Ramp Up Mathematics

–Calculus–

Lecture Notes

Prof. Dr. Volker Bach

Technische Universität Braunschweig
Summer Semester 2026

Version as of 03-Jun-2026

Contents

I. Convergence

I.1. Real Numbers

Recall that $\mathbb{N} := \{1, 2, 3, \ldots\}$ is the set of **natural numbers**, $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ is the set of **integers**, and

$$\mathbb{Q} \; := \; \left\{ \frac{a}{b} \,\middle|\, a \in \mathbb{Z}, \; b \in \mathbb{N} \right\}, \tag{I.1}$$

is the set of **rational numbers**. These sets of numbers are nested,

$$\mathbb{N} \; \subseteq \; \mathbb{Z} \; \subseteq \mathbb{Q} \tag{I.2}$$

and they are each **totally ordered**. In case of $\mathbb{Q}$ the latter means that,

$$\forall p, q \in \mathbb{Q} : \qquad \text{either} \quad p < q \quad \text{or} \quad p > q \quad \text{or} \quad p = q \tag{I.3}$$

holds true. Moreover, $\mathbb{Q}$ is an **ordered field** which additionally to (I.3) requires that

$$(p < q) \wedge (r \in \mathbb{Q}) \;\Rightarrow\; p + r < q + r \quad \text{and} \quad (p > 0) \wedge (q > 0) \;\Rightarrow\; p \cdot q > 0\,. \tag{I.4}$$

We consider the sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ with respect to the basic algebraic operations:

- Sums $m + n$ and products $m \cdot n$ of any two natural numbers $m, n \in \mathbb{N}$ are again natural numbers, but differences and quotients are not, in general.
- Sums $m + n$, differences $m - n$, and products $m \cdot n$ of any two integers $m, n \in \mathbb{Z}$ are again integers, but quotients are not, in general.
- For two rational numbers $p, q \in \mathbb{Q}$, their sum $p + q$, their difference $p - q$, their product $p \cdot q$ are again rational numbers. If additionally $q \neq 0$, then their quotient p/q is defined and is a rational number, too.

This means that within $\mathbb{Q}$ we can carry out all calculations that we need for everyday purpose and that a (simple) handcalulator can perform.

Yet, $\mathbb{Q}$ is lacking the important property of *completeness* which we now describe.

Definition I.1. Let $\mathbb{F}$ be an ordered field and $A \subseteq \mathbb{F}$ a subset.

(i) If there exists $b \in \mathbb{F}$ such that $b \geq a$, for all $a \in A$, then A is said to be **bounded above** and b is called an **upper bound (on A)**.

(ii) If there exists $c \in \mathbb{F}$ such that $c \leq a$, for all $a \in A$, then A is said to be **bounded below** and c is called a **lower bound (on A)**.

(iii) If A is bounded above and also bounded below, then A is said to be **bounded**.

Definition I.2. Let $\mathbb{F}$ be an ordered field and $A \subseteq \mathbb{F}$ a subset.

(i) A number $b \in \mathbb{F}$ is called **supremum of A**, if b is an upper bound on A and any number $b' \in \mathbb{F}$ strictly smaller than b is not an upper bound on A.

(ii) A number $c \in \mathbb{F}$ is called **infimum of A**, if c is a lower bound on A and any number $c' \in \mathbb{F}$ strictly bigger than c is not a lower bound on A.

Remarks and Examples.

- If $A \subseteq \mathbb{F}$ possesses a supremum $b \in \mathbb{F}$ then this number is the unique least upper bound on A and is denoted $b =: \sup\{A\}$.
- If $A \subseteq \mathbb{F}$ possesses an infimum $c \in \mathbb{F}$ then this number is the unique greatest lower bound on A and is denoted $c =: \inf\{A\}$.

Definition I.3 (LUB Axiom)**.** An ordered field $\mathbb{F}$ fulfills be **LUB axiom** or **supremum axiom**, if any subset $A \subseteq \mathbb{F}$, which is bounded above, possesses a supremum.

Remarks and Examples.

- The ordered field $\mathbb{Q}$ of rational numbers does not fulfill the LUB axiom. To see this define

$$A \ := \ \left\{ p \in \mathbb{Q} \,\middle|\, p^2 < 2 \right\} . \tag{I.5}$$

  Then A is bounded above by, e.g., $2 \in \mathbb{Q}$, but A does not have a supremum. For if $b \in \mathbb{Q}$ was a supremum of A then b would fulfill $b^2 = 2$. This equation, however, has not rational solution.

- Fortunately, we can enlarge the ordered field $\mathbb{Q}$ and pass to the ordered field $\mathbb{R} \supset \mathbb{Q}$ which does fulfill the LUB axiom.

Theorem I.4. There is a unique ordered field $\mathbb{R} \supset \mathbb{Q}$ which fulfills the LUB axiom and which is called the **real numbers**.

I.2. Convergent and Cauchy Sequences

- The real numbers $\mathbb{R}$ may be identified with the decimal numbers $\mathbb{D}$. Decimal numbers between zero and one, for instance, are defined as

$$\mathbb{D} \cap [0,1) \;=\; \Big\{ 0, a_1 a_2 a_3 a_4 \ldots \;\Big|\; a_j \in \{0,1,\ldots,9\}\,,\;\; \forall\, k \in \mathbb{N}\; \exists j \geq k :\; a_j \leq 8 \Big\}\,. \tag{I.6}$$

  The condition $\forall\, k \in \mathbb{N}\; \exists j \geq k :\; a_j \leq 8$ rules out decimal numbers with period 9, like $0,17999\ldots$ (which equals $0,18$).

- The set $\mathbb{Q}$ of rational numbers is *countable*, i.e., its elements can be written in a sequence $(p_n)_{n=1}^\infty$, so that $\mathbb{Q} = \{p_1, p_2, p_3, \ldots\}$.

- The set $\mathbb{R}$ of real numbers is not countable, i.e., $\mathbb{R}$ is *uncountable*.

- Any real number can be approximated by rational numbers to arbitrary accuracy. For example, consider the irrational number $\sqrt{2} = 1,4142136\ldots \in \mathbb{R} \setminus \mathbb{Q}$. We define

$$\begin{aligned} p_1 \;&:=\; 1,4 \;<\; \sqrt{2} \;\leq\; p_1 + \tfrac{1}{10}\,, \\ p_2 \;&:=\; 1,41 \;<\; \sqrt{2} \;\leq p_2 \;+ \tfrac{1}{100}\,, \\ p_3 \;&:=\; 1,414 \;<\; \sqrt{2} \;\leq p_3 \;+ \tfrac{1}{1.000}\,, \\ &\;\;\vdots \end{aligned} \tag{I.7}$$

- This way we obtain a sequence $(p_n)_{n=1}^\infty \in \mathbb{Q}^{\mathbb{N}}$ of rational numbers p_n such that $p_n < \sqrt{2} \leq p_n + 10^{-n}$ and thus

$$\forall\, n \in \mathbb{N} : \quad |p_n - \sqrt{2}| \;\leq\; 10^{-n}\,. \tag{I.8}$$

  So, given any prescribed accuracy $\varepsilon > 0$, we can ensure that the approximand p_k for $\sqrt{2}$ has accuracy ε or better, provided we choose $k \in \mathbb{N}$ so large that $10^{-k} \leq \varepsilon$.

- These concepts leads to the notion of *convergence* of a sequence which we describe below.

