

Mathematik für Ingenieure und Informatiker I

Frei nach Meyberg-Vachenauer "Höhere Mathematik I"
und dem Skript von Ruzicka und Wolke

KAPITEL 1

Zahlen und Vektoren

1.1. Mengen und Abbildungen

1.1.1. Mengen.

Sprechweise. Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterscheidbaren Objekten unserer Anschauung zu einem Ganzen.

Bezeichnungen. • Objekte von Mengen heißen *Elemente*.

$a \in M$ „ a ist Element der Menge M “

$a \notin M$ „ a ist nicht Element der Menge M “

• Schreibweisen:

$A = \{1, 2, 3\}$ Aufzählung der Elemente von A

$B = \{x \in \mathbb{Z} \mid 1 \leq x \leq 3\}$ Aussonderung aus der Menge \mathbb{Z}

$C = \{x^2 \mid x \in \mathbb{Z}\}$ Konstruktion aus der Menge \mathbb{Z}

• Zwei Mengen sind *gleich*, wenn sie die gleichen Elemente enthalten. Es kommt weder auf die Reihenfolge der Elemente an, noch darauf, wie oft sie angegeben werden.

$$\{1, 2, 3\} = \{2, 3, 2, 1\}.$$

- Die *leere Menge* $\emptyset = \{ \}$ enthält kein Element.
- Eine Menge B heißt *Teilmenge* von A , geschrieben $B \subset A$, wenn für alle $b \in B$ auch $b \in A$ gilt. BSP.: $\emptyset \subset A$, $A \subset A$. Eine Menge B heißt *echte Teilmenge* von A , geschrieben $B \subsetneq A$, wenn $B \subset A$ und nicht $B = A$ gilt. Dann existiert immer ein $a \in A$ mit $a \notin B$.

Bemerkung. Die obige Definition von Mengen ist zu ungenau. Beispielsweise kann

$$M = \{A \text{ Menge} \mid A \neq A\}$$

keine Menge sein, denn $M \in M$ gilt genau dann, wenn $M \neq M$ gilt. Bei den „Mengen“ im Verlauf dieser Vorlesung wird dieses Problem aber keine Rolle spielen.

1.1.2. Aussagen.

Definition. Eine *Aussage* ist ein sinnvoller Satz unserer Sprache, der entweder *wahr* oder *falsch* ist.

Seien S, T Aussagen. Man erhält neue Aussagen:

„nicht S “ (*Negation*), „ S und T “ (*Konjunktion*), „ S oder T “ (*Alternative*), „wenn S , dann T “ (*Implikation*, $S \Rightarrow T$) „genau dann S , wenn T “ (*Äquivalenz*, $S \Leftrightarrow T$).

Bemerkung. „wenn S , dann T “ ist gleichbedeutend zu „ T oder nicht S “.

Wenn A eine Menge ist, x eine Variable, und S eine Aussage, bilde

$\forall x \in A : S$ „für alle $x \in A$ gilt S “ (*Allquantor*)

$\exists x \in A : S$ „es gibt ein $x \in A$, für das S gilt“ (*Existenzquantor*)

HINWEIS: Exzessive Verwendung logischer Symbole macht einen Text unleserlich.

1.1.3. Verknüpfungen von Mengen. Es seien A, B Mengen. Man definiert

$$A \cup B = \{x \mid x \in A \text{ oder } x \in B\}$$

$$A \cap B = \{x \mid x \in A \text{ und } x \in B\}$$

$$A \setminus B = \{x \in A \mid x \notin B\}$$

Das *kartesische Produkt* zweier Mengen wird konstruiert als Menge von *Paaren* bzw. *h -Tupeln*.

$$A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\}$$

$$A_1 \times \cdots \times A_n = \{(a_1, \dots, a_n) \mid a_1 \in A_1, \dots, a_n \in A_n\}$$

$$A^n = \underbrace{A \times \cdots \times A}_{n \text{ Faktoren}}$$

Es gilt $(a, b) = (c, d)$ genau dann, wenn $a = c$ und $b = d$. Es gilt $(a_1, \dots, a_m) = (b_1, \dots, b_n)$ genau dann, wenn $m = n$ und $a_1 = b_1, \dots, a_n = b_n$.

Die *Potenzmenge* einer Menge A ist die Menge aller Teilmengen.

$$\text{Pot}(A) = \{B \text{ Menge} \mid B \subset A\}.$$

1.1.4. Abbildungen.

Es seien A und B Mengen.

