$

if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\ &\quad \text{there exists } z \in \mathbb{R},\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad x < z \implies \mathrm{dist}_Y(f(x), q) < \epsilon\end{aligned}$$

We can also combine divergence to and at infinity.

Definition 12.7.3 **(Divergence to ∞ at ∞)** Let $D \subseteq \mathbb{R}$ be unbounded from above. Let $f : D \to \mathbb{R}$ be a function. We say that f diverges to ∞ as $x \to \infty$, and write

$$\lim_{x\to\infty} f(x) = \infty$$

if

$$\begin{aligned}&\text{for all } M \in \mathbb{R},\\ &\quad \text{there exists } z \in \mathbb{R},\\ &\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad x > z \implies f(x) > M\end{aligned}$$

Definition 12.7.4 **(Divergence to $-\infty$ at ∞)** Let $D \subseteq \mathbb{R}$ be unbounded from above. Let $f : D \to \mathbb{R}$ be a function. We say that f diverges to $-\infty$ as $x \to \infty$, and write

$$\lim_{x\to\infty} f(x) = -\infty$$

if

$$\begin{aligned}&\text{for all } M \in \mathbb{R},\\&\quad\text{there exists } z \in \mathbb{R},\\&\quad\quad\text{for all } x \in D,\\&\quad\quad\quad x > z \implies f(x) < M\end{aligned}$$

12.8 The Intermediate Value Theorem

Theorem 12.8.1 **(Intermediate Value Theorem)** Let $f : [a, b] \to \mathbb{R}$ be a continuous function and let $c \in \mathbb{R}$ be a value between $f(a)$ and $f(b)$. Then, there exists an $x \in [a, b]$ such that $f(x) = c$.

12.9 The Extreme Value Theorem

The *Extreme Value Theorem* states that a continuous, real-valued function defined on a non-empty, compact domain K always attains both a maximum and a minimum on K.

Theorem 12.9.1 **(Extreme Value Theorem)** Let (X, dist_X) be a metric space, $K \subseteq X$ be a non-empty compact subset and $f : K \to \mathbb{R}$ be continuous. Then f attains a maximum and a minimum on K.

12.10 Equivalence of norms

Definition 12.10.1 **(Equivalent norms)** Let V be a vector space and let $\|\cdot\|_A$ and $\|\cdot\|_B$ be two different norms on V. We say that the norms $\|\cdot\|_A$ and $\|\cdot\|\_B$ are *equivalent* if there exists a constant $c_1 > 0$ and $c_2 > 0$ such that for all $v \in V$

$$c_1\|x\|_A \leq \|x\|_B \leq c_2\|x\|_A.$$

Theorem 12.10.2 **(Equivalence of norms on finite-dimensional vector spaces)** Let V be a finite-dimensional vector space and let $\|\cdot\|_A$ and $\|\cdot\|_B$ be two norms on V. Then the norms $\|\cdot\|_A$ and $\|\cdot\|_B$ are equivalent.

Theorem 12.10.3 Let $(V, \|\cdot\|)$ be a finite-dimensional normed vector space. Then $(V, \|\cdot\|)$ is complete.

Theorem 12.10.4 **(Heine-Borel Theorem for finite-dimensional normed vector spaces)** Let $(V, \|\cdot\|)$ be a finite-dimensional normed vector space. Then a subset $A \subseteq V$ is compact if and only if A is closed and bounded.

12.11 Bounded linear maps and operator norms

Definition 12.11.1 (Linear map) Let V and W be two vector spaces. A function $L : V \to W$ is called a *linear map* if both

1. for all $a, b \in V$,
$$L(a + b) = L(a) + L(b)$$
2. for all $\lambda \in \mathbb{R}$ and $a \in V$,
$$L(\lambda a) = \lambda L(a)$$

Definition 12.11.2 (Bounded linear map) Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. We say that a linear map $L : V \to W$ is *bounded* if the image under L of the closed unit ball

$$\bar{B}_V(0, 1) = \{v \in V \mid \|v\|_V \le 1\}$$

is a bounded subset of $(W, \|\cdot\|_W)$, i.e. if
$$L(\bar{B}_V(0, 1))$$
is a bounded subset of $(W, \|\cdot\|_W)$.

Proposition 12.11.3 (Alternative characterization of bounded linear maps) Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. A linear map $L : V \to W$ is bounded if and only if there exists an $M > 0$ such that for all $v \in V$,
$$\|L(v)\|_W \le M\|v\|_V$$

Proposition 12.11.4 The space of bounded linear maps between one normed vector space to another is itself again a vector space, that we denote by $\mathrm{BLin}(V, W)$. Addition and scalar multiplication are defined pointwise, that means that if $L : V \to W$ and $K : V \to W$ are two linear maps and $\lambda \in \mathbb{R}$ is a scalar, then the linear map $L + K : V \to W$ is defined by

$$(L + K)(v) = L(v) + K(v)$$

and the map
$$(\lambda L)(v) = \lambda(L(v)).$$
The zero-element in this vector space $\mathrm{BLin}(V, W)$ is the map that maps every vector to the zero-element of W.

We now define the operator norm on the space of bounded linear maps.

Proposition 12.11.5 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. Consider the vector space $\mathrm{BLin}(V, W)$ of bounded linear maps $L : V \to W$. Then the function $\|\cdot\|_{V\to W} : \mathrm{BLin}(V, W) \to \mathbb{R}$ defined by
$$\|L\|_{V\to W} := \sup_{x\in\bar{B}_V(0,1)} \|L(x)\|_W$$
is a norm on $\mathrm{BLin}(V, W)$.

Definition 12.11.6 (Operator norm) The norm $\|\cdot\|_{V\to W}$ on the vectors space $\mathrm{BLin}(V, W)$ is called the *operator norm.*

Proposition 12.11.7 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. Let $L : V \to W$ be a bounded linear map. Then for all $v \in V$,

$$\|L(v)\|_W \le \|L\|_{V\to W}\|v\|_V$$

and in fact
$$\|L\|_{L\to W} = \min\{C \ge 0 \mid \forall v \in V, \|L(v)\|_W \le C\|v\|_V\}$$

Theorem 12.11.8 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces and assume that V is finite-dimensional. Let $L : V \to W$ be a linear map. Then L is bounded.

Theorem 12.11.9 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. Let $L : V \to W$ be a linear map. The function L is continuous if and only if it is bounded.

13 Differentiability

13.1 Definitions

Definition 13.1.1 (Differentiability in a point) Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two vector spaces. Let $\Omega \subseteq V$ be an open subset of V. Let $f : \Omega \to W$ be a function and let $a \in \Omega$. We say that f is differentiable in a if there exists a bounded linear map $L_a : V \to W$ such that if we define the error function $\mathrm{Err}_a : \Omega \to W$ by

$$\mathrm{Err}_a(x) := f(x) - f(a) - L_a(x - a)$$

it holds that

$$\lim_{x \to a} \frac{\|\mathrm{Err}_a(x)\|_W}{\|x - a\|_V} = 0$$

We call L_a the derivative of f in a and instead of L_a we often write $(Df)_a$.

Definition 13.1.2 (DFifferentiability on an open set) Let $(V, \|\cdot\|V)$ and $(W, \|\cdot\|_W)$ be two vector spaces. Let $\Omega \subseteq V$ be an open subset of V. We say that a function $f : \Omega \to W$ is differentiable on Ω if for every $a \in \Omega$, the function f is differentiable in a.

Proposition 13.1.3 Let $\Omega \subseteq \mathbb{R}$ be an open subset of $\mathbb{R}$ and consider a function $f : \Omega \to W$ interpreted as a function from the subset Ω of the normed vector space $(\mathbb{R}, |\cdot|)$ to the normed vector space $(W, \|\cdot\|_W)$. Then f is differentiable in $a \in \Omega$ if and only if the limit

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

exists. Moreover, if this limit exists, we call it $f'(a)$, and then for all $h \in \mathbb{R}$

$$f'(a) \cdot h = (Df)_a(h).$$

Remark 13.1.4. The limit above exists if and only if the limit

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

exists, and then they have the same value.

Example 13.1.5 Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x.$$

Let $a \in \mathbb{R}$, then

$$\begin{aligned} \lim_{x \to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x \to a} \frac{x - a}{x - a} \\ &= 1 \end{aligned}$$

13.2 The derivative as a function

Definition 13.2.1 (The derivative as a function) Let $f : \Omega \to W$ be a function from an open domain Ω in a finite-dimensional normed vector space $(V, \|\cdot\|_V)$ to a finite-dimensional normed vector space $(W, \|\cdot\|_W)$. Suppose that f is differentiable on Ω. Then we define the *derivative of f* as the function

$$Df : \Omega \to \mathrm{Lin}(V, W)$$

that maps every $a \in \Omega$ to the derivative of f in a.

13.3 Constant and linear maps are differentiable

Proposition 13.3.1 **(Constant maps are differentiable)** Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. Let $b \in W$ and consider the constant function $f : V \to W$ given by $f(v) = b$ for all $v \in V$. Then f is differentiable and for all $a \in V$, $(Df)_a = 0$ (zero map).

Proposition 13.3.2 **(Linear maps are differentiable)** Let $A : V \to W$ be a linear map between the finite-dimensional normed vector spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$. Then A is differentiable and for all $a \in V$, $(DA)_a = A$. Hence, the derivative of A is the constant function $DA : V \to \mathrm{Lin}(V, W)$ given by

$$a \mapsto A.$$

13.4 Bases and coordinates

Definition 13.4.1 **(Coordinate projections)** Let $i \in \{1, \ldots, m\}$ and consider the map $P^i : \mathbb{R}^m \to \mathbb{R}$ giiven by

$$P^i(x) = e_i^\top x = (e_i, x) = x_i.$$

The map P^i is called the projection to the i-th coordinate.

Proposition 13.4.2 Coordinate projections P^i in the above definition are linear.

Definition 13.4.3 **(Coordinate map)** Let W be a finite-dimensional vector space and assume that $w_1, \ldots, w_m$ is a basis of W. The map $\Psi : W \to \mathbb{R}^m$ that assigns to every $v \in W$ its coordinates with respect to the basis $w_1, \ldots, w_m$ is called the *coordinate map* with respect to the basis $w_1, \ldots, w_m$.

Proposition 13.4.4 The coordinate map $\Psi : W \to \mathbb{R}^m$ with respect to a basis $w_1, \ldots, w_m$ is linear.

Corollary 13.4.5 The derivative $D\Psi : W \to \mathrm{Lin}(W, \mathbb{R}^m)$ is given by

$$(D\Psi)_a = \Psi$$

for all $a \in W$.