- To this end we recall that, if $\mathcal{M}$ is a set, then a **sequence in $\mathcal{M}$** is a map $a : \mathbb{N} \to \mathcal{M}$. The collection of all sequences in $\mathcal{M}$ is denoted $\mathcal{M}^{\mathbb{N}} := \{\mathbb{N} \to \mathcal{M}\}$. It is customary to write $(a_n)_{n=1}^\infty$ for the map $a : \mathbb{N} \to \mathcal{M}$ and a_n for its values [instead of $a(n)$].

  Specifically, sequences $(a_n)_{n=1}^\infty \in \mathbb{R}^{\mathbb{N}}$ in $\mathbb{R}$ are called **real sequences**.

Definition I.5. Let $(a_n)_{n=1}^\infty \in \mathbb{R}^{\mathbb{N}}$ a real sequence.

(i) $(a_n)_{n=1}^\infty \in \mathbb{R}^{\mathbb{N}}$ is **convergent** $:\Leftrightarrow$

$$\exists\, a \in \mathbb{R}\; \forall\, \varepsilon > 0\; \exists n_0 \in \mathbb{N}\; \forall\, n \geq n_0 : \quad |a_n - a| \;\leq\; \varepsilon\,. \tag{I.9}$$

In this case, a is called the **limit of $(a_n)_{n=1}^\infty$**, and we write $a =: \lim_{n\to\infty}\{a_n\}$ or $a_n \to a$, $n \to \infty$.

If $(a_n)_{n=1}^\infty$ is not convergent, it is called **divergent**.

(ii) $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ is a **Cauchy sequence** :$\Leftrightarrow$

$$\forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall m > n \geq n_0 : \quad |a_m - a_n| \leq \varepsilon . \tag{I.10}$$

Remarks and Examples.

- If $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ is convergent then it is a Cauchy sequence, too. Indeed, given $\varepsilon > 0$ we use (I.9) with $\frac{1}{2}\varepsilon$ and obtain $n_0 \in \mathbb{N}$ such that $|a_k - a| \leq \frac{1}{2}\varepsilon$, for all $k \geq n_0$. So, if $m > n \geq n_0$, we have that

$$|a_m - a_n| = \big|(a_m - a) - (a_n - a)\big| \leq |a_m - a| + |a_n - a| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon . \tag{I.11}$$

  Hence, $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ is a Cauchy sequence.

- A fundamental fact about the real numbers $\mathbb{R}$ is that the converse is true, as well: Any real Cauchy sequence is convergent, too. We formulate this in a theorem.

Theorem I.6. The real numbers are **complete**, i.e., a real sequence $(a_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N}$ is convergent if, and only if, it is a Cauchy sequence.

Proof. The proof of this theorem goes back to Cauchy and makes essential use of the fact that the real numbers $\mathbb{R}$ fulfill the LUB axiom. ☐

Remarks and Examples.

- If $0 < q < 1$ then $(q^n)_{n=1}^\infty$ is convergent and $\lim_{n\to\infty}\{q^n\} = 0$.
- Conversely, if $q > 1$ then $(q^n)_{n=1}^\infty$ is divergent.
- The sequence $\big((-1)^n\big)_{n=1}^\infty$ is divergent.
- The sequence $(a_n)_{n=1}^\infty$ below is convergent, namely,

$$a_n := \frac{n^3 + 4n - 7}{3n^2 + 4n^3 + 20} = \frac{1 + \frac{4}{n^2} - \frac{7}{n^3}}{\frac{3}{n} + 4 + \frac{20}{n^3}} \to \frac{1}{4} , \; n \to \infty . \tag{I.12}$$

- If $k \in \mathbb{N}$ and $q > 1$ then

$$\frac{n^k}{q^n} \to 0 , \; n \to \infty , \tag{I.13}$$

  i.e., any exponential function of n grows faster than any (fixed) power of n. To see this, define $\varepsilon := q^{1/k} - 1 > 0$ so that $q = (1+\varepsilon)^k$. Then

$$\frac{n^k}{q^n} = \left(\frac{n}{(1+\varepsilon)^n}\right)^k . \tag{I.14}$$

Now, if $n \geq 2$ then

$$(1+\varepsilon)^n \;=\; \sum_{\ell=0}^{n} \binom{n}{\ell} \varepsilon^{\ell} \;\geq\; \binom{n}{2} \varepsilon^2 \;=\; \frac{n(n-1)}{2}\, \varepsilon^2 \;\geq\; \frac{n^2\, \varepsilon^2}{4}\,, \tag{I.15}$$

by the binomial theorem, which implies that

$$0 \;\leq\; \frac{n^k}{q^n} \;\leq\; \left(\frac{4}{\varepsilon^2\, n} \right)^k \;\leq\; \frac{4^k}{\varepsilon^{2k}}\, \frac{1}{n^k} \;\to\; 0\,,\; n \to \infty\,. \tag{I.16}$$

II. Norms and Scalar Products

II.1. Normed Spaces and Banach Spaces

In this section we define norms on real vector spaces and real Banach spaces and collect some of their basic properties.

Definition II.1. Let X be a real vector space. A map $\| \cdot \| : X \to \mathbb{R}_0^+$ is called **Norm (on X)** $:\Leftrightarrow$

(i)

$$\forall\, x \in X : \qquad \big\{ \|x\| = 0 \;\Leftrightarrow\; x = 0 \big\} \tag{II.1}$$

(ii)

$$\forall\, x \in X,\ \lambda \in \mathbb{K} : \qquad \|\lambda x\| \;=\; |\lambda| \cdot \|x\|, \tag{II.2}$$

(iii)

$$\forall\, x, y \in X : \qquad \|x + y\| \;\leq\; \|x\| + \|y\|. \tag{II.3}$$

In this case $(X, \| \cdot \|)$ is said to be a **normed (vector) space**. We denote by

$$B_X(x, r) \;:=\; \big\{ y \in X \,\big|\, \|x - y\| < r \big\} \tag{II.4}$$

the **open ball about $x \in X$ of radius $r > 0$**.

Remarks and Examples.

- Property (i) is frequently referred to as *definiteness* of the norm $\| \cdot \|$,
- Property (ii) is known as *homogeneity* of the norm $\| \cdot \|$,
- Inequality (iii) is called the *triangle inequality*.

- For $d \in \mathbb{N}$ the set
$$\mathbb{R}^d := \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix} \,\middle|\, \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_d \in \mathbb{R} \right\} \tag{II.5}$$
is a real vector space with respect to componentwise operations:
$$\forall\, \vec{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix},\ \vec{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} \in \mathbb{R}^d,\ \lambda \in \mathbb{R} : \tag{II.6}$$
$$\vec{x} + \lambda \vec{y} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} + \lambda \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} := \begin{pmatrix} \alpha_1 + \lambda\beta_1 \\ \vdots \\ \alpha_d + \lambda\beta_d \end{pmatrix}.$$

- There are many norms one can define on $\mathbb{R}^d$; here are the most important ones:

ℓ^1-norm:
$$\left\| \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_1 := |\alpha_1| + |\alpha_2| + \ldots + |\alpha_d|, \tag{II.7}$$

ℓ^∞-norm or **maximum norm:**
$$\left\| \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_\infty := \max\big\{ |\alpha_1|,\ |\alpha_2|,\ \ldots,\ |\alpha_d| \big\}, \tag{II.8}$$

ℓ^2-norm or **euclidean norm:**
$$\left\| \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_{\mathrm{eucl}} := \left\| \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_2 := \sqrt{\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_d^2}\,. \tag{II.9}$$

- If the vector space X is of finite dimension $\dim[X] \in \mathbb{N}$ then all norms on X are **equivalent**. That is, if $\|\cdot\|_a, \|\cdot\|_b : X \to \mathbb{R}_0^+$ are two norms on X then there exist a constant $m > 0$ such that
$$\forall\, \vec{x} \in X : \quad m \cdot \|\vec{x}\|_a \le \|\vec{x}\|_b \le \frac{1}{m} \cdot \|\vec{x}\|_a\,. \tag{II.10}$$

- In particular, there are constants $0 < c_d < c_d' < c_d'' < \infty$ which depend only on $d \in \mathbb{N}$ such that

$$\forall\, \vec{x} \in \mathbb{R}^d : \quad c_d\, \|\vec{x}\|_1 \;\leq\; \|\vec{x}\|_{\mathrm{eucl}} \;\leq\; c_d'\, \|\vec{x}\|_\infty \;\leq\; c_d''\, \|\vec{x}\|_1 \,. \tag{II.11}$$

Lemma II.2. $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_{\mathrm{eucl}} : \mathbb{R}^d \to \mathbb{R}_0^+$ all define norms on $\mathbb{R}^d$.