Definition. Eine *Abbildung* (oder *Funktion*) $f: A \rightarrow B$ ordnet jedem $a \in A$ genau ein $b \in B$ zu, schreibe $f(a) = b$. Zwei Funktionen $f: A \rightarrow B, g: C \rightarrow D$ sind gleich, wenn $A = C, B = D$ und $f(a) = g(a)$ für alle $a \in G$ gilt, schreibe $f = g$.

Bezeichnungen. Es sei $f: A \rightarrow B$ eine Abbildung, dann heisst A der *Definitionsbereich* von f , B der *Wertebereich* von f (ACHTUNG: anders als bei Meyberg-Vachenauer). Sei $f(a) = b$, dann ist b der *Wert* (das *Bild*) von f an der *Stelle* (am *Punkt*, am *Argument*) a .

Sei $C \subset A$, dann heisst

$$f(C) = \{f(a) \mid a \in C\}$$

das *Bild* von C unter f , und $f(A) \subset B$ das *Bild* von f .

Sei $D \subset B$, dann heisst

$$f^{-1}(D) = \{a \in A \mid f(a) \in D\}$$

das *Urbild* von D unter f . Es gilt $f^{-1}(B) = A$.

Definition. Eine Abbildung $f: A \rightarrow B$ heisst

- *injektiv* \Leftrightarrow Für alle $a, b \in A$ folgt aus $f(a) = f(b)$, dass $a = b$.
- *surjektiv* \Leftrightarrow Für alle $b \in B$ existiert $a \in A$ mit $f(a) = b$.
- *bijektiv* \Leftrightarrow f ist injektiv und surjektiv.

Bezeichnungen. • Sei A eine Menge. Die Abbildung $id_A: A \rightarrow A$ mit $id_A(a) = a$ für alle $a \in A$ heisst die *Identität* auf A .

- Sei $f: A \rightarrow B$ bijektiv. Definiere die *Umkehrabbildung* $f^{-1}: B \rightarrow A$ so, dass $f^{-1}(b)$ für alle $b \in B$ das eindeutig bestimmte $a \in A$ ist mit $f(a) = b$.
- Seien $f: A \rightarrow B, g: B \rightarrow C$ Abbildungen, dann definiere die *Verkettung* $g \circ f: A \rightarrow C$ durch

$$g \circ f(a) = g(f(a)) \in C \quad \text{für alle } a \in A.$$

Sei $f: A \rightarrow B$ bijektiv, dann gilt $f^{-1} \circ f = id_A$ und $f \circ f^{-1} = id_B$.

- Sei $C \subset A$, dann heisst $f|_C: C \rightarrow B$ mit $f|_C(c) = f(c)$ für alle $c \in C$ die *Einschränkung* von f auf C .
- Sei $D \subset B$ Teilmenge mit $f(A) \subset D$. Dann können wir den Wertebereich von $f: A \rightarrow B$ einschränken, und die Funktion $f: A \rightarrow D$ betrachten.

1.2. Zahlen

1.2.1. Natürliche Zahlen und vollständige Induktion.

Bezeichnungen.

$$\mathbb{N} = \mathbb{N}_{>} = \{1, 2, 3, \dots\} \quad \text{natürliche Zahlen}$$

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$$

$$\begin{aligned} \mathbb{Z} &= \{n \mid n \in \mathbb{N}_0 \text{ oder } -n \in \mathbb{N}\} \\ &= \{\dots, -1, 0, 1, \dots\} \quad \text{ganze Zahlen} \end{aligned}$$

Es sei $A(n)$ eine Aussage, die von einer Zahl $n \in \mathbb{Z}$ abhängt.

Prinzip der vollständigen Induktion: Die Aussage $A(n)$ gilt genau dann für alle $n \in \mathbb{N}_0$, wenn

- 1) $A(0)$ wahr ist, und
- 2) $A(n+1)$ aus $A(n)$ folgt für alle $n \in \mathbb{N}_0$.

Allgemeiner: $A(n)$ gilt für alle $n \in \mathbb{Z}$ mit $n \geq n_0 \in \mathbb{Z}$, wenn

- 1) $A(n_0)$ wahr ist, und
- 2) $A(n+1)$ aus $A(n)$ folgt für alle $n \in \mathbb{Z}$, $n \geq n_0$.

Rekursive Definitionen: Sei M eine Menge. Um eine Funktion $f: \mathbb{N}_0 \rightarrow M$ anzugeben, reicht es

- 1) $f(0) \in M$ anzugeben, und
- 2) für alle $n \in \mathbb{N}_0$ den Wert $f(n+1) \in M$ mit Hilfe von $n+1$ und $f(n)$ anzugeben.