Proposition 13.4.6 **(Dual basis)** If W is a finite-dimensional normed vector space, and $w_1, \ldots, w_m$ is a basis of W, then there exist linear maps $\Psi_i : W \to \mathbb{R}$ for $i = 1, \ldots, m$ such that for all $v \in W$,

$$v = \Psi_1(v)w_1 + \Psi_2(v)w_2 + \cdots + \Psi_m(v)w_m.$$

Together, the function $\Psi_1, \ldots, \Psi_m$ form a basis of the vecetor space $\mathrm{Lin}(W, \mathbb{R})$ and they are called *the dual basis* of $w_1, \ldots, w_m$.
Every Ψ_i is a linear map from W to $\mathbb{R}$, and

$$\Psi_i(w_j) = \begin{cases} 1 & \text{if } i = j \\ 2 & \text{if } i \neq j \end{cases}$$

13.5 The matrix representation

Definition 13.5.1 **(Jacobian with respect to bases)** We will sometimes call the matrix representation of a derivative $(Df)_a : V \to W$ the Jacobian of f (with respect to bases $v_1, \ldots, v_d$ and $w_1, \ldots, w_m$) in the point a, and we will denote it by $[Df]_a$.

13.6 The chain rule

Theorem 13.6.1 **(Chain rule)** Let $(U, \|\cdot\|_U), (V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces. Let $\Omega \subseteq U$ and $E \subseteq V$ both be open. Let $f : \Omega \to V$ be such that $f(\Omega) \subseteq E$. Let $g : E \to W$.
If f is differentiable in a point $a \in \Omega$ and g is differentiable in the point $f(a)$, then the function $g \circ f$ is differentiable in a and

$$(D(g \circ f))_a = (Dg)_{f(a)} \circ (Df)a.$$

13.7 Sum, product and quotient rules

Theorem 13.7.1 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces. Let $\Omega \subseteq V$ be an open and let $f : \Omega \to W$ and $g : \Omega \to W$ be two functions, that are both differentiable in a point $a \in \Omega$, with derivative $(Df)_a : V \to W$ and $(Dg)_a : V \to W$ respectively.
Then the function $f+g : \Omega \to W$ is also differentiable in a with derivative $(D(f+g))_a = (Df)_a + (Dg)_a$

Theorem 13.7.2 Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces. Let $\Omega \subseteq V$ be open and let $f : \Omega \to W$ and $g : \Omega \to \mathbb{R}$ be two functions. Assume both f and g are differentiable in a point $a \in \Omega$, with derivatives $(Df)_a : V \to W$ and $(Dg)_a : V \to \mathbb{R}$ respectively. Then

1. **Product rule:** The function $f \cdot g$ is differentiable in a with derivative given by

$$(D(f \cdot g))_a(h) = f(a)(Dg)_a(h) + g(a)(Df)_a(h)$$

for all $h \in V$

2. **Quotient rule:** If $g(a) \neq 0$, then the function f/g is differentiable in a with derivative given by

$$(D(f/g))_a(h) = \frac{1}{(}g(a)^2 g(a)(Df)_a(h) - f(a)(Dg)_a(h))$$

for all $h \in V$.

13.8 Differentiability of components

Proposition 13.8.1 Let $w_1, \ldots, w_m$ be a basis of W and let $\Psi_1, \ldots, \Psi_m$ be the dual basis.
Then a function $f : \Omega \to W$ is differentiable in a point $a \in \Omega$ if and only if for every $i \in \{1,$
$dots, m\}$, the function

$$\Psi_i \circ f$$

is differentiable in $a \in \Omega$. Moreover, if the function f is differentiable in $a \in \Omega$, tehn for every $v \in V$,

$$(Df)_a(v) = \sum_{i=1}^{m} w_i D(\Psi_i \circ f)_a(v)$$

Corollary 13.8.2 A function $f : \Omega \to \mathbb{R}^m$ is differentiable in a point $a \in \Omega$ if and only if for $i = 1, \ldots, m$ the component function $f_1 : \Omega \to \mathbb{R}$ given by $f_i = P^i \circ f$ is differentiable. Moreover, if f is differentiable in a, then for all $v \in V$.

$$(Df)_a(v) = \sum_{i=1}^{m} e_i (Df_i)_a(v) = ((Df_1)_a(v), \ldots, (Df_m)_a(v))$$

where e_i denote the standard unit vectors.
If in fact Ω is a subset of $\mathbb{R}$, then

$$f'(a) = (f_1'(a), \ldots, f_m'(a)).$$

13.9 Differentiability implies continuity

Theorem 13.9.1 Let $\Omega \subseteq V$ be open and suppose $f : D \to W$ is differentiable in a point $a \in \Omega$. Then f is continuous in a.

13.10 Derivative vanishes in local maxima and minima

Theorem 13.10.1 Let Ω be an open subset of a normed vector space V. Suppose $f : \Omega \to \mathbb{R}$ is differentiable in $a \in \Omega$. Suppose that $f(a)$ is a local maximum or minimum, i.e. suppose there exists an $r > 0$ such that either

$$\text{for all } x \in B(a, r),$$
$$f(x) \leq f(a)$$

or

$$\text{for all } x \in B(a, r),$$
$$f(x) \geq f(a).$$

Then $(Df)_a = 0$

13.11 The mean-value theorem

Theorem 13.11.1 **(Rolle's theorem)** Let $f : [a.b] \to \mathbb{R}$ be continuous, assume that f is differentiable on (a, b) and that $f(a) = f(b)$. Then there exists a $c \in (a, b)$ such that $f'(c) = 0$.

Theorem 13.11.2 **(Mean-value theorem)** Let $f : [a, b] \to \mathbb{R}$ be continuous, and assume that f is differentiable on (a, b). Then there exists a $c \in (a, b)$ such that

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

14 Differentiability of standard functions

- The constant function is differentiable
- Linear functions between finite dimensional normed vector spaces are always differentiable
- Sums, products, compositions and at times quotients of differentiable functions are differentiable.

14.1 Global context

We will consider two normed vector spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ and a function $f : \Omega \to W$ where $\Omega \subseteq V$ is an open subset of V. We will assume that V and W are finite-dimensional, and we will denote $v_1, \dots, v_d$ a basis in V, with corresponding coordinate map Φ, and by $w_1, \dots, w_m$ a basis of W with corresponding coordinate map Ψ.

14.2 Polynomials and rational functions are differentiable

Proposition 14.2.1 **(Differentiability of polynomials in one variable)** For every $n \in \mathbb{N}$, it holds that the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^n$ is differentiable with

$$f'(x) = nx^{n-1}.$$

In other words, the derivative of f. i.e. $(Df) : \mathbb{R} \to \mathrm{Lin}(\mathbb{R}, \mathbb{R})$ is given by

$$x \mapsto (h \mapsto nx^{n-1}h)$$

Proposition 14.2.2 **(Every polynomial is differentiable)** Every polynomial on $\mathbb{R}^d$ is differentiable

Proposition 14.2.3 **(Every rational function is differentiable on its domain)** Let $p : \mathbb{R}^d \to \mathbb{R}$ and $q : \mathbb{R}^d \to \mathbb{R}$ be two polynomials. Let

$$D := \{x \in \mathbb{R}^d \mid q(x) \neq 0\}.$$

Then the function $f : D \to \mathbb{R}$ given by

$$f(x) = \frac{p(x)}{q(x)}$$

is differentiable.

14.3 Differentiability of other standard functions

Proposition 14.3.1 The functions

$$\begin{aligned} &\exp : \mathbb{R} \to \mathbb{R} && \ln : (0, \infty) \to \mathbb{R} \\ &\sin : \mathbb{R} \to \mathbb{R} && \cos : \mathbb{R} \to \mathbb{R} \\ &\tan : (-\pi/2, \pi/2) \to \mathbb{R} && \arctan : \mathbb{R} \to \mathbb{R} \end{aligned}$$

are all differentiable on their domain, while the functions

$$\arcsin : [-1, 1] \to \mathbb{R} \qquad \arccos : [-1, 1] \to \mathbb{R}$$

are both differentiable on the interval $(-1, 1)$.
The derivatives are given by:

$$\exp'(t) = \exp(t) \qquad \ln'(t) = 1/t$$
$$\sin'(t) = \cos(t) \qquad \cos'(t) = -sin(t)$$
$$\tan'(t) = \frac{1}{\cos^2(t)} \qquad \arctan'(t) = \frac{1}{1+t^2}$$
$$\arcsin'(t) = \frac{1}{\sqrt{1-t^2}} \qquad \arccos'(t) = -\frac{1}{\sqrt{1-t^2}}$$

Example 14.3.2 Consider the function $f : \mathbb{R} \to \mathbb{R}^2$ given by

$$f(t) = (t^2, \sin(t))$$

The component functions $f_1 : \mathbb{R} \to \mathbb{R}$ and $f_2 : \mathbb{R} \to \mathbb{R}$ are given by

$$f_1(t) = t^2$$

and

$$f_2(t) = \sin(t)$$

Since these components are differentiable standard functions, we find that f is differentiable as well and

$$f'(t) = (f_1'(t), f_2'(t)) = (2t, \cos(t))$$

15 Directional and partial derivatives

15.1 A recurring and very important construction

15.2 Directional derivatives

Definition 15.2.1 **(Directional derivative)** Let $f : \Omega \to W$ be a function from an open domain Ω in a finite-dimensional normed vector space V to a finite-dimensional normed vector space W. Let $a \in \Omega$ and $v \in V$.
Then we say the *directional derivative* in the direction of v of f exists in the point $a \in \Omega$ if there exists a $\delta > 0$ such that the function

$$g := f \circ \ell_{a,v} : (-\delta, \delta) \to W$$

is differentiable at 0, where the function $\ell_{a,v} : (-\delta, \delta) \to V$ is defined by

$$\ell_{a,v} := a + tv.$$

Moreover, if it exists, we define the directional derivative in the direction of v of f in the point a as

$$(D_v f)_a := g'(0) = \lim_{h \to 0} \frac{f(a + hv) - f(a)}{h}.$$

Proposition 15.2.2 Suppose $f : \Omega \to W$ is differentiable in a point $a \in \Omega$. Then for all $v \in V$, the directional derivative of f at a in the direction of v

$$(D_v f)_a$$

exists and is equal to the derivative of f at the point a applied to the vector v

$$(Df)_a(v).$$

Warning There are functions $f : \Omega \to W$ that are *not differentiable in a point* a even though for every $v \in V$, the directional derivative $(D_v f)_a$ exists.