Proof. We give the proof only for the maximum norm $\|\cdot\|_\infty$.
Definiteness:

$$\begin{aligned} \left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_\infty = 0 \;\;\Leftrightarrow\;\; & \max\big\{|\alpha_1|, \ldots, |\alpha_d|\big\} = 0 \;\;\Leftrightarrow\;\; |\alpha_1| = \cdots = |\alpha_d| = 0 \\ & \Leftrightarrow\;\; \alpha_1 = \cdots = \alpha_d = 0 \;\;\Leftrightarrow\;\; \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} . \end{aligned} \tag{II.12}$$

Homogeneity:

$$\begin{aligned} \left\| \lambda \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_\infty = \left\| \begin{pmatrix} \lambda\alpha_1 \\ \vdots \\ \lambda\alpha_d \end{pmatrix} \right\|_\infty & = \max\big\{|\lambda\alpha_1|, \ldots, |\lambda\alpha_d|\big\} \;=\; |\lambda| \cdot \max\big\{|\alpha_1|, \ldots, |\alpha_d|\big\} \\ & = |\lambda| \cdot \left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_\infty . \end{aligned} \tag{II.13}$$

Triangle Inequality: For $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_d, \beta_d \in \mathbb{R}$ let $\mu, \nu, \kappa \in \{1, 2, \ldots, d\}$ be such that

$$|\alpha_\mu| \;=\; \max\big\{|\alpha_1|, \ldots, |\alpha_d|\big\}, \quad |\beta_\nu| \;=\; \max\big\{|\beta_1|, \ldots, |\beta_d|\big\}, \tag{II.14}$$

$$\text{and} \quad |\alpha_\kappa + \beta_\kappa| \;=\; \max\big\{|\alpha_1 + \beta_1|, \ldots, |\alpha_d + \beta_d|\big\} . \tag{II.15}$$

Then

$$\left\| \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_d + \beta_d \end{pmatrix} \right\|_\infty = \; |\alpha_\kappa + \beta_\kappa| \;\leq\; |\alpha_\kappa| + |\beta_\kappa| \;\leq\; |\alpha_\mu| + |\beta_\nu| \;=\; \left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \right\|_\infty + \left\| \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} \right\|_\infty , \tag{II.16}$$

where we use the triangle inequality $|\alpha + \beta| \leq |\alpha| + |\beta|$ for real numbers $\alpha, \beta \in \mathbb{R}$ in the second step. ☐

Definition II.3. Let $(X, \|\cdot\|)$ be a normed space.

(i) A sequence $(\vec{x}_n)_{n=1}^\infty \in X^{\mathbb{N}}$ is **convergent** :⇔

$$\exists \vec{x} \in X \, \forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall n \geq n_0 : \quad \|\vec{x}_n - \vec{x}\| \leq \varepsilon. \tag{II.17}$$

(ii) A sequence $(\vec{x}_n)_{n=1}^\infty \in X^{\mathbb{N}}$ is called **Cauchy sequence** :⇔

$$\forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall m > n \geq n_0 : \quad \|\vec{x}_m - \vec{x}_n\| \leq \varepsilon. \tag{II.18}$$

(iii) The normed space $(X, \|\cdot\|)$ is said to be **complete** :⇔

$$(\vec{x}_n)_{n=1}^\infty \in X^{\mathbb{N}} \text{ is convergent} \quad \Leftrightarrow \quad (\vec{x}_n)_{n=1}^\infty \in X^{\mathbb{N}} \text{ is a Cauchy sequence}\,. \tag{II.19}$$

A complete normed space $(X, \|\cdot\|)$ is called a **Banach space**.

Theorem II.4. If $(X, \|\cdot\|)$ is a finite-dimensional normed space then $(X, \|\cdot\|)$ is complete, i.e., a Banach space.

Corollary II.5. If $d \in \mathbb{N}$ then each $\big(\mathbb{R}^d, \|\cdot\|_1\big)$, $\big(\mathbb{R}^d, \|\cdot\|_{\mathrm{eucl}}\big)$, and $\big(\mathbb{R}^d, \|\cdot\|_\infty\big)$ is a Banach space.

II.2. Scalar Products (Inner Products)

In this section we define (real) scalar products and collect some basic facts about these.

Definition II.6. Let X be a real vector space.

(i) A map $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{R}$ is a **bilinear form (on X)**

$$\begin{aligned} :\Leftrightarrow \; & \forall\, \alpha, \beta \in \mathbb{R} \, \forall\, \vec{x}, \vec{y}, \vec{w}, \vec{z} \in X : \\ & \langle \alpha\vec{x} + \vec{y} \,|\, \beta\vec{w} + \vec{z} \rangle \;=\; \alpha\beta\langle \vec{x}|\vec{w}\rangle + \alpha\langle \vec{x}|\vec{z}\rangle + \beta\langle \vec{y}|\vec{z}\rangle + \langle \vec{y}|\vec{z}\rangle. \end{aligned} \tag{II.20}$$

(ii) If a bilinear form $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{R}$ fulfills

$$:\Leftrightarrow \forall\, \vec{x}, \vec{y} \in X : \quad \langle \vec{x}|\vec{y}\rangle \;=\; \langle \vec{y}|\vec{x}\rangle\,, \tag{II.21}$$

then it is called **symmetric**.

(iii) A symmetric bilinear form $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{R}$ is called **positiv definite**

$$:\Leftrightarrow \quad \forall\, \vec{x} \in X \setminus \{\vec{0}\} : \quad \langle \vec{x}|\vec{x}\rangle \;>\; 0\,. \tag{II.22}$$

(iv) A positive definite symmetric bilinear form $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{R}$ is called **scalar product (on X)** or **inner product (on X)**.

Theorem II.7 (Cauchy-Schwarz Inequality)**.** Let X be a real vector space equipped with a scalar product $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{R}$. Then

$$\forall \vec{x}, \vec{y} \in X : \quad \big|\langle\vec{x}|\vec{y}\rangle\big| \;\leq\; \sqrt{\langle\vec{x}|\vec{x}\rangle}\,\sqrt{\langle\vec{y}|\vec{y}\rangle}. \tag{II.23}$$

Corollary II.8. If X be a real vector space equipped with a scalar product $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{R}$. Then

$$\|\cdot\| : X \;\to\; \mathbb{R}_0^+, \quad \|\vec{x}\| \;:=\; \langle\vec{x}|\vec{x}\rangle^{1/2} \tag{II.24}$$

defines a norm on X, the **norm induced by** $\langle\cdot|\cdot\rangle$.