Beispiel. • *Potenzen* a^n sind definiert durch

$$a^0 = 1 \text{ und } a^{n+1} = a^n \cdot a.$$

- Die *Fakultät* $n!$ ist definiert durch

$$0! = 1 \text{ und } (n+1)! = (n+1) \cdot n!$$

1.2.1. Satz. *Es sei M eine Menge mit verschiedenen Elementen. Dann gibt es genau $n!$ Möglichkeiten, die Elemente von M aufzuzählen.*

Beispiel. Es seien $a_i \in \mathbb{R}$ für alle $i \in \mathbb{N}$. Definiere endliche *Summen* und *Produkte* durch

$$\begin{aligned} \sum_{i=1}^0 a_i = 0 & \quad \text{und} & \quad \sum_{i=1}^{n+1} a_i = a_{n+1} + \sum_{i=1}^n a_n, \\ \prod_{i=1}^0 a_i = 1 & \quad \text{und} & \quad \prod_{i=1}^{n+1} a_i = a_{n+1} \cdot \prod_{i=1}^n a_n. \\ \Rightarrow \sum_{i=1}^n a_i = a_1 + \cdots + a_n, & & \quad \prod_{i=1}^n a_i = a_1 \cdots a_n. \end{aligned}$$

Unendliche Summen und Produkte lassen sich so nicht definieren.

1.2.2. Rationale und reelle Zahlen.

Bezeichnungen.

$$\begin{aligned} \mathbb{Q} &= \left\{ \frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N} \right\} && \text{rationale Zahlen} \\ \mathbb{R} &= ? \supset \mathbb{Q} && \text{reelle Zahlen} \end{aligned}$$

Wir führen die reellen Zahlen axiomatisch ein als Menge \mathbb{R} mit Elementen $0, 1$, Verknüpfungen $+, \cdot : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ (d.h. $+(a, b) = a + b$, $\cdot(a, b) = ab$) und der Relation $<$. Es müssen Körper-, Ordnungs- und Vollständigkeitsaxiome gelten.

Körperaxiome. für alle $a, b \in \mathbb{R}$ muss gelten:

$(a + b) + c = a + (b + c)$	Assoziativität der Addition
$a + 0 = a$	Neutrales Element der Addition
$a + (-a) = 0$	Negatives Element
$a + b = b + a$	Kommutativität der Addition
$(a \cdot b) \cdot c = a \cdot (b \cdot c)$	Assoziativität der Multiplikation
$a \cdot 1 = a$	Neutrales Element der Multiplikation
$a \cdot \frac{1}{a} = 1$ falls $a \neq 0$	Inverses Element
$a \cdot b = b \cdot a$	Kommutativität der Multiplikation
$a \cdot (b + c) = a \cdot b + a \cdot c$	Distributivgesetz
$0 \neq 1$	

Bezeichnungen. Seien $x, y \in \mathbb{R}$. Schreibe

$$\begin{aligned} x > 0 & \quad \text{für } y < x, \\ x \leq y & \quad \text{falls } x < y \text{ oder } x = y, \\ x \geq y & \quad \text{falls } y \leq x. \end{aligned}$$

Ordnungsaxiome.

- Für $x, y \in \mathbb{R}$ gilt genau eine der drei Aussagen

$$x < y, \quad x = y \quad \text{oder} \quad x > y$$

- Aus $x \leq y$ und $y \leq z$ folgt $x \leq z$
- Aus $x \leq y$ und $z \leq w$ folgt $x + y \leq z + w$
- Aus $x \leq y$ und $0 \leq z$ folgt $x \cdot z \leq y \cdot z$
- Für jede Zahl $x \in \mathbb{R}$ existiert $n \in \mathbb{N}$, so dass

$$x \leq n = 1 + \dots + 1 \in \mathbb{R}. \quad (\text{Archimedisches Axiom})$$

Folgerung. • Aus $x \leq y$ folgt $-y \leq -x$.

- Aus $0 < x \leq y$ folgt $0 < \frac{1}{y} \leq \frac{1}{x}$.

Definition. Eine Teilmenge $S \subset \mathbb{R}$ heisst *nach oben beschränkt*, wenn es $s \in \mathbb{R}$ gibt mit $x \leq s$ für alle $x \in S$, und s heisst dann *obere Schranke* von S .

Sei s obere Schranke und $t \geq s$, dann ist auch t obere Schranke.

Vollständigkeitsaxiom. Jede nicht leere, nach oben beschränkte Teilmenge $S \subset \mathbb{R}$ besitzt eine kleinste obere Schranke.