Example 15.2.3 Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x_1, x_2) := \begin{cases} x_1, & x_2 \neq 0 \\ 0, & x_2 = 0 \end{cases}$$

Let us verify that for all $v \in \mathbb{R}^2$, the directional derivative $(D_v f)_0$ exists. Let $v \in \mathbb{R}^2$. If $v_2 \neq 0$, then

$$\begin{aligned} (D_v f)_0 &= \lim_{t \to 0} \frac{f(0 + tv) - f(0)}{t} \\ &= \lim_{t \to 0} \frac{tv_1 - 0}{t} \\ &= \lim_{t \to 0} v_1 = v_1 \end{aligned}$$

while if $v_2 = 0$ then

$$\begin{aligned} (D_v f)_0 &= \lim_{t \to 0} \frac{f(0 + tv) - f(0)}{t} \\ &= \lim_{t \to 0} \frac{0 - 0}{t} = 0 \end{aligned}$$

In both cases, the directional derivative exists.

We now claim, that f is not differentiable in 0. Indeed, if f would be differentiable in 0, then the derivative $(Df)_0$ would be a linear map from $\mathbb{R}^2 \to \mathbb{R}$. Since $(Df)_0(e_1) = 0$ and $(Df)_0(e_2) = 0$, in fact $(Df)_0$ maps every vector to zero. In particular, $(D_{(1,1)} f)_0 = 0$. However, the computation above shows that $(D_{(1,1)} f)_0 = 1$. This is a contradiction.

15.3 Partial derivatives

Partial derivative are special types of directional derivatives, for functions that are defined on the vector space $\mathbb{R}^d$.

Definition 15.3.1 (Partial derivative) Let $f : \Omega \to W$ be a function defined on an open domain $\Omega \subseteq \mathbb{R}^d$. The i-th *partial derivative* in a point $a \in \Omega$, denoted by

$$\frac{\partial f}{\partial x_i}(a),$$

is the directional derivative in the direction of the i-th unit vector e_i

$$\frac{\partial f}{\partial x_i}(a) := \frac{d}{dt} f(a + te_i)\bigg|_{t=1} = \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}.$$

Example 15.3.2 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x_1, x_2) = x_1^2 + 2x_1 x_2 + 3x_2^4.$$

Let us determine whether the partial derivative

$$\frac{\partial f}{\partial x_2}$$

exists in the point $a := (a_1, a_2)$.

To do so, by definition, we need to see if the directional derivative of f in the direction e^2 in the point a, namely

$$(D_{e_2} f)_a$$

exists. I.e. we need to verify whether the derivative of the function $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(t) := (f \circ \ell_{a,e_2}(t) = f(a + te_2)$$
$$= f((a_1, a_2 + t)) = a_1^2 + 2a_1(a_2 + t) + 3(a_2 + t)^4$$

exists in the point $t = 0$. Since g is a polynomial in one variable, it is indeed differentiable, and the derivative in the point $t = 0$ exists and

$$g'(0) = 2a_1 + 12a_2^3.$$

Therefore, the partial derivative of f in the point (a_1, a_2) exists and equals

$$\frac{\partial f}{\partial x_2}(a) = (D_{e_2} f)_a = 2a_1 + 12a_2^3.$$

In general, there are manyy different expressions for the partial derivative of a function in some point a. Here are a few of them

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) = (D_{e_i} f)_a &= \frac{d}{dt} f(a + te_i)\bigg|_{t=0} \\ &= \frac{d}{dt} f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_d)\bigg|_{t=0} \\ &= \frac{d}{ds} f(a_1, \dots, a_{i-1}, s, a_{i+1}, \dots, a_d)\bigg|_{s=a_i}\end{aligned}$$

Proposition 15.3.3 Let $f : \Omega \to W$ be a function from an open domain Ω in $\mathbb{R}^d$ to a normed vector space $(W, \|\cdot\|_W)$. Let $a \in \Omega$.

The i-th partial derivative of f in the point a exists if and only if the function

$$t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_d)$$

is differentiable in the point a_i, and in this case

$$\frac{\partial f}{\partial x_i}(a) = \frac{d}{dt} f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_d)\Big|_{t=a_i}$$

Example 15.3.4 Consider again the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^4.$$

By the previous proposition, to determine whether the partial derivative

$$\frac{\partial f}{\partial x_2}(x_1, x_2)$$

exists in a point (x_1, x_2) and to determine its value, we just verify that

$$\frac{d}{dt} f(x_1, t)\Big|_{t=x_2} = \frac{d}{dt}(x_2^2 + 2x_1t + 3t^4)\Big|_{t=x_2} = 2x_1 + 12x_2^3.$$

We conclude as before that

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 2x_1 + 12x_2^3.$$

15.4 The Jacobian of a map

Proposition 15.4.1 Suppose $f : \Omega \to \mathbb{R}^m$ is a function defined on an open domain $\Omega \subseteq \mathbb{R}^d$, and suppose f is differentiable in a point $a \in \Omega$. Then the Jacobian matrix of f (with respect to the standard bases) is given by

$$[Df]_a := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_d}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_d}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_d}(a) \end{pmatrix}$$

In other words, for all $x \in \mathbb{R}^d$, it holds that

$$(Df)_a(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_d}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_d}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_d}(a) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

Proposition 15.4.2 Let $f : \Omega \to W$ with $\Omega \subseteq V$ open, and let $v_1, \dots, v_d$ be a basis of V with coordinate map Φ and let $w_1, \dots, w_m$ be a basis of W with coordinate map Ψ. Then the Jacobian of f with respect to these bases is given by

$$[Df]_a = [D\bar{f}]_{\Phi(a)}$$

where $\bar{f} := \Psi \circ f \circ \Phi^{-1}$ is the coordinate representation of f.

15.5 Linearization and tangent planes

Definition 15.5.1 **(Linearization)** Let $f : \Omega \to W$ be a differentiable function in a point $a \in \Omega$. Then the *linearization* of f is the function ${}_a : V \to W$ given by

$$L_a(x) = f(a) + (Df)_a(x - a).$$

Recall that the *graph* of a function $f : \Omega \to \mathbb{R}$ is the following subset of $\Omega \times \mathbb{R}$:

$$\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in \Omega\}$$

Definition 15.5.2 Let $f : \Omega \to \mathbb{R}$, where Ω is a subset of a normed vector space V. Assume f is differentiable in $a \in \Omega$. Then the *tangent plane to the graph of* f at a is the graph of the linearization L_a of f, i.e.

$$T_a := \{(v, {}_a(v)) \mid v \in V\}$$

Definition 15.5.3 Let $f : \Omega \to \mathbb{R}$ where Ω is s subset of a normed vector space V. Let $a \in \Omega$, and set $c := f(a)$. Assume f is differentiable in a with $(Df)_a \neq 0$. Then the *tangent plane to the level set*

$$f^{-1}(c) = \{x \in V \mid f(x) = c\}$$

at a is given by

$$\{x \in V \mid L_a(x) = c\}.$$

15.6 The gradient of a function

Definition 15.6.1 **(Gradient)** Let $f : \Omega \to \mathbb{R}$ be a function from an open domain Ω in $(\mathbb{R}^d, \|\cdot\|_2)$ to $\mathbb{R}, |\cdot|)$ and suppose f is differentiable in the point $a \in \Omega$. Then we call the vector

$$\boldsymbol{\nabla} f(a) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_d}(a) \end{pmatrix}$$

the *gradient* of f in the point a.

If a function $f : \Omega \to \mathbb{R}$ is differentiable in a point a, then the derivative $(Df)_a$ relates to the gradient as follows.

Proposition 15.6.2 Let $f : \Omega \to \mathbb{R}$ be a function from an open domain Ω in $(\mathbb{R}^d, \|\cdot\|_2)$ to $(\mathbb{R}, |\cdot|)$ and suppose f is differentiable in the point $a \in \Omega$. Then for all $v \in \mathbb{R}^d$, it holds that

$$(Df)_a(v) = (\boldsymbol{\nabla} f(a), v) = (\boldsymbol{\nabla} f(a))^\top v = \sum_{i=1}^{d} \frac{\partial f}{\partial x_1} v_i$$

where $(\cdot, \cdot)$ denotes the standard inner product on $\mathbb{R}^d$.

Proposition 15.6.3 Let $f : \Omega \to \mathbb{R}$ from an open domain Ω in $(\mathbb{R}^d, \|\cdot\|_2)$. Assume f is differentiable in a point $a \in \Omega$ with $(Df)_a \neq 0$. Set $c := f(a)$. Then the tangent plane to the level set $f^{-1}(c)$ at a is given by

$$a + \{x \in \mathbb{R}^d \mid (\boldsymbol{\nabla} f(a), x) = 0\}.$$

16 The Mean-Value Inequality

16.1 The mean-value inequality for functions defined on an interval

Theorem 16.1.1 (Mean-value inequality (v0)) Let $\gamma : [a, b] \to V$ be continuous on $[a, b]$ and differentiable on (a, b). Then

$$\|\gamma(b) - \gamma(a)\|_V \leq \sup_{t \in (a,b)} \|\gamma'(t)\|_V (b - a).$$

16.2 The mean-value inequality for functions defined on a domain

Recap: operator norm

Given a linear map $L : V \to W$, the norm $\|L\|_{V \to W}$ is the smallest constant $K \in \mathbb{R}$ such that for all $v \in V$,

$$\|L(v)\|_W \leq K\|v\|_V.$$

In other words,

1. for all $v \in V$, and every $L \in \mathrm{Lin}(V, W)$,
$$\|L(v)\|_W \leq \|L\|_{V \to W} \|v\|_V$$
2. for every $L \in \mathrm{Lin}(V, W)$ we have the following: if $K \in \mathbb{R}$ is a constant such that for all $v \in V$
$$\|L(v)\|_W \leq K\|V\|_V,$$
then
$$\|L\|_{V \to W} \leq K.$$

Proposition 16.2.1 (Mean-value inequality) Let $f : \Omega \to W$ be a differentiable function on Ω, and assume that for all $\tau \in [0, 1], p, q \in \Omega$, $\gamma(\tau) = (1 - \tau)p + \tau q \in \Omega$. Then

$$\|f(p) - f(q)\|_W = \sup_{\tau \in [0,1]} \left\|(Df)_{(1-\tau)p+\tau q}\right\|_{V \to W} \|p - q\|_V.$$

Proof. Note that $\gamma : [0, 1] \to V$ satisfies by assumption that for all $\tau \in [0, 1]$, $\gamma(\tau) \in \Omega$. So γ is a path in Ω.