Remarks and Examples.

- If $d \in \mathbb{N}$ and $X = \mathbb{R}^d$ then

  $$\forall\, \vec{x} = (\alpha_1, \ldots, \alpha_d)^t\,,\; \vec{y} = (\beta_1, \ldots, \beta_d)^t \in \mathbb{R}^d : \quad \big\langle\vec{x}|\vec{y}\big\rangle_{\mathrm{eucl}} \;:=\; \sum_{\nu=1}^{d} x_\nu\, y_\nu \tag{II.25}$$

  defines the **Euclidean scalar product**.

- Let $X = \mathbb{R}[x]$ be the space of real polynomials in the variable x, i.e., for vectors $p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_M x^M$ and $q(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_N x^N$, the vector operations are defined by

  $$(p + \lambda q)(x) \;:=\; (\alpha_0 + \lambda\beta_0) + (\alpha_1 + \lambda\beta_1)x + (\alpha_2 + \lambda\beta_2)x^2 + \ldots + (\alpha_L + \lambda\beta_L)x^L\,, \tag{II.26}$$

  where $L = \max(M, N)$. A scalar product on $\mathbb{R}[x]$ is defined by

  $$\langle p|q\rangle \;:=\; \int_{-\infty}^{\infty} p(x)\, q(x)\, \frac{e^{-x^2}\, dx}{\sqrt{2\pi}}\,. \tag{II.27}$$

III. Real Functions of One Real Variable

In this chapter we define and discuss real-valued functions $f : I \to \mathbb{R}$ defined on $I \subseteq \mathbb{R}$. We always assume that I is an **interval** by which we mean that there are numbers $a \in \mathbb{R}$ and $b \in \mathbb{R}$ with $a < b$ such that either

$$I = (a\,,\, b) \quad \text{or} \quad I = (-\infty\,,\, b) \quad \text{or} \quad I = (a\,,\, \infty) \quad \text{or} \quad I = \mathbb{R}\,. \tag{III.1}$$

III.1. Continuity

Definition III.1. Let I be an interval, $f : I \to \mathbb{R}$, and $x_0 \in I$.

(i) $$\lim_{x \to x_0} \{f(x)\} = y \quad :\Leftrightarrow \tag{III.2}$$

$$\forall\, (x_n)_{n=1}^{\infty} \in (I \setminus \{x_0\})^{\mathbb{N}} : \Big(\lim_{n\to\infty} \{x_n\} = x_0\Big) \Rightarrow \Big(\lim_{n\to\infty} \{f(x_n)\} = y\Big)\,.$$

(ii) f is **continuous at $\boldsymbol{x_0}$** $$:\Leftrightarrow \quad \lim_{x \to x_0} \{f(x)\} = f(x_0)\,. \tag{III.3}$$

(iii) f is **continuous on $\boldsymbol{I}$** $$:\Leftrightarrow \quad \forall\, x_0 \in I : \ f \text{ is continuous at } x_0\,. \tag{III.4}$$

Remarks and Examples. Most of the elementary functions are continuous on their maximal domain of definition. For example,

- Polynomials $p : \mathbb{R} \to \mathbb{R}$ are continuous on $\mathbb{R}$, where $p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_M x^M$.
- Rational functions $p/q : D \to \mathbb{R}$ are continuous on I for any interval $I \subseteq D := \{x \in \mathbb{R} |\, q(x) \neq 0\}$, where p and q are polynomials.
- Trigonometric and exponential functions $\sin$, $\cos$, and $\exp$ are continuous on $\mathbb{R}$, the logarithm $\ln$ is continuous on $\mathbb{R}^+ = (0, \infty)$.

- If $f, g : I \to \mathbb{R}$ are continuous at $x_0 \in I$ and $\alpha \in \mathbb{R}$, then $f + \alpha g : I \to \mathbb{R}$ is continuous at x_0.
- If $g : I \to \mathbb{R}$ is continuous at $x_0 \in I$ and $f : J \to \mathbb{R}$ is defined on an interval $J \supseteq g(I)$ and is continuous at $g(x_0)$, then the composition $f \circ g$ is defined on I by $(f \circ g)[x] = f[g(x)]$ and $f \circ g : I \to \mathbb{R}$ is continuous at x_0.
- The function $\theta : \mathbb{R} \to \mathbb{R}$,
$$\theta(x) \ := \ \begin{cases} 1\,, & x \geq 0\,, \\ 0\,, & x < 0\,, \end{cases} \tag{III.5}$$
is not continuous at $x_0 = 0$, but for all $x_0 \in \mathbb{R} \setminus \{0\}$.
- Rule of Thumb: A function is continuous if its graph can be drawn without lifting the pen off the paper.

III.2. Differentiability

Definition III.2. Let I be an interval, $f : I \to \mathbb{R}$, and $x_0 \in I$.

(i) f is **differentiable at $\boldsymbol{x_0}$** $\quad :\Leftrightarrow \quad \exists\, y \in \mathbb{R} : \quad \lim_{x \to x_0} \left\{ \frac{f(x) - f(x_0)}{x - x_0} \right\} \ = \ y\,.$ (III.6)

In this case $y =: f'(x_0)$ is called the **derivative of $\boldsymbol{f}$ at $\boldsymbol{x_0}$**.

Another customary notation for $f'(x_0)$ is $\dfrac{df(x_0)}{dx}$.

(ii) f is **differentiable on $\boldsymbol{I}$** $\quad :\Leftrightarrow \quad \forall\, x_0 \in I : \ f$ is differentiable at x_0. (III.7)

Remarks and Examples. Most of the elementary functions are differentiable on their maximal domain of definition. For example,

- Polynomials $p : \mathbb{R} \to \mathbb{R}$ are differentiable on $\mathbb{R}$, and for $p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_M x^M$, we have that
$$p'(x) \ = \ \alpha_1 + 2\alpha_2 x + \ldots + M \alpha_M x^{M-1}\,. \tag{III.8}$$
- Rational functions $p/q : D \to \mathbb{R}$ are differentiable on I for any interval $I \subseteq D := \{x \in \mathbb{R} |\, q(x) \neq 0\}$, where p and q are polynomials, and
$$\left(\frac{p}{q}\right)'(x) \ = \ \frac{p'(x)q(x) - p(x)q'(x)}{q^2(x)}\,. \tag{III.9}$$
- Trigonometric and exponential functions sin, cos, and exp are differentiable on $\mathbb{R}$, the logarithm ln is differentiable on $\mathbb{R}^+ = (0, \infty)$, and
$$\sin'(x) = \cos(x)\,, \quad \cos'(x) = -\sin(x)\,, \quad \exp'(x) = \exp(x)\,, \quad \ln'(x) = \frac{1}{x}\,. \tag{III.10}$$

The next two theorems provide rules for the differentiation of functions which enable us to calculate the derivative of a vast variety of functions.