Mit diesen Axiomen sind die reellen Zahlen eindeutig beschrieben.

Definition. Die kleinste obere Schranke heisst *Supremum*, die größte untere Schranke heisst *Infimum*.

Beispiel. • $\sup\{x \in \mathbb{Q} \mid x^2 \leq 2\} = \sqrt{2}$.

- $\inf\{\frac{1}{n} \mid n \in \mathbb{N}\} = 0$.

1.2.3. Intervalle, Betrag.

Bezeichnungen. Für $a, b \in \mathbb{R}$ mit $a \leq b$ schreibe

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad \text{abgeschlossenes Intervall}$$

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\} \quad \text{offenes Intervall}$$

analog $[a, b)$, $(a, b]$ halboffene Intervalle.

Mit dem Symbol „ ∞ “ definiert man

$$(-\infty, a] = \{x \in \mathbb{R} \mid x \leq a\}$$

$$(-\infty, a) = \{x \in \mathbb{R} \mid x < a\}$$

und analog $[a, \infty)$, (a, ∞) und $(-\infty, \infty) = \mathbb{R}$.

Definition. Der Betrag einer reellen Zahl $x \in \mathbb{R}$ ist definiert durch

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

Rechenregeln.

$$\begin{aligned}
& -|x| \leq x \leq |x| \\
& | -x | = |x| \\
& |xy| = |x| |y|, \\
& \left| \frac{x}{y} \right| = \frac{|x|}{|y|} \quad \text{falls } y \neq 0, \\
& |x| \leq y \quad \Longrightarrow \quad -y \leq x \leq y \\
& |x + y| \leq |x| + |y|
\end{aligned}$$

*Dreiecksungleichung.***1.2.4. Die binomische Formel.**

Definition. Es seien $k, n \in \mathbb{N}_0$ mit $0 \leq k \leq n$. Dann definiere den *Binomialkoeffizienten* $\binom{n}{k} \in \mathbb{N}$ (lies „ n über k “) durch

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

Es gilt $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$ und $\binom{n}{k} = \binom{n}{n-k}$.

1.2.2. Satz. Eine n -elementige Menge besitzt genau $\binom{n}{k}$ verschiedene Teilmengen mit genau k Elementen.

1.2.3. Satz. Es gilt

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

1.3. Geometrie der Ebene, komplexe Zahlen

1.3.1. Kartesische Koordinaten. In der Ebene E (Zeichenebene, Tafel Ebene) wähle

- Ursprung $O = (0, 0) \in E$,
- Zwei Geraden durch O im rechten Winkel, die x -Achse und die y -Achse,
- Zwei Punkte $(1, 0)$ auf der x -Achse, $(0, 1)$ auf der y -Achse, im gleichen Abstand von O .

Dann lässt sich jeder Punkt P in E als Paar (x, y) schreiben. Dazu fälle das Lot von P auf die x - und die y -Achse.

Wir können Punktfolgen in E als Teilmengen von \mathbb{R}^2 darstellen.

Beispiel. $I \subset \mathbb{R}$ Intervall, $f: I \rightarrow \mathbb{R}$ Funktion. Der *Graph* von f ist die Menge $G_f := \{(x, y) \in \mathbb{R}^2 \mid x \in I, y = f(x)\}$.

Beispiel. Sei $C \subset E$ eine Kurve. Wähle kartesische Koordinaten. Falls $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Funktion ist mit

$$C = \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\} = F^{-1}(\{0\})$$

heißt $F(x, y) = 0$ eine *Gleichung* der Kurve C .

a) Seien $P = (a, b)$, $Q = (c, d) \in \mathbb{R}^2$ mit $P \neq Q$. Die Gerade durch P und Q hat Gleichung

$$(d - b)(x - a) - (c - a)(y - b) = 0.$$

b) Kreis um (a, b) mit Radius $R > 0$

$$(x - a)^2 + (y - b)^2 - R^2 = 0.$$

c) Nicht jedes F liefert eine Kurve, Bsp.: $x \cdot y = 0$

1.3.2. Vektoren und Verschiebungen. Zu zwei Punkten $P, Q \in E$ gibt es genau eine Parallelverschiebung, die P in Q überführt. Wir schreiben \vec{PQ} für den „Pfeil“ von P nach Q . Wenn die obige Parallelverschiebung den Punkt R nach S verschiebt, haben \vec{PQ} und \vec{RS} die gleiche Länge und die gleiche Richtung.

Definition. Ein Pfeil \vec{PQ