Moreover, $\tau \mapsto (1 - \tau)p + \tau q$ is affine, therefore continuous and differentiable. So γ is continuous on $[0, 1]$ and differentiable on $(0, 1)$.

$$\gamma' = (q - p)$$

Define $g : [0, 1] \to W$, given by

$$g := f \circ \gamma.$$

Since f is differentiable, it is continuous. Since γ is continuous, g is continuous. By the chain rule g is differentiable on $(0, 1)$.

$$\begin{aligned}\|f(p) - f(q)\|_W &= \|f(\gamma(0)) - f(\gamma(1))\|_W = \|g(0) - g(1)\|_W \\ &\leq \sup_{\tau \in (0,1)} \|g'(\tau)\|_W (1 - 0)\end{aligned}$$

By chain rule

$$\begin{aligned}g'(\tau) &= (Df)_\tau(1) \\ &= (D(f \circ \gamma))_\tau(1) \\ &= (Df)_{g(\tau)}(D\gamma)_\tau(1) \\ &= (Df)_{g(\tau)}((D\gamma)_\tau(1)) \\ &= (Df)_{g(\tau)}(\gamma'(\tau)).\end{aligned}$$

Therefore,

$$\|f(p)-f(q)\|_W \leq \sup_{\tau\in(0,1)} \left\|(Df)_{\gamma(\tau)}(\gamma'(\tau)\right\|_W \quad = \sup_{\tau\in(0,1)} \|(Df)_\tau\|_W. = \sup_{\tau\in(0,1)} \left\|(Df)_{(1-\tau)p+\tau q}(q-p)\right\|_W$$

□

Lemma 16.2.2 Suppose $f:\Omega\to W$ is differentiable on Ω, and suppose its derivative $Df:\Omega\to \mathrm{Lin}(V,W)$ is bounded. Let $a\in\Omega$ and assume $r>0$ is such that $B(a,r)\subseteq\Omega$. Then for all $x\in B(a,r)$,

$$\left\|Err_a^f(x)\right\|_W \leq \sup_{z\in B(a,r)} \|(Df)_z-(Df)_a\|_{V\to W}\|x-a\|_V.$$

16.3 Continuous partial derivatives implies differentiability

Theorem 16.3.1 Let $f:\Omega\subseteq\mathbb{R}^d\to W$. Suppose for all $i,a\in\Omega$

$$\frac{\partial f}{\partial x_i}(a)$$

exists. Moreover, assume that $\frac{\partial f}{\partial x_i}:\Omega\to W$ are continuous. Then f is differentiable on Ω, and the derivative $(Df):\Omega\to\mathrm{Lin}(\mathbb{R}^d,W)$ is continuous.
We say that f is continuously differentiable on Ω.

$$(Df)(a)=(Df)_a$$

Proposition 16.3.2 Let $f:\Omega\to W$ be a function defined on some open set $\Omega\subseteq\mathbb{R}^d$ and let $a\in\Omega$. Assume that there exists a radius $r>0$ such that for all $x\in B(a,r)$, and for all $i\in\{1,\dots,d\}$, the partial derivative

$$\frac{\partial f}{\partial x_i}(x)$$

exists and the function

$$\frac{\partial f}{\partial x_i}:\Omega\to W$$

is continuous on $B(a,r)$.
Then the function f is continuously differentiable on $B(a,r)$.

17 Higher order derivatives

List of most important message for this chapter:

- That the $(n+1)$-th derivative is the derivative of the n-th derivative.
- The interpretation of the n-th derivative in terms of iterated directional derivatives.
- Concluding higher-order differentiability from continuity of higher order derivatives.
- The symmetry of n-th order derivatives.

17.1 Definition

Definition 17.1.1 We set $\mathrm{Lin}_1(V, W) := \mathrm{Lin}(V, W)$ and for every $n \in \mathbb{N}\{0\}$, we define $\mathrm{Lin}_{n+1}(V, W) := \mathrm{Lin}(V, \mathrm{Lin}_n(V, W))$.

Definition 17.1.2 **(Higher-order derivatives)** Let $n \in \mathbb{N}\{0, 1\}$. Suppose $f : \Omega \to W$ is n times differentiable on a ball $B(a, r) \subseteq \Omega$. We then say that f is $(n + 1)$ times differentiable in the point a if the function

$$D^n f : B(a, r) \to \mathrm{Lin}_n(V, W)$$

is differentiable in a. The $(n + 1)$-th derivative in the point a is then defined as

$$(D^{n+1} f)_a := (D(D^n f))_a \in \mathrm{Lin}_{n+1}(V, W).$$

17.2 Multilinear maps

Definition 17.2.1 **(Multilinear maps)** A map $L : V^n \to W$ is called multilinear, or n-linear, if for every $i \in \{1, \ldots, n\}$ and every $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$, the map

$$u \mapsto L(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_n)$$

is linear.
We will denote the vector space of n-linear maps from V^n to W by $\mathrm{MLin}(V^n, W)$.

Proposition 17.2.2 $\mathcal{J}_n$ is invertible with inverse $\mathcal{K}_n$, and it preserves the norm.

17.3 Relation to n-fold directional derivatives

Proposition 17.3.1 Suppose a function $f : \Omega \to W$ is n times differentiable in a point $a \in \Omega$. Then all directional n-fold derivatives exist in a and for all $v_1, \ldots, v_n \in V$,

$$(D^n f)_a(v_n, v_{n-1}, \ldots, v_2, v_1) = (D_{v_n}(D_{v_{n-1}} \ldots (D_{v_2}(D_{v_1} f)) \ldots))_a.$$

17.4 A criterion for higher differentiability

Theorem 17.4.1 Let $f : \Omega \to W$ where Ω is an open subset of $\mathbb{R}^d$. If all partial derivatives of f of order less than or equal to n exist, and if all partial derivatives of order n are continuous on Ω, then f is n timers differentiable on Ω.

17.5 Symmetry of second order derivatives

Lemma 17.5.1 Let $f : \Omega \to W$ be a function defined on an open domain $\Omega \subseteq V$. Let $a \in \Omega$ and assume that f is twice differentiable in a. Then for all $u, v \in V$,

$$(D^2 f)_a(u, v) = (D^2 f)_a(v, u).$$

17.6 Symmetry of higher-order derivatives

Definition 17.6.1 We denote by $\mathrm{Sym}_n(V, W)$ the collection of symmetric, n-linear maps from V^n to W. That is, a map $\mathcal{S} : V^n \to W$ is in $\mathrm{Sym}_n(V, W)$ if and only if it is linear in every argument and if for every permutation $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ it holds that

$$\mathcal{S}(v_1, \dots, v_n) = \mathcal{S}(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

In conclusion, we may as well write

$$D^n f : \Omega \to \mathrm{Sym}_n(V, W)$$

18 Polynomials and approximation by polynomials

18.1 Homogeneous polynomials

Definition 18.1.1 (multi-index) A d-dimensional multi-index α of order $k \in \mathbb{N}$ is a map $\{1, \dots, d\} \to \mathbb{N}$ such that

$$\alpha_1 + \alpha_2 + \cdots + \alpha_d = k$$

We write $|\alpha|$ for the order of a multi-index α.

If α is a multi-index, we use the notation

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_d^{\alpha_d}.$$

Similarly, for a function $f : \Omega \to W$ where $\Omega \subseteq \mathbb{R}^d$, we will use the notation

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \dots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d} f$$

We also define

$$\alpha! := \alpha_1! \alpha_2! \dots \alpha_d!$$

Note that

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} x^\alpha = \alpha!$$

Example 18.1.2

$$\frac{\partial^{14}}{\partial x_1^3 \partial x_2^7 \partial x_3^4} ((x_1)^3 (x_2^7)(x_3)^4) = 3!7!4!$$

Proposition 18.1.3 Every homogeneous polynomial $f : \mathbb{R}^d \to \mathbb{R}$ of degree n can be written as

$$\sum_{|\alpha|=n} \frac{1}{\alpha!} s_\alpha x^\alpha$$

for some coefficients $s_\alpha \in \mathbb{R}$. Moreover, the coefficients $s_a lpha$ are precisely determined by

$$s_\alpha = \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(0)$$

Lemma 18.1.4 Given a basis $v_1, \dots, v_d$ of the vector space V, there is a one-to-one correspondence between homogeneous polynomials of degree n and $\mathrm{Sym}_n(V, \mathbb{R})$. More precisely there is an invertible linear map $\mathcal{F}$ from $\mathrm{Sym}_n(V, \mathbb{R})$ to the vector space of homogeneous polynomials in d variables of degree n. With the linear map $\kappa : \mathbb{R}^d \to V$ defined as

$$\kappa(x) = x_1 v_1 + \cdots + x_d v_d$$

the n-linear symmetric map $\mathcal{S} \in \mathrm{Sym}_n(V, \mathbb{R})$ gets mapped to the homogeneous polynomial $F := \mathcal{F}(\mathcal{S}) : \mathbb{R}^d \to \mathbb{R}$ given by

$$\mathcal{F}(\mathcal{S})(x) = \frac{1}{n!} \mathcal{S}(\kappa(x), \dots, \kappa(x)).$$

Then the following equality holds for all $x \in \mathbb{R}^d$

$$\mathcal{F}(\mathcal{S})(x) = \frac{1}{n!} \mathcal{S}(\kappa(x), \dots, \kappa(x)) = \sum_{|\alpha|=n} \frac{1}{\alpha!} x^\alpha \mathcal{S}^{(\alpha)}$$

where

$$S^{(\alpha)} = \mathcal{S}(v_{i_1}, v_{i_2}, \dots, v_{i_n}$$

where $i_1, \dots, i_n \in \{1, \dots, d\}$ are such that v_1 appears α_1 times, v_2 appears α_2 times, and so on. In particular,

$$\mathcal{S}^{(\alpha)} = \frac{\partial^{|\alpha|} F}{\partial x^\alpha}(0).$$

In particular, an element of $\mathrm{Sym}_n(V, \mathbb{R})$ is completely determined by the values on the diagonal, i.e. if $\mathcal{S}, \mathcal{T} \in \mathrm{Sym}_n(V, \mathbb{R})$, then $\mathcal{S} = \mathcal{T}$ if and only if for all $v \in V$,

$$\mathcal{S}(v, \dots, v) = \mathcal{T}(v, \dots, v).$$

18.2 Taylor's theorem

Definition 18.2.1 **(Taylor expansion)** Let $f : \Omega \to W$ be n times differentiable in a point $a \in \Omega$. Then the function $T_{a,n} : V \to W$ given by $T_{a,n}(x) := f(a) + \sum_{k=1}^n \frac{1}{k!}(D^k f)_a(x - a, \dots, x - a)$ is called the Taylor expansion of f around a.