Theorem III.3. Let I be an interval, $f, g : I \to \mathbb{R}$ differentiable at $x_0 \in I$, and $\alpha \in \mathbb{R}$. Then $f + \alpha g, f \cdot g : I \to \mathbb{R}$ are differentiable at $x_0 \in I$, and

$$\underline{\text{Linearity:}} \qquad (f + \alpha g)'(x_0) \;=\; f'(x_0) + \alpha g'(x_0)\,, \tag{III.11}$$

$$\underline{\text{Leibniz Rule:}} \qquad (f \cdot g)'(x_0) \;=\; f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)\,. \tag{III.12}$$

If furthermore $g(x_0) \neq 0$ then $f/g : I' \to \mathbb{R}$ is defined and differentiable at x_0, for some interval $x_0 \in I' \subseteq I$, and

$$\underline{\text{Quotient Rule:}} \qquad \left(\frac{f}{g}\right)'(x_0) \;=\; \frac{f'(x_0) \cdot g(x_0) - f(x_0) \cdot g'(x_0)}{g^2(x_0)}\,. \tag{III.13}$$

Theorem III.4 (Chain Rule)**.** Let I, J be two intervals, $g : I \to J$ differentiable at $x_0 \in I$, and $f : J \to \mathbb{R}$ differentiable at $g(x_0) \in J$. Then $f \circ g : I \to \mathbb{R}$ is differentiable at $x_0 \in I$ and

$$\underline{\text{Chain Rule:}} \qquad \big(f \circ g\big)'(x_0) \;=\; f'[g(x_0)] \cdot g'(x_0)\,. \tag{III.14}$$

Remarks and Examples.

$$\left(\sin\left[\sqrt{\frac{x}{x^2+1}}\right]\right)' \;=\; \cos\left[\sqrt{\frac{x}{x^2+1}}\right] \cdot \left[\Big(\frac{x}{x^2+1}\Big)^{1/2}\right]' \tag{III.15}$$

$$= \cos\left[\sqrt{\frac{x}{x^2+1}}\right] \cdot \frac{1}{2}\Big(\frac{x}{x^2+1}\Big)^{-1/2} \cdot \Big(\frac{x}{x^2+1}\Big)' \tag{III.16}$$

$$= \frac{1}{2}\cos\left[\sqrt{\frac{x}{x^2+1}}\right] \cdot \Big(\frac{x}{x^2+1}\Big)^{-1/2} \cdot \Big(\frac{1 \cdot (x^2+1) - x \cdot (2x)}{(x^2+1)^2}\Big) \tag{III.17}$$

$$= \frac{1}{2}\cos\left[\sqrt{\frac{x}{x^2+1}}\right] \cdot \Big(\frac{x}{x^2+1}\Big)^{-1/2} \cdot \Big(\frac{1 - x^2}{(x^2+1)^2}\Big)\,, \tag{III.18}$$

where we use the chain rule in the first equation and also to get from (III.16) to (III.17). The quotient rule gets us from (III.17) to (III.18).

III.3. Taylor's Theorem (for one Real Variable)

Definition III.5. Let $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ be a nonnegative integer, I be an interval, and $f : I \to \mathbb{R}$.

f is $\boldsymbol{n}$ **times continuously differentiable on** $\boldsymbol{I}$ $:\Leftrightarrow$

$$\begin{aligned}
& f =: f^{(0)} \text{ is differentiable on } I\text{, and its derivative} \\
& f' =: f^{(1)} \text{ is differentiable on } I\text{, and its derivative} \\
& f'' =: f^{(2)} \text{ is differentiable on } I\text{, and} \\
& \qquad \vdots \\
& f^{(n-1)} \text{ is differentiable on } I\text{, and its derivative } f^{(n)} \text{ is continuous on } I.
\end{aligned} \tag{III.19}$$

The next theorem yields a very important method to approximate functions locally by polynomials.

Theorem III.6 (Taylor)**.** Let $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ be a nonnegative integer, I be an interval, $f : I \to \mathbb{R}$ be an $n + 1$ times continuously differentiable function, and $x_0 \in I$. Then, for any $x \in I \setminus \{x_0\}$, there exists $x' \in I \setminus \{x_0\}$ with $0 < |x' - x_0| < |x - x_0|$, such that

$$f(x) \;=\; T_{n;\,x_0}[x] \;+\; R_{n+1;\,x_0}[x]\,, \tag{III.20}$$

with

$$\begin{aligned}
T_{n;\,x_0}[x] \;&:=\; \sum_{k=0}^{n} \frac{(x-x_0)^k}{k!}\, f^{(k)}(x_0) \qquad \text{(III.21)} \\
&= f(x_0) \;+\; f'(x_0)\,(x-x_0) \;+\; \frac{f''(x_0)}{2}(x-x_0)^2 \;+\; \ldots \;+\; \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n
\end{aligned}$$

and

$$R_{n+1;\,x_0}[x] \;:=\; \frac{f^{(n+1)}(x')}{n!}(x-x_0)^{n+1}\,. \tag{III.22}$$

Remarks and Examples. Let $n \in \mathbb{N}$, $x \in \mathbb{R}$, and $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \exp[x] := e^x$. Then f is $n + 1$ times continuously differentiable on $\mathbb{R}$ and

$$f(x) \;=\; f'(x) \;=\; f''(x) \;=\; \ldots \;=\; f^{(n+1)}(x) \;=\; e^x\,. \tag{III.23}$$

In particular, for $x_0 := 0$, we have that

$$f(0) \;=\; f'(0) \;=\; f''(0) \;=\; \ldots \;=\; f^{(n)}(0) \;=\; 1 \tag{III.24}$$

and hence

$$T_{n;\,0}[x] \;=\; 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots + \frac{x^n}{n!}\,, \tag{III.25}$$

for any $x \in \mathbb{R}$. Taylor's Theorem III.6 implies that, given $x > 0$, there exists $x' \in (0, x)$ such that

$$\left| e^x - \Big(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots + \frac{x^n}{n!} \Big) \right| \;=\; \frac{x^{n+1}}{(n+1)!}\, e^{x'} \;\leq\; \frac{x^{n+1}}{(n+1)!}\, e^{x} \,, \qquad \text{(III.26)}$$

since $(e^x)' = e^x > 0$ and thus $x \mapsto e^x$ is monotonically increasing.

Choosing $x := 1$ and $n = 10$, we obtain

$$\left| e - \Big(1 + 1 + \frac{1}{2} + \frac{1}{6} + \ldots + \frac{1}{10!} \Big) \right| \;=\; \frac{x^{n+1}}{(n+1)!}\, e^{x'} \;\leq\; \frac{1}{11!}\, e \,, \qquad \text{(III.27)}$$

which yields a concrete estimate for Euler's number e of four decimals accuracy,

$$2.71827 \;\leq\; e \;\leq\; 2.71829 \,. \qquad \text{(III.28)}$$

As another application of Taylor's theorem we discuss minima and maxima of functions. First we distinguish *local* from *global minima* and *local* from *global maxima*.

Definition III.7. Let $A \subseteq \mathbb{R}$ be a set (not necessarily an interval) and $f : A \to \mathbb{R}$ a real function on A. A point $x_0 \in A$ is a

$$\textbf{local minimum of } \boldsymbol{f} \quad :\Leftrightarrow \quad \exists\, \delta > 0\, \forall\, x \in A,\ |x - x_0| < \delta : \ f(x) \;\geq\; f(x_0) \,; \qquad \text{(III.29)}$$

$$\textbf{global minimum of } \boldsymbol{f} \quad :\Leftrightarrow \quad \forall\, x \in A : \ f(x) \;\geq\; f(x_0) \,; \qquad \text{(III.30)}$$

$$\textbf{local maximum of } \boldsymbol{f} \quad :\Leftrightarrow \quad \exists\, \delta > 0\, \forall\, x \in A,\ |x - x_0| < \delta : \ f(x) \;\leq\; f(x_0) \,; \qquad \text{(III.31)}$$

$$\textbf{global maximum of } \boldsymbol{f} \quad :\Leftrightarrow \quad \forall\, x \in A : \ f(x) \;\leq\; f(x_0) \,. \qquad \text{(III.32)}$$

A key application of derivatives and Taylor's theorem is the detection and identification of local minima and maxima. The first theorem below states that at local minima or maxima the derivative of the function under investigation vanishes.

Theorem III.8. Let $A \subseteq \mathbb{R}$ be a set, $f : A \to \mathbb{R}$, and $x_0 \in A$ such that there is an interval $I \ni x_0$ such that $I \subseteq A$. Assume furthermore that f is differentiable at x_0. Then

$$x_0 \text{ is a local minimum or a local maximum of } f \quad \Rightarrow \quad f'(x_0) \;=\; 0 \,. \qquad \text{(III.33)}$$

Having found all candidates for local minima and maxima, the following theorem provides a criterion to decide whether a point with vanishing derivative of f actually is a local minimum or a local maximum or something else.

Theorem III.9. Let $I \subseteq \mathbb{R}$ be an interval, $n \in \mathbb{N}$, a function $f : I \to \mathbb{R}$ which is n times continuously differentiable, $x_0 \in I$ and

$$f'(x_0) \;=\; f''(x_0) \;=\; f'''(x_0) \;=\; \cdots \;=\; f^{(n-1)}(x_0) \;=\; 0 \,, \;\; f^{(n)}(x_0) \;\neq\; 0 \,. \qquad \text{(III.34)}$$

Then

(i) If n is even and $f^{(n)}(x_0) > 0$ then x_0 is a local minimum of f.
(ii) If n is even and $f^{(n)}(x_0) < 0$ then x_0 is a local maximum of f.
(iii) If n is odd then x_0 is a point of inflexion of f.

Proof. [only (i)] If $x \in I$ then

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \\ &\qquad \ldots + \frac{f^{(n-1)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n)}(x')}{n!}(x - x_0)^n\,, \\ &= f(x_0) + \frac{f^{(n)}(x')}{n!}(x - x_0)^n\,, \end{aligned} \tag{III.35}$$

for some $x' \in I$ with $0 < |x' - x_0| < |x - x_0|$, according to Taylor's Theorem III.6, additionally using Assumption (III.34). Since we further assumed that $f^{(n)}$ is continuous at x_0, we have $f^{(n)}(x') \geq \frac{1}{2} f^{(n)}(x_0) > 0$, provided x' is sufficiently close to x_0. That is, if $\delta > 0$ is sufficiently small we have for all $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ that

$$f(x) \;=\; f(x_0) + \frac{f^{(n)}(x')}{n!}(x - x_0)^n \;\geq\; f(x_0) + \frac{f^{(n)}(x_0)}{2\,n!}(x - x_0)^n \;>\; f(x_0)\,, \tag{III.36}$$

since n is even and hence $(x - x_0)^n > 0$. ☐

Remarks and Examples. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) := x^4$. Then

$$f'(x_0) \;=\; 4x_0^3 \;=\; 0 \quad\Leftrightarrow\quad x_0 \;=\; 0\,, \tag{III.37}$$

and possible local extrema may only occur at $x_0 = 0$. We further differentiate f' and find

$$f''(x) \;=\; 12x^2\,,\;\; f'''(x) \;=\; 24x\,,\;\; f^{(iv)}(x) \;=\; 24\,. \tag{III.38}$$

Inserting $x = 0$, we obtain

$$f'(0) \;=\; f''(0) \;=\; f'''(0) \;=\; 0\,,\;\; \text{and}\;\; f^{(iv)}(0) \;=\; 24 \;>\; 0\,, \tag{III.39}$$

and hence Theorem III.9 (i) implies that $x_0 = 0$ is a local minimum. Since there are no other local extrema, $x_0 = 0$ is also a global minimum, and f possesses no local nor global maxima.

IV. Vector-Valued Functions of Several Variables

In this chapter we discuss functions $f : X \to Y$ that map a real Banach space $\big(X, \|\cdot\|_X\big)$ into another real Banach space $\big(Y, \|\cdot\|_Y\big)$. First, we generalize the notions of continuity and differentiablilty which we defined in Chapter III for $\big(X, \|\cdot\|_X\big) = \big(Y, \|\cdot\|_Y\big) = \big(\mathbb{R}, |\cdot|\big)$. To this end, we recall from Chapter II that

- a real *Banach space* $\big(X, \|\cdot\|\big)$ is a real vector space X, on which a norm $\|\cdot\| : X \to \mathbb{R}_0^+$ is defined, such that every Cauchy sequence in X converges and
- that a sequence $(\vec{x}_n)_{n=1}^\infty \in X^{\mathbb{N}}$ is *convergent* with limit $\vec{x} := \lim_{n\to\infty}\{\vec{x}_n\} \in X$, if $\|\vec{x}_n - \vec{x}\| \to 0$, as $n \to \infty$.

IV.1. Total Differentiability and Partial Differentiability

Definition IV.1. Let $\big(X, \|\cdot\|_X\big)$ and $\big(Y, \|\cdot\|_Y\big)$ be two real Banach space, $f : X \to Y$, and $\vec{x}_0 \in X$.

(i) $$\boldsymbol{\lim_{\vec{x}\to\vec{x}_0} \big\{f(\vec{x})\big\} = \vec{y}} \;:\Leftrightarrow \tag{IV.1}$$

$$\forall\, (\vec{x}_n)_{n=1}^\infty \in (X \setminus \{\vec{x}_0\})^{\mathbb{N}} : \Big(\lim_{n\to\infty}\{\vec{x}_n\} = \vec{x}_0\Big) \Rightarrow \Big(\lim_{n\to\infty}\{f(\vec{x}_n)\} = \vec{y}\Big).$$

(ii) $$f \text{ is } \textbf{continuous at } \boldsymbol{\vec{x}_0} \;:\Leftrightarrow\; \lim_{\vec{x}\to\vec{x}_0}\big\{f(\vec{x})\big\} = f(\vec{x}_0). \tag{IV.2}$$

(iii) f is **totally differentiable at $\vec{x}_0$** $:\Leftrightarrow$

There exists a continuous linear map $A : X \to Y$ so that

$$\lim_{\vec{x}\to\vec{x}_0}\left\{\frac{\|f(\vec{x}) - f(\vec{x}_0) - A(\vec{x} - \vec{x}_0)\|_Y}{\|\vec{x} - \vec{x}_0\|_X}\right\} = 0. \tag{IV.3}$$

In this case $A =: f'(\vec{x}_0)$ is called **total derivative of f at $\vec{x}_0$.**

Having defined the notion of *total* differentiabilty, we next define *partial* differentiability.

Definition IV.2. Let $M, N \in \mathbb{N}$ be two natural numbers, $f = (f_1, f_2, \ldots, f_N)^T \mathbb{R}^M \to \mathbb{R}^N$ a function, and $\vec{x} = (x_1, x_2, \ldots, x_M)^T \in \mathbb{R}_M$ a vector.