Theorem 18.2.2 Let $\Omega \subseteq V$ be open, let $a \in \Omega$ and suppose $f : \Omega \to W$ is n times differentiable in a point a. Then there exists a function $\mathrm{Err}_{a,n} : \Omega \to W$ such that

$$f(v) = f(a) + \sum_{k=1}^n \frac{1}{k!}(D^k f)_a(x - a, \dots, x - a) + \mathrm{Err}_{a,n}(v)$$

and such that

$$\lim_{v \to a} \frac{\|\mathrm{Err}_{a,n}(v)\|_W}{\|v - a\|_V^n} = 0$$

Proposition 18.2.3 Let $a \in V$, $k \in \mathbb{N}$, $\mathcal{S} \in \mathrm{Sym}_k(V, W)$ and consider the function $f : V \to W$ defined by

$$f(x) := \frac{1}{k!}\mathcal{S}(x - a, \dots, x - a).$$

Then

i for all $b \in V$,

$$(D^k f)_b = \mathcal{S}$$

ii for all $b \in V$ and all $j > k$,

$$(D^j f)_b = 0$$

iii for all $b \in V$, all $j < k$, all $u_1, \dots, u_j \in V$,

$$(D^j f)_b(u_1, \dots, u_j) = \frac{1}{(k - j)!}\mathcal{S}(u_1, \dots, u_j, b - a, \dots, b - a).$$

Proposition 18.2.4 Suppose $f : \Omega \to W$ and $g : \Omega \to W$ are both n times differentiable in $a \in \Omega$ and

$$\lim x \to a \frac{\|f(x) - g(x)\|_W}{\|x - a\|_V^n} = 0$$

Then for all $k = 0, \dots, n$

$$(D^k f)_a = (D^k g)_a$$

Proposition 18.2.5 Let $\Omega \subseteq \mathbb{R}^d$ be open, let $a \in \Omega$ and suppose $f : \Omega \to \mathbb{R}^m$ is n times differentiable in $a \in \Omega$. Then for all $k = 1, \dots, n$ and all $x \in \mathbb{R}^d$,

$$\frac{1}{k!}(D^k f)_a(x, \dots, x) = \sum_{|\alpha|=k} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) x^\alpha.$$

Example 18.2.6 Let $f : \mathbb{R}^2 \to \mathbb{R}$, $a \in \mathbb{R}^2$ and suppose we want to find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that for all $u \in \mathbb{R}^2$,

$$(D^3 f)_a(u, u, u) = 2(u_1)^2(u_2).$$

Note that it is necessary for the right-hand side to be a homogeneous polynomial of degree 3, otherwise such an f cannot be found. We call this homogeneous polynomial $q : \mathbb{R}^2 \to \mathbb{R}$.

By the earlier proposition, we know that

$$q(u) = \sum_{|\alpha|=} \frac{1}{\alpha!} s_\alpha u^\alpha,$$

where

$$s_a = \frac{\partial^3 q}{\partial x^\alpha}(0).$$

We know by the previous proposition that if such a function f exist that then for all $u \in \mathbb{R}^2$,

$$\sum_{|\alpha|=3} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) u^\alpha = \frac{1}{3!}(D^3 f)_a(u, u, u) = \frac{2}{3!}(u_1)^2(u_2).$$

If we compare the left-hand side and right-hand side this may suggest us to find a function such that

$$\frac{1}{3!} \frac{\partial^3 f}{(\partial x_1)^2 \partial x_2}(a) = \frac{2}{3!},$$

and all other partial derivatives in a vanish. Now the polynomial $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x) \mapsto \frac{1}{3}(x_1 - a_1)^2(x_2 - a_2)$ is such a polynomial.

Theorem 18.2.7 (Taylor's theorem in coordinates) Let $\Omega \subseteq \mathbb{R}^d$ be open, let $a \in \Omega$ and suppose $f : \Omega \to \mathbb{R}^m$ is n times differentiable in the point $a \in \Omega$. Then, defining the function $\mathrm{Err}_{a,n} : \Omega \to \mathbb{R}^m$ by

$$\mathrm{Err}_{a,n}(x) := f(x) - \left(f(a) + \sum_{1 \le |\alpha| \le n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a)(x - a)^\alpha \right)$$

we have that

$$\lim_{x \to a} \frac{\|\mathrm{Err}_{a,n}(x)\|_2}{\|x - a\|_2^n} = 0.$$

Definition 18.2.8 Let $\Omega \subseteq \mathbb{R}^d$ be open and $a \in \Omega$. Suppose $f : \Omega \to \mathbb{R}$ is n times differentiable in a. Then the polynomial

$$T_{a,n}(x) = f(a) + \sum_{1 \le |\alpha| \le n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a)(x - a)^\alpha$$

is called the n-th Taylor polynomial of f around the point a.

Corollary 18.2.9 (Taylor's theorem for functions of one variable) Let $\Omega \subseteq \mathbb{R}$ be a function such that f is n times differentiable in a point $a \in \Omega$. Then there exists a function $\mathrm{Err}_{a,n} : \Omega \to \mathbb{R}$ such that

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(a) \cdot (x - a)^k + \mathrm{Err}_{a,n}(x)$$

and such that

$$\lim_{x \to a} \frac{|\mathrm{Err}_{a,n}(x)|}{|x-a|^n} = 0.$$

Theorem 18.2.10 (Taylor's theorem with Lagrangge remainder) Let $f : \Omega \to \mathbb{R}$ be $(n+1)$ times differentiable on Ω, $a \in \Omega$. Then there exists a $\theta \subseteq (0,1)$ such that

$$f(x) = f(a) + \sum_{1 \le |\alpha| \le n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x-a)^\alpha \qquad + \frac{1}{(n+1)!}(D^(n_1)f)_{a+\theta(x-a)}(x-a, \dots, x-a).$$

18.3 Taylor approximations of standard functions

Corollary 18.3.1 For every $n \in \mathbb{N}$, it holds that

$$\begin{aligned}
\exp(x) &= \sum_{k=0}^{n} \frac{x^k}{k!} + O(|x|^{n+1}) \\
\sin(x) &= \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + O(|x|^{2n+3}) \\
\cos(x) &= \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} + O(|x|^{2n+2}) \\
\ln(1+x) &= \sum_{k=1}^{n} \frac{(-1)^{k-1} x^k}{k} + O(|x|^{n+1})
\end{aligned}$$

The notation

$$f(x) = g(x) + O(|x|^N)$$

should be read as that there exists a $C \geq 0$ and a $\delta > 0$ such that for all $x \in (-\delta, \delta)$,

$$|f(x) - g(x)| \leq C|x|^N.$$

19 Banach fixed point theorem

19.1 The Banach fixed point theorem

Definition 19.1.1 (Contraction) Let (X, dist_X) and (Y, dist_Y) be metric spaces. Let $D \subseteq X$ be a subset of X and let $f : D \to Y$ be a function. We say that f is a contraction if there exists a $\kappa \in [0, 1)$ such that for all $x, y \in D$, it holds that

$$\mathrm{dist}_Y(f(x), f(z)) \leq \kappa \mathrm{dist}_X(x, z).$$

If f satisfies the inequality for all x, z, and a constant $\kappa \in [0, 1)$, then we will also sometimes say that f is a κ-contraction.

Definition 19.1.2 (Banach fixed point theorem, metric space version) Let (X, dist_X) be a metric space, $D \subseteq X$ be a non-empty and complete subset of X, $f : D \to D$ be a function and $\kappa \in [0, 1)$. Assume that for all $x, z \in D$, it holds that

$$\mathrm{dist}_X(f(x), f(z)) \leq \kappa \mathrm{dist}_X(x, z).$$

Then there exists a unique point $p \in D$ such that $f(p) = p$.
Moreover, for all $q \in D$, if we define the sequence $(x^{(n)})_n$ by

$$\begin{aligned} x^{(0)} &= q, \\ x^{(n+1)} &= f(x^{(n)}), \end{aligned}$$

then the sequence $(x^{(n)})_n$ converges to p and for all $n \in \mathbb{N}$,

$$\mathrm{dist}_X(x^{(n)}, p) \leq \frac{\kappa^n}{1-\kappa} \mathrm{dist}_X(q, x^{(0)}).$$

Theorem 19.1.3 (Banach fixed point theorem, $\mathbb{R}^d$ version) Let $\subseteq \mathbb{R}^d$ be a closed non-empty set, $f : D \to D$ be a function and $\kappa \in [0, 1)$. Assume that for all $x, z \in D$, ti holds that

$$\|f(x) - f(z)\|_2 \leq \kappa \|x - z\|_2.$$

Then there exists a unique point $p \in D$ such that $f(p) = p$. Moreover, for all $q \in D$ we define the sequence $(X^{(n)})_n$ by

$$\begin{aligned} x^{(0)} &= q, \\ x^{(n+1)} &= f(x^{(n)}), \qquad \text{for } n \in \mathbb{N}. \end{aligned}$$

Then the sequence $(x^{(n)})_n$ converges to p and for all $n \in \mathbb{N}$,

$$\left\| x^{(n)} - p \right\|_2 \leq \frac{\kappa^n}{1-\kappa} \left\| x^{(1)} - x^{(0)} \right\|_2.$$

20 Implicit function theorem

20.1 The objective

We will be considering continuously differential functions $f : \Omega \subset \mathbb{R}^{d+m} \to \mathbb{R}^m$. It is good to think of the vector space $\mathbb{R}^{d+m}$ as the vector space $\mathbb{R}^d \bigoplus \mathbb{R}^m$, i.e. as the vector space of pairs (x, y) of vectors $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$.

20.2 Notation

Consider a function $f : \mathbb{R}^{d+m} \to \mathbb{R}^m$. Then for a point $(a, b) \in \mathbb{R}^{d+m}$

$$[Df]_{(a,b)} = \left(\begin{array}{ccc|ccc} \frac{\partial f_1}{\partial x_1}((a,b)) & \dots & \frac{\partial f_d}{\partial x_1}((a,b)) & \frac{\partial f_{d+1}}{\partial x_1}((a,b)) & \dots & \frac{\partial f_{d+m}}{\partial x_1}((a,b)) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_m}((a,b)) & \dots & \frac{\partial f_d}{\partial x_m}((a,b)) & \frac{\partial f_{d+1}}{\partial x_m}((a,b)) & \dots & \frac{\partial f_{d+m}}{\partial x_m}((a,b)) \end{array} \right)$$

We will denote by

$$(D_1 f)_{(a,b)} : \mathbb{R}^d \to \mathbb{R}^m$$

the restriction of the derivative $(Df)_{(a,b)} : \mathbb{R}^{d+m} \to \mathbb{R}^m$ to the subspace $\mathbb{R}^d \subset \mathbb{R}^{d+m}$. In other words, for all $h \in \mathbb{R}^d$,

$$(D_1 f)_{(a,b)}(h) = (Df)_{(a,b)}((h, 0)).$$

Similarly, we will denote by

$$(D_2 f)_{a,b} : \mathbb{R}^m \to \mathbb{R}^m$$

the restriction of the derivative $(Df)_{(a,b)} : \mathbb{R}^{d+m} \to \mathbb{R}^m$ to the subspace $\mathbb{R}^m \subset \mathbb{R}^{d+m}$. In other words, for all $k \in \mathbb{R}^m$.

$$(D_2 f)_{(a,b)}(k) = (Df)_{(a,b)}((0, k)).$$

By linearity of the derivative, we have that for all $h \in \mathbb{R}^d$ and $k \in \mathbb{R}^m$,

$$(Df)_{(a,b)}((h, k)) = (D_1 f)_{(a,b)}(h) + (D_2 f)_{(a,b)}(k).$$

We will denote the matrix representations of the maps $(D_1 f)_{(a,b)}$ and $(D_2 f)_{(a,b)}$ by $[D_1 f]_{(a,b)}$ and $[D_2 f]_{(a,b)}$ respectively.
So

$$[D_1 f]_{(a,b)} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}((a,b)) & \dots & \frac{\partial f_d}{\partial x_1}((a,b)) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_m}((a,b)) & \dots & \frac{\partial f_d}{\partial x_m}((a,b)) \end{pmatrix}$$

and

$$[D_2 f]_{(a,b)} = \begin{pmatrix} \frac{\partial f_{d+1}}{\partial x_1}((a,b)) & \dots & \frac{\partial f_{d+m}}{\partial x_1}((a,b)) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{d+1}}{\partial x_m}((a,b)) & \dots & \frac{\partial f_{d+m}}{\partial x_m}((a,b)) \end{pmatrix}$$

Then

$$\begin{aligned} (Df)_{(a,b)}((h, k)) &= (D_1 f)_{(a,b)}(h) + (D_2 f)_{(a,b)}(k) \\ &= [D_1 f]_{(a,b)} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} + [D_2 f]_{(a,b)} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}. \end{aligned}$$

20.3 The implicit function theorem

Theorem 20.3.1 (Implicit function theorem) Let $\Omega \subset \mathbb{R}^{d+m}$, $f : \Omega \to \mathbb{R}^m$ be a continuously differentiable function and let $(a, b) \in \Omega$ be a point such that $f(a, b) = 0$ and the matrix $[D_2 f]_{(a,b)}$ is invertible. Then there exists an $r_1 > 0$ and an $r_2 > 0$, and a continuously differentiable function $g : B(a, r_1) \to \mathbb{R}^m$ such that for all $x \in B(a, r_1)$ and all $y \in B(b, r_2)$,

$$f(x, y) = 0 \qquad \text{if and only if} \qquad y = g(x).$$

Moreover, for all $x \in B(a, r_1)$,

$$(Dg)_x = -(D_2 f)^{-1}_{(x,g(x))} \circ (D_1 f)_{(x,g(x))}.$$

20.4 The inverse function theorem

Theorem 20.4.1 (Inverse function theorem) Let $\Omega \subset \mathbb{R}^m$ be open and let $h : \Omega \to \mathbb{R}^m$ be a function which is continuously differentiable. Suppose $b \in \Omega$ and suppose that $(Dh)_b$ is non-singular. Then there exists an $r_1 > 0$ and an $r_2 > 0$, and a continuously differentiable function $g : B(h(b), r_1) \to \mathbb{R}^m$ such that for all $x \in B(h(b), r_1)$ and all $y \in B(b, r_2)$,

$$x = h(y) \qquad \text{if and only if} \qquad y = g(x).$$

Moreover, for all $x \in B(h(b), r_1)$,

$$(Dg)_x = (Dh)^{-1}_{g(x)}.$$

In particular, for $r_3 > 0$ small enough, the function h restricted to $B(b, r_3)$ mapping to $h(B(b, r_3))$ is invertible with continuously differentiable inverse g.

21 Function sequences

21.1 Point-wise convergence

Definition 21.1.1 (Pointwise convergence) Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$. We say that a sequence of functions $f : \mathbb{N} \to (D \to Y)$ converges pointwise to a function $f^* : D \to Y$ if

$$\begin{aligned}&\text{for all } x \in D,\\ &\quad \lim_{n\to\infty} f_n(x) = f^*(x).\end{aligned}$$

Example 21.1.2 We consider a case in which $(X, \text{dist}_X) = (\mathbb{R}, \text{dist}_\mathbb{R})$ and $D = [0,] \subset \mathbb{R}$. We consider the sequence of functions $f : \mathbb{N} \to ([0, 1] \to \mathbb{R})$ defined by

$$f_n(x) = x^n.$$

Then the sequence (f_n) converges pointwise to the function $f^* : [0, 1] \to \mathbb{R}$ defined by

$$f^*(x) = \begin{cases} 0 & \text{if } x \in [0, 1) \\ 1 & \text{if } x = 1 \end{cases}.$$

To show this, let $x \in [0, 1]$. Then we consider two cases.
In case $x \in [0, 1)$, then

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0 = f^*(x)$$

In case $x = 1$, then

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} 1^n = 1 = f^*(x).$$

21.2 Uniform convergence

Definition 21.2.1 Let (X, dist_X) and (Y, dist_Y) be two metric spaces and let $D \subseteq X$. We say that a sequence of functions $f : \mathbb{N} \to (D \to Y)$ converges uniformly to a function $f^* : D \to Y$ if

$$\begin{aligned}&\text{for all } \epsilon > 0,\\ &\quad \text{there exists } N \in \mathbb{N} \text{ such that}\\ &\quad\quad \text{for all } n \geq N,\\ &\quad\quad\quad \text{for all } x \in D,\\ &\quad\quad\quad\quad \text{dist}_Y(f_n(x), f^*(x)) < \epsilon.\end{aligned}$$

Proposition 21.2.2 Let (X, dist_X) and (Y, dist_Y) be two metric spaces, let $D \subseteq X$, and assume that a sequence of functions $f : \mathbb{N} \to (D \to Y)$ converges uniformly to a function $f^* : D \to Y$. Then the sequence of functions f converges pointwise to f^*.

Corollary 21.2.3 Suppose a sequence of functions $f : \mathbb{N} \to (D \to Y)$ converges *pointwise* to a function $f^* : D \to Y$. Then (f_n) converges uniformly on D if and only if (f_n) converges uniformly to f^* on D.

21.3 Preservation of continuity under uniform convergence

Theorem 21.3.1 Let (f_n) be a sequence of continuous functions from a domain D in the metric space (X, dist_X) to the metric space (Y, dist_Y) that converges uniformly to a function $g : D \to Y$. Then the function g is also continuous on D.

Example 21.3.2 Consider the sequence of functions (f_n) from $[0, 1]$ to $\mathbb{R}$ defined by

$$f_n(x) = x^n.$$

We have seen that the pointwise limit is $g : [0,1] \to \mathbb{R}$ given by

$$g(x) := \begin{cases} 0, & \text{if } x \\ in[0,1) \\ 1, & \text{if } x = 1. \end{cases}$$

Because the function g is not continuous, but for every $n \in \mathbb{N}$ the function f_n is continuous, we converge that the sequence f_n does not converge uniformly to g.

21.4 Differentiability theorem

Theorem 21.4.1 Let (f_n) be a sequence of functions from an open domain Ω in a vector space V to $\mathbb{R}$ and suppose the sequence converges pointwise to a function $g : \Omega \to \mathbb{R}$. Suppose moreover that the functions f_n are continuously differentiable on Ω and suppose the sequence of functions $Df_n : \Omega \to \mathrm{Lin}(V, \mathbb{R})$ converges uniformly to a function $\Delta : \Omega \to \mathrm{Lin}(V, \mathbb{R})$. Then the function g is differentiable on Ω as well and

$$Dg = \Delta$$

21.5 The normed vector space of bounded functions

Definition 21.5.1 Let D be a set. The normed vector space $(\mathcal{B}(D), \|\cdot\|_\infty)$ is defined as the vector space of *bounded* functions D to $\mathbb{R}$ with norm $\|\cdot\|_\infty$ given by

$$\|f\|_\infty = \sup_{x \in D} |f(x)|.$$

Proposition 21.5.2 Let (f_n) be a sequence of functions from D to $\mathbb{R}$ and let f be a function. Then the sequence (f_n) converges uniformly to f if and only if there exists an $N \in \mathbb{N}$ such that for every $n \geq N$, the function $(f_n - f)$ is bounded, and such that the sequence $n \mapsto (f_{N+n} - f)$ converges to 0 in $\mathcal{B}(D)$.

Proposition 21.5.3 Let (f_n) be a sequence of functions from a domain D in the metric space (X, dist_X) to the metric space (Y, dist_Y) and let $g : D \to Y$ be a function. Then (f_n) converges to g uniformly if and only if there exists an $N \in \mathbb{N}$ such that for every $n \geq N$, the function h_n given by

$$h_n(x) := \mathrm{dist}_Y(f_n(x), g(x))$$

is bounded and the sequence $n \mapsto h_{N+n}$ converges to 0 in $\mathcal{B}(D)$.

22 Power series

Definition 22.0.1 **(Power series)** A *power series* at a point $c \in \mathbb{R}$ is a function series of the form

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

where $a : \mathbb{N} \to \mathbb{R}$ is a real-valued sequence.

22.1 Convergence of power series

Lemma 22.1.1 Suppose a power series

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

converges at a point $z \in \mathbb{R}$. Let $\delta > 0$ be such that $\delta < |z-c|$. Then the power series converges absolutely and uniformly on the interval $[c-\delta, c+\delta]$.