(i) The function f is **partially differentiable at $\vec{x}$** $:\Leftrightarrow$

$$\begin{aligned} &\forall\, m \in \mathbb{Z}_1^M,\ n \in \mathbb{Z}_1^N : \\ &\frac{\partial f_n(\vec{x})}{\partial x_m} := \lim_{h \to 0} \left\{ \frac{f_n(x_1, \ldots, x_m + h, \ldots, x_M) \ - \ f_n(x_1, \ldots, x_m, \ldots, x_M)}{h} \right\} \text{ exists.} \end{aligned} \tag{IV.4}$$

(ii) The function f is **continuously differentiable at $\vec{x}$** $:\Leftrightarrow$

$$\begin{aligned} &f \text{ is partially differentiable at } \vec{x} \text{ and} \\ &\forall\, m \in \mathbb{Z}_1^M,\ n \in \mathbb{Z}_1^N : \quad \frac{\partial f_n(\vec{x})}{\partial x_m} \text{ is continuous at } \vec{x}. \end{aligned} \tag{IV.5}$$

(iii) The space of **continuously differentiable functions from $\mathbb{R}^M$ to $\mathbb{R}^M$ is defined as**

$$\begin{aligned} &C^1\big(\mathbb{R}^M; \mathbb{R}^N\big) \\ &:= \Big\{ f : \mathbb{R}^M \to \mathbb{R}^N \ \Big|\ \forall\, \vec{x} \in \mathbb{R}^M : f \text{ is continuously differentiable at } \vec{x} \Big\}. \end{aligned} \tag{IV.6}$$

If the Banach spaces in Definition IV.1 are chosen as $\big(X, \|\cdot\|_X\big) = \big(\mathbb{R}^M, \|\cdot\|_{\text{eucl}}\big)$ and $\big(Y, \|\cdot\|_Y\big) = \big(\mathbb{R}^N, \|\cdot\|_{\text{eucl}}\big)$, then the notions of total and partial differentiability become comparable.

Theorem IV.3. Let $M, N \in \mathbb{N}$ be two natural numbers, $f = (f_1, f_2, \ldots, f_N)^T : \mathbb{R}^M \to \mathbb{R}^N$ a function, and $\vec{x} = (x_1, x_2, \ldots, x_M)^T \in \mathbb{R}_M$ a vector.

(i) If f is totally differentiable at $\vec{x}$, then f is partially differentiable at $\vec{x}$.

(ii) If f is continuously differentiable at $\vec{x}$, then f is totally differentiable at $\vec{x}$, and the derivative f' is continuous at $\vec{x}$. In this case, the derivative of f is

$$J_f := f' = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_M} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_M} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_N}{\partial x_1} & \frac{\partial f_N}{\partial x_2} & \cdots & \frac{\partial f_N}{\partial x_M} \end{pmatrix}, \tag{IV.7}$$

the **Jacobian J_f of f**.

IV.2. Continuous Differentiability

Theorem IV.3 states that, in general, total differentiability is stronger than partial differentiablity. If we additionally require the continuity of the derivative with respect to $\vec{x}$, then the two notions coincide.

We recursively define $C^1(\mathbb{R}^M;\mathbb{R}^N)$ by (IV.6) and, for $k \in \mathbb{N}$, the space of *k-times continuously differentiable functions from $\mathbb{R}^M$ to $\mathbb{R}^N$* by

$$C^k\big(\mathbb{R}^M;\mathbb{R}^N\big) := \Big\{ f \in C^{k-1}\big(\mathbb{R}^M;\mathbb{R}^N\big) \;\Big|\; \forall\, m_1, m_2, \ldots, m_{k-1} \in \mathbb{Z}_1^M : \; \frac{\partial^{m-1} f(\vec{x})}{\partial x_{m_{k-1}} \cdots \partial x_{m_1}} \in C^1(\mathbb{R}^M;\mathbb{R}^N) \Big\}, \tag{IV.8}$$

and by

$$C^\infty\big(\mathbb{R}^M;\mathbb{R}^N\big) := \bigcap_{k=1}^{\infty} C^k\big(\mathbb{R}^M;\mathbb{R}^N\big) \tag{IV.9}$$

the space of *smooth functions from $\mathbb{R}^M$ to $\mathbb{R}^N$*.

Remarks and Examples.

- Most of the elementary functions are smooth on $\mathbb{R}^M$.
- If $N = 1$ then $f \in C^1(\mathbb{R}^M;\mathbb{R})$ is real-valued, and in this case the (transpose of the) Jacobian of f is often called its **gradient** and is usually denoted by

$$\nabla f := J_f^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_1} \end{pmatrix}. \tag{IV.10}$$

- Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$f(x,y,z) := \sin\big(x^2+y^2\big)\, e^{z^2} - \cos\left(\frac{x}{z^3+1}\right), \tag{IV.11}$$

where we denote the three real variables as $x := x_1$, $y := x_2$, and $z := x_3$. Using the rules (III.11)-(III.14), we compute the partial derivatives of f as

$$\frac{\partial f(x,y,z)}{\partial x} = 2x \cos\big(x^2+y^2\big)\, e^{z^2} + \frac{1}{z^3+1} \sin\left(\frac{x}{z^3+1}\right), \tag{IV.12}$$

$$\frac{\partial f(x,y,z)}{\partial y} = 2y \cos\big(x^2+y^2\big)\, e^{z^2}, \tag{IV.13}$$

$$\frac{\partial f(x,y,z)}{\partial z} = 2z \sin\big(x^2+y^2\big)\, e^{z^2} - \frac{2xz}{(z^3+1)^2} \sin\left(\frac{x}{z^3+1}\right). \tag{IV.14}$$

Each of these partial derivatives are continuous functions on $\mathbb{R}^3$, therefore $f \in C^1(\mathbb{R}^3;\mathbb{R})$.

[In fact, $f \in C^\infty(\mathbb{R}^3;\mathbb{R})$ is smooth.] We obtain the gradient of f as

$$[\nabla f](x,y,z) \;=\; \begin{pmatrix} 2x\,\cos\big(x^2+y^2\big)\,e^{z^2} \;+\; \frac{1}{z^3+1}\,\sin\left(\frac{x}{z^3+1}\right) \\ 2y\,\cos\big(x^2+y^2\big)\,e^{z^2} \\ 2z\,\sin\big(x^2+y^2\big)\,e^{z^2} \;-\; \frac{2xz}{(z^3+1)^2}\,\sin\left(\frac{x}{z^3+1}\right) \end{pmatrix}. \qquad \text{(IV.15)}$$

- The continuity of the partial derivatives not only ensure that continuous partial differentiability and continuous total differentiablity are equivalent, but it also guarantees that the order of taking the partial derivatives does not matter. More precisely, if $k \in \mathbb{N}$, with $k \geq 2$, and $f \in C^k(\mathbb{R}^M;\mathbb{R}^N)$, then the order of the partial derivatives is immaterial. That is, if $j_1, j_2, \ldots, j_k \in \mathbb{Z}_1^M$ and $n \in \mathbb{Z}_1^N$ then

$$\frac{\partial}{\partial x_{j_1}}\left[\frac{\partial}{\partial x_{j_2}}\Big(\cdots\Big(\frac{\partial f}{\partial x_k}\Big)\cdots\Big)\right] \;=\; \frac{\partial}{\partial x_{j_{\pi(1)}}}\left[\frac{\partial}{\partial x_{j_{\pi(2)}}}\Big(\cdots\Big(\frac{\partial f}{\partial x_{j_{\pi(k)}}}\Big)\cdots\Big)\right], \qquad \text{(IV.16)}$$

for any permutation $\pi : \mathbb{Z}_1^M \to \mathbb{Z}_1^M$. This justifies the notation

$$\frac{\partial^k f}{\partial x_{j_1}\,\partial x_{j_2}\cdots\partial x_k} \;:=\; \frac{\partial}{\partial x_{j_1}}\left[\frac{\partial}{\partial x_{j_2}}\Big(\cdots\Big(\frac{\partial f}{\partial x_k}\Big)\cdots\Big)\right]. \qquad \text{(IV.17)}$$