Corollary 22.1.2 For every power series

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

around a point $c \in \mathbb{R}$ exactly one of the following occurs:

i The series converges for $x = c$ and diverges for $x \neq c$. In this case we say that the radius of convergence of the power series is 0.

ii There exists an $R > 0$ such that for all $x \in (c-R, c+R)$ the power series converges and for all $x \in \mathbb{R} \setminus [c-R, c+R]$ the series diverges. In this case we say the radius of convergence equals R.

iii The series converges for all $x \in \mathbb{R}$. In this case we say the radius of convergence is ∞.

Proposition 22.1.3 Let

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

be a power series around c and define the (extended real) number

$$L := \limsup_{k \to \infty} \sqrt[k]{|a_k|}$$

Then

i If $L = 0$, then the radius of convergence is ∞.

ii If $L = \infty$, then the radius of convergence is 0.

iii If $L \in (0, \infty)$, then the radius of convergence is $1/L$.

Theorem 22.1.4 **(Root test, $\limsup$ version)** Let (b_k) be a sequence of non-negative real numbers.

1. If

$$\limsup_{k \to \infty} \sqrt[k]{b_k} < 1,$$

then the series $\sum_{k=0}^{\infty} b_k$ converges.

2. If
$$\limsup_{k\to\infty} \sqrt[k]{b_k} > 1,$$
then the series $\sum_{k=0}^{\infty} b_k$ diverges.

22.2 Standard functions defined as power series

Proposition 22.2.1 The power series
$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k$$
has a radius of convergence $R = \infty$.

Definition 22.2.2 The function $\exp : \mathbb{R} \to \mathbb{R}$ is defined as the power series
$$\exp(x) := \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$
It has radius of convergence $R = \infty$.

Definition 22.2.3 The function $\sin : \mathbb{R} \to \mathbb{R}$ is defined as the power series
$$\sin(x) := \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k+1)!} x^{2k+1}.$$
It has a radius of convergence $R = \infty$.

Definition 22.2.4 The function $\cos : \mathbb{R} \to \mathbb{R}$ is defined as the power series
$$\cos(x) := \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k)!} x^{2k}.$$
It has a radius of convergence $R = \infty$.

22.3 Operations with power series

Proposition 22.3.1 **(Sums of power series)** Let
$$\sum_{k=0}^{\infty} a_k (x-c)^k \quad \text{and} \quad \sum_{k=0}^{\infty} b_k (x-c)^k$$
be two power series around c, with radii of convergence R_1 and R_2 respectively. The sum of these functions is the power series
$$\sum_{k=0}^{\infty} (a_k + b_k)(x-c)^k$$
and the radius of convergence R for this new power series satisfies
$$R \geq \min(R_1, R_2).$$

Proposition 22.3.2 (Products of power series) Let
$$\sum_{k=0}^{\infty} a_k (x-c)^k \quad \text{and} \quad \sum_{k=0}^{\infty} b_k (x-c)^k$$

be two power series around c, with radii of convergence R_1 and R_2 respectively. The product of these functions is the power series

$$\sum_{k=0}^{\infty} c_k (x-c)^k$$

where

$$c_k = \sum_{\ell=0}^{k} a_{\ell_{k-\ell}}.$$

The radius of convergence R for this new power series satisfies

$$R \geq \min(R_1, R_2).$$

22.4 Differentiation of power series

Proposition 22.4.1 Let

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

be a power series with radius of convergence R. Then, the power series is differentiable on the interval $(c-R, c+R)$ and its derivative is the power series

$$\sum_{\ell=0}^{\infty} (\ell+1) a_{\ell+1} (x-c)^{\ell}$$

which has the same radius of convergence R.

Corollary 22.4.2 Let

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

be a power series with radius of convergence R. Then the power series is infinitely many times differentiable on the interval $(c-R, c+R)$, and for every $\ell \in \mathbb{N}$, the ℓ-th derivative has the same radius of convergence.

Theorem 22.4.3 **(Identification of coefficients)** Let $R > 0$ and let $f : (c-R, c+R) \to \mathbb{R}$ be given by a power series

$$f(x) := \sum_{k=0}^{\infty} a_k (x-c)^k.$$

Then for all $k \in \mathbb{N}$,

$$a_k = \frac{f^{(k)}(c)}{k!}.$$

Theorem 22.4.4 **(Identity theorem for power series)** Let $R > 0$ and let $f, g : (c-R, c+R) \to \mathbb{R}$ be given by power series

$$f(x) = \sum_{k=0}^{\infty} a_k (x-c)^k \quad \text{and} \quad g(x) = \sum_{k=0}^{\infty} b_k (x-c)^k$$

and assume for all $x \in (c-R, c+R)$,

$$f(x) = g(x).$$

Then for all $k \in \mathbb{N}$,

$$a_k = b_k$$

22.5 Taylor series

Definition 22.5.1 **(Taylor series)** Let $f : (c - R, c + R) \to \mathbb{R}$ be a function that is infinitely many times differentiable. The Taylor series of f around c is the power series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!} (x - c)^k.$$

23 Riemann integration in one dimension

Main message:

- Every continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.
- The fundamental theorem of calculus, mainly the part that when $F : [a, b] \to \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ satisfy that $F' = f$ and f is bounded and Riemann integrable, then

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

23.1 Riemann integrable functions and the Riemann integral

Definition 23.1.1 **(Partition)** A *partition* P of an interval $[a, b]$ (with n intervals) is a subset $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ such that $a = x_0 < x_1 < \dots < x_n = b$

Definition 23.1.2 **(Upper/Lower sum)** Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $P = (x_0, x_1, \dots, x_n)$ be a partition of $[a, b]$. Then the upper sum of f with respect to P is defined as

$$U(P, f) := \sum_{k=1}^{n} M_k \Sigma x_k$$

where $\Sigma x_k := (x_k - x_{k-1}$ and

$$M_k := \sup_{x \in [x_{k-1}, x_k]} f(x).$$

Similarly, we define the lower sup of f with respect to P as

$$L(P, f) := \sum_{k=1}^{n} m_k \Sigma x_k$$

where

$$m_k := \inf_{x \in [x_{k-1}, x_k]} f(x).$$

Definition 23.1.3 **(Refinement)** Let $\tilde{P}$ be a partition of $[a, b] \subseteq \mathbb{R}$. A partition P is called a *refinement* of $\tilde{P}$ if $\tilde{P} \subset P$.
If $\tilde{P}, \tilde{Q}$ are two partitions of $[a, b]$, then a partition P is called a *common refinement* of $\tilde{P}$ and $\tilde{Q}$ if P is both a refinement of $\tilde{P}$ and $\tilde{Q}$.

Note that two partitions P, Q always have a common refinement, namely $P \cup Q$.

Proposition 23.1.4 For every bounded $f : [a, b] \to \mathbb{R}$ and every partition P of $[a, b]$ we have

$$L(P, f) \leq U(P, f).$$

Definition 23.1.5 **(Darboux integral)** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We define the upper Darboux integral of f as

$$\overline{\int_a^b} f dx := \inf\{U(P, f) \mid P \text{ partition of } [a, b]\}$$

and the lower Darboux integral of f as

$$\underline{\int_a^b} f dx := \sup\{L(P, f) \mid P \text{ partition of } [a, b]\}.$$

Proposition 23.1.6 Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then

$$\underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx.$$

Definition 23.1.7 (Riemann integral) Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We say that f is *Riemann integrable* if

$$\underline{\int_a^b} f dx = \overline{\int_a^b} f dx.$$

In this case we say that the *Riemann integral* of f over $[a, b]$ is

$$\int_a^b f(x)dx := \underline{\int_a^b} f dx = \overline{\int_a^b} f dx.$$

Proposition 23.1.8 (Alternative characterization of Riemann integrability) Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if

for every $\epsilon > 0$,
there exists a partition P of $[a, b]$,
$U(P, f) - L(P, f) < \epsilon$

Definition 23.1.9 We denote the set of bounded, Riemann-integrable functions $f : [a, b] \to \mathbb{R}$ by $\mathcal{R}[a, b]$.

23.2 Sums, products of Riemann integrable functions

Proposition 23.2.1 (R[a,b) is a vector space] Let $f, g \in \mathcal{R}[a, b]$ and let $\lambda \in \mathbb{R}$. Then

1. The function $f + g \in \mathcal{R}[a, b]$ and

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

2. The function $\lambda f \in \mathcal{R}[a, b]$ and

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx$$

Definition 23.2.2 If $f : [a, b] \to \mathbb{R}$ is bounded and Riemann integrable on $[a, b]$, then we define

$$\int_b^a f(x)dx := -\int_a^b f(x)dx.$$

Proposition 23.2.3 (Further properties) Let $f, g[a, b] \to \mathbb{R}$ be two bounded and Riemann integrable functions.

i. We have

$$\int_a^b 1dx = b - a.$$

ii. If for all $x \in [a, b]$, we have $f(x) \leq g(x)$, then (monotonicity)

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

iii. If $z \in (a,b)$, then f is integrable on $[a,z]$ and $[z,b]$ and (restriction)

$$\int_a^b f(x)dx = \int_a^z f(x)dx + \int_z^b f(x)dx.$$

iv. We have (triangle inequality)

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx.$$

v. The function $f \cdot g$ is Riemann integrable on $[a,b]$.

23.3 Continuous functions are Riemann integrable

Proposition 23.3.1 Let $f : [a,b] \to \mathbb{R}$ be continuous. Then f is Riemann integrable.

23.4 The fundamental theorem of calculus

Theorem 23.4.1 (Fundamental theorem of calculus) i. Let $f : [a,b] \to \mathbb{R}$ be continuous. Then the function

$$F : [a,b] \to \mathbb{R}$$

given by

$$F(x) := \int_a^x f(s)ds$$

is differentiable on (a,b) and for all $x \in (a,b)$

$$F'(x) = f(x).$$

ii. Let $F : [a,b] \to \mathbb{R}$ be an anti-derivative of a function $f : [a,b] \to \mathbb{R}$, i.e. for all $x \in (a,b)$,

$$F'(x) = f(x)$$

and suppose that f is bounded and Riemann integrable on $[a,b]$. Then

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

24 Riemann integration in multiple dimensions

24.1 Partitions in multiple dimensions

Definition 24.1.1 **(Closed rectangle)** By a *closed rectangle* in $\mathbb{R}^d$ we mean a set R of the form

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d].$$

Definition 24.1.2 **(Partition of a rectangle)** Let $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ be a closed rectangle in $\mathbb{R}^d$. By a partition Q of R we mean a Cartesian product

$$Q = P^1 \times \cdots \times P^d$$

where for $i \in \{1, \ldots, d\}$, the partition $P^i = \{x^i_1, \ldots, x^i_{n_i}\}$ is a partition of $[a_i, b_i]$.