- As an example, we compute the second partial derivatives of $f \in C^\infty(\mathbb{R}^3;\mathbb{R})$ defined in (IV.11) with respect to x and y. Starting from (IV.12) and (IV.13), we obtain

$$\begin{aligned} \frac{\partial}{\partial y}\left[\frac{\partial f(x,y,z)}{\partial x}\right] &= \frac{\partial}{\partial y}\left[2x\,\cos\big(x^2+y^2\big)\,e^{z^2} \;+\; \frac{1}{z^3+1}\,\sin\left(\frac{x}{z^3+1}\right)\right] \\ &= 2x\,e^{z^2}\,\frac{\partial[\cos(x^2+y^2)]}{\partial y} \;=\; -4xy\,e^{z^2}\,\sin(x^2+y^2), \end{aligned} \qquad \text{(IV.18)}$$

and

$$\begin{aligned} \frac{\partial}{\partial x}\left[\frac{\partial f(x,y,z)}{\partial y}\right] &= \frac{\partial}{\partial x}\left[2y\,\cos\big(x^2+y^2\big)\,e^{z^2}\right] \\ &= 2y\,e^{z^2}\,\frac{\partial[\cos(x^2+y^2)]}{\partial x} \;=\; -4xy\,e^{z^2}\,\sin(x^2+y^2), \end{aligned} \qquad \text{(IV.19)}$$

which agree, indeed.

IV.3. Taylor's Theorem for several Real Variables

The next, final theorem is the analogue of Taylor's theorem (Theorem III.6) for functions of $M \geq 2$ variables. Just as its one-dimensional counterpart it yields an important local approximation method for functions by polynomials.

Theorem IV.4 (Taylor)**.** Let $M \in \mathbb{N}$, $k \in \mathbb{N}_0$, and $f \in C^{k+1}(\mathbb{R}^M;\mathbb{R})$. Then, for any $x_0, z = (z_1, z_2, \ldots, z_M)^T \in \mathbb{R}^M$, we have that

$$f(x) = \sum_{\kappa=0}^{k} \frac{1}{\kappa!} \sum_{m_1,\ldots,m_\kappa=1}^{M} \left(\frac{\partial^\kappa f(x_0)}{\partial x_{m_1}\,\partial x_{m_2}\cdots\partial x_{m_\kappa}} \right) z_{m_1}\, z_{m_2} \cdots z_{m_\kappa} \tag{IV.20}$$

$$+ \sum_{m_1,\ldots,m_{\kappa+1}=1}^{M} \left(\int_0^1 \frac{(1-s)^k}{k!} \, \frac{\partial^{k+1} f(x_0 + sz)}{\partial x_{m_1}\cdots\partial x_{m_\kappa}\,\partial x_{m_{\kappa+1}}}\, ds \right) z_{m_1} \cdots z_{m_\kappa}\, z_{m_{\kappa+1}} \, .$$

Remarks and Examples. We study the example $f \in C^\infty(\mathbb{R}^2;\mathbb{R})$ given by

$$f(x,y) := \exp\left[\tfrac{1}{2}(x^2+y^2)\right] . \tag{IV.21}$$

Before taking derivatives, we observe that $f \geq 1$ and

$$\big\{ f(x,y) = 1 \big\} \Leftrightarrow \big\{ x^2 + y^2 = 0 \big\} . \tag{IV.22}$$

Next, we compute all first and second partial derivatives,

$$\frac{\partial f(x,y)}{\partial x} = x \exp\left[\tfrac{1}{2}(x^2+y^2)\right] , \quad \frac{\partial f(x,y)}{\partial y} = y \exp\left[\tfrac{1}{2}(x^2+y^2)\right] , \tag{IV.23}$$

and

$$\frac{\partial^2 f(x,y)}{\partial x^2} = (x^2+1) \exp\left[\tfrac{1}{2}(x^2+y^2)\right] , \tag{IV.24}$$

$$\frac{\partial^2 f(x,y)}{\partial y^2} = (y^2+1) \exp\left[\tfrac{1}{2}(x^2+y^2)\right] , \tag{IV.25}$$

$$\frac{\partial^2 f(x,y)}{\partial x\,\partial y} = \frac{\partial^2 f(x,y)}{\partial y\,\partial x} = xy \exp\left[\tfrac{1}{2}(x^2+y^2)\right] . \tag{IV.26}$$

For $(x_0, y_0) = (0,0)$ and $z = (z_1, z_2)$, Theorem IV.4 with $k = 1$ yields

$$\begin{aligned} f(z_1,z_2) &= f(0,0) + z_1 \frac{\partial f(0,0)}{\partial x} + z_2 \frac{\partial f(0,0)}{\partial y} + z_1^2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x^2}\, ds \right) \\ &\quad + z_2^2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial y^2}\, ds \right) + 2 z_1\, z_2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x\,\partial y}\, ds \right) \\ &= 1 + z_1^2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x^2}\, ds \right) + z_2^2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial y^2}\, ds \right) \\ &\quad + 2 z_1\, z_2 \left(\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x\,\partial y}\, ds \right) . \end{aligned} \tag{IV.27}$$

Now we assume that $z_1^2 + z_2^2 < r^2 \leq \frac{1}{4}$ and estimate

$$\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x^2}\, ds \;=\; \int_0^1 (1-s)\,(z_1^2+1)\, e^{s^2(z_1^2+z_2^2)/2}\, ds \;\geq\; \int_0^1 (1-s)\, ds \;=\; \frac{1}{2}\,, \tag{IV.28}$$

from below and

$$\begin{aligned}\int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x^2}\, ds \;&=\; \int_0^1 (1-s)\,(z_1^2+1)\, e^{s^2(z_1^2+z_2^2)/2}\, ds \\ &\leq (r^2+1)\, e^{r^2/2} \int_0^1 (1-s)\, ds \;=\; \frac{1}{2}(r^2+1)\, e^{r^2/2} \\ &\leq \frac{1}{2} + \frac{1}{2}\big(e^{r^2/2}-1\big) + \frac{r^2}{2} e^{r^2/2} \;\leq\; \frac{1}{2} + \sqrt{e} r^2\,, \end{aligned} \tag{IV.29}$$

from above. A similar estimate for the partial derivative w.r.t. y holds true,

$$\frac{1}{2} \;\leq\; \int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial y^2}\, ds \;\leq\; \frac{1}{2} + \sqrt{e} r^2\,. \tag{IV.30}$$

On the other hand,

$$\begin{aligned}\bigg| \int_0^1 (1-s)\, \frac{\partial^2 f(sz_1, sz_2))}{\partial x\, \partial y}\, ds \bigg| \;&=\; |z_1|\,|z_2| \bigg| \int_0^1 s^2\,(1-s)\, e^{s^2(z_1^2+z_2^2)/2}\, ds \bigg| \\ &\leq r^2\, e^{r^2/2} \int_0^1 s^2(1-s)\, ds \;=\; \frac{1}{12} r^2\, e^{r^2/2} \;\leq\; \frac{r^2}{6}\,. \end{aligned} \tag{IV.31}$$

It follows that, for $z_1^2 + z_2^2 < r^2 \leq \frac{1}{4}$, we have

$$1 \;+\; \frac{1}{2}\big(z_1^2+z_2^2\big)\big[1 - \tfrac{1}{6} r^2\big] \;\leq\; f(z_1, z_2) \;\leq\; 1 \;+\; \frac{1}{2}\big(z_1^2+z_2^2\big)\big[1 + (\tfrac{1}{6} - \sqrt{e}) r^2\big]\,. \tag{IV.32}$$