24.2 Riemann integral on rectangles in $\mathbb{R}^d$

Definition 24.2.1 Let $R \subset \mathbb{R}^d$ be a closed rectangle,

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d],$$

$f : R \to \mathbb{R}$ be a bounded function and

$$Q = P^1 \times \cdots \times P^d$$

be a partition of R.
Then the *upper sum* of f with respect to Q is defined as

$$U(Q, f) := \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} M_{k_1,\ldots,k_d} \Delta x^1_{k_1} \ldots \Delta x^d_{k_d}$$

where where $\Delta x^i_{k_i} := (x^i_{k_i} - x^i_{k_i-1})$ and

$$M_{k_1,\ldots,k_d} := \sup_{x \in R_{k_1,\ldots,k_d}} f(x).$$

Similarly, we define the *lower sum* of f with respect to Q as

$$L(Q, f) := \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} m_{k_1,\ldots,k_d} \Delta x^1_{k_1} \ldots \Delta x^d_{k_d}$$

where

$$m_{k_1,\ldots,k_d} := \inf_{x \in R_{k_1,\ldots,k_d}} f(x).$$

Definition 24.2.2 Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and $f : R \to \mathbb{R}$ be a bounded function. We define the upper Darboux integral of f as

$$\overline{\int_R} f d\mathbf{x} := \inf\{U(P, f) \mid P \text{ partition of } R\}$$

and the lower Darboux integral of f as

$$\underline{\int_R} f d\mathbf{x} := \sup\{L(P, f) \mid P \text{ partition of } R\}.$$

Definition 24.2.3 (Riemann integral in multiple dimensions) Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and $f : R \to \mathbb{R}$ be a bounded function. We say that f is *Riemann integrable* if

$$\underline{\int_R} f d\mathbf{x} = \overline{\int_R} f d\mathbf{x}.$$

In this case we say that the *Riemann integral* of f over R is

$$\int_R f d\mathbf{x} := \underline{\int_R} f d\mathbf{x} = \overline{\int_R} f d\mathbf{x}.$$

$$\forall \epsilon > 0 \exists P \text{ partition of } R : U(P, f) - L(P, f) < \epsilon.$$

Proposition 24.2.4 (Alternative characterization of Riemann integrability) Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and $f : R \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if

for all $\epsilon > 0$,
there exists a partition P of R,
$U(P, f) - L(P, f) < \epsilon$

24.3 Properties of the multidimensional Riemann integral

Proposition 24.3.1 Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and $f, g : R \to \mathbb{R}$ be bounded and Riemann integrable on R. Then

1. (volume)

$$\int_R 1 d\mathbf{x} = (b_1 - a_1) \cdot (b_2 - a_2) \dots (b_d - a_d) =: \mathrm{Vol}(R).$$

2. If for all $x \in R, f(x) \leq g(x)$, then (monotonicity)

$$\int_R f(x) dx \leq \int_R g(x) dx.$$

3. The function $|f|$ is Riemann integrable on R and (triangle inequality)

$$\left| \int_R f(x) f x \right| \leq \int_R |f(x)| dx.$$

4. If Q is a closed rectangle contained in R, then f is integrable on Q. Moreover, if $Q_1, \dots, Q_N$ are finitely many closed rectanlges,

   - if their interiors are disjoint, i.e. $\mathrm{int} Q_i \cap \mathrm{int} Q_j = \emptyset$ if $i \neq j$ and
   - if the union of Q_i's equals R, i.e.
   $$\bigcup_{i=1}^N Q_i = R,$$

   then
   $$\int_R f(x) dx = \sum_{i=1}^N \int_{Q_i} f(x) dx.$$

24.4 Continuous functions are Riemann integrable

Theorem 24.4.1 Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and $f : R \to \mathbb{R}$ be a continuous function. Then f is bounded and Riemann integrable on R.

24.5 Fubini's theorem

Theorem 24.5.1 **(Fubini)** Let $R = A \times B$ be a rectangle in $\mathbb{R}^{d+m}$. Let $f : R \to \mathbb{R}$ be bounded and Riemann integrable on R, and suppose for every $x \in \mathbb{R}^d$ the function $h_x : \mathbb{R}^m \to \mathbb{R}$ defined by

$$h_x(y) := f(x, y)$$

is Riemann integrable. Then the function $F : \mathbb{R}^d \to \mathbb{R}$ given by

$$F(x) := \int_B f(x, y)dy$$

is Riemann integrable and

$$\int_R f(z)dz = \int_A \left(\int_B f(x, y)dy \right) dx.$$

24.6 The (topological) boundary of a set

Definition 24.6.1 **(Topological boundary)** Let E be a subset of the normed vector space $(\mathbb{R}^d, \|\cdot\|)$. The boundary of E is defined as

$$\partial E := \mathbb{R}^d \setminus ((\mathrm{int} E) \cup (\int (\mathbb{R}^d \setminus E))).$$

24.7 Jordan content

Definition 24.7.1 **(Volume of a rectangle)** Let $R \subseteq \mathbb{R}^d$ be a closed rectangle. The volume of R is defined as

$$\mathrm{Vol}(R) := (b_1 - a_1) \cdot (b_2 - a_2) \ldots (b_d - a_d).$$

Definition 24.7.2 **(Cube)** We say that a closed rectangle $R \subseteq \mathbb{R}^d$ is a *cube* if all sides have the same length, i.e. if $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ then for all $i, j \in \{1, \ldots, d\}$,

$$b_i - a_i = b_j - a_j$$

Definition 24.7.3 **(Jordan content zero)** We say that a subset $S \subseteq \mathbb{R}^d$ has *Jordan content zero* if

for all $\epsilon > 0$,
there exists $N \in \mathbb{N}$,
there exists rectangles $R_1, \ldots, R_N$,

$$S \subseteq \bigcup_{i=1}^{N} R_i \quad \text{and} \quad \sum_{i=1}^{N} \mathrm{Vol}(R_i) \leq \epsilon.$$

Lemma 24.7.4 Suppose a set $S \subseteq \mathbb{R}^d$ has Jordan content zero. Then for all $\epsilon > 0$, there exists an $M \in \mathbb{N}$ and cubes $Q_1, \ldots, Q_M$ such that

$$S \subseteq \bigcup_{i=1}^{M} Q_i \quad \text{and} \quad \sum_{i=1}^{M} \mathrm{Vol}(Q_i) \leq \epsilon.$$

Proposition 24.7.5 Let $S \subseteq \mathbb{R}^d$ be a subset with Jordan content zero and let $F : S \to \mathbb{R}^d$ be Lipschitz. Then $F(S)$ has Jordan content zero.

Proposition 24.7.6 Let E be a bounded subset on $\mathbb{R}^d$, and let $F : E \to \mathbb{R}^{d+m}$ be Lipschitz where $m \geq 1$. Then $F(E)$, as a subset of $\mathbb{R}^{d+m}$, has Jordan content zero.

24.8 Integration over general domains

Definition 24.8.1 **(Integration over bounded subsets)** Let E be a bounded subset of $\mathbb{R}^d$. We say that a function $f : E \to \mathbb{R}$ is integrable on E if, with some rectangle R in $\mathbb{R}^d$ containing E, the function $f_E : R \to \mathbb{R}$ defined by

$$f_E(x) := \begin{cases} f(x) & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

is integrable on R. Moreover, we define

$$\int_E f(x)dx := \int_R f_E(x)dx.$$

Definition 24.8.2 **(Jordan set)** Let $E \subseteq \mathbb{R}^d$. We say that E is a *Jordan set* if the topological boundary ∂E of E has Jordan content zero.

Proposition 24.8.3 Let $R \subseteq \mathbb{R}^d$ be a closed rectangle and assume that $E \subseteq R$ is a bounded subset of $\mathbb{R}^d$ and assume E is a Jordan set. Let f be a bounded and Riemann integrable function on R. Then f is integrable on E.

24.9 The volume of bounded sets

Definition 24.9.1 **(Characterisistic function of a set)** Let $E \subseteq \mathbb{R}^d$. The characteristic function of E is the function $\mathbb{1}_E : \mathbb{R}^d \to \mathbb{R}$ given by

$$\mathbb{1}_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 \text{ if } x \notin E & \end{cases}.$$

Definition 24.9.2 **(Volume of a set)** Let E be a bounded set such that the characteristic function $\mathbb{1}_E : \mathbb{R}^d \to \mathbb{R}$ is Riemann integrable. Then the *volume* of E is defined as

$$\text{Vol}(E) := \int_E 1dx.$$

25 Change-of-variables Theorem

25.1 Change-of-variables Theorem

Theorem 25.1.1 (Change-of-variable theorem) Let $\Omega \subset \mathbb{R}^d$ be open, $E \subset \Omega$ be a Jordan set such that also its closure $\bar{E} \subset \Omega$, $\Phi : \Omega \to \mathbb{R}^d$ be continuously differentiable and injective. Asssume that also the inverse function Φ^{-1} is differentiable.
Assume that $f : \Phi(E) \to \mathbb{R}$ is integrable on $\Phi(E)$ and assume that $f \circ \Phi$ is integrable on E.
Then

$$\int_{\Phi(E)} f(x)dx = \int_E f(\Phi(y))|\det([D\Phi]_y)|dy.$$

25.2 Polar coordinates

The transformation is given by the function

$$\Phi_{pol} : (0, \infty) \times (0, 2\pi) \to \mathbb{R}^2$$

defined by

$$\Phi_{pol}(r, \phi) := (r\cos\phi, r\sin\phi)$$

Here,

$$\det[D\Phi_{pol}]_{(r,\phi)} = r$$

25.3 Cylindrical coordinates

The transformation is given by the function

$$\Phi_{cyl} : (0, \infty) \times (0, 2\pi) \times \mathbb{R} \to \mathbb{R}^3$$

defined by

$$\Phi_{cyl}(r, \phi, z) := (r\cos\phi, r\sin\phi, z)$$

Here,

$$\det[D\Phi_{cyl}]_{(r,\phi,z)} = r$$

25.4 Spherical coordinates

The transformation is given by the function

$$\Phi_{sph} : (0, \infty) \times (0, 2\pi) \times (0, \pi) \to \mathbb{R}^3$$

given by

$$\Phi_{sph}(\rho, \phi, \theta) := (\rho\cos\phi\sin\theta, \rho\sin\phi\cos\theta, \rho\cos\theta)$$

Here,

$$\det[D\Phi_{sph}]_{(\rho,\phi,\theta)} = \rho^2\sin\theta.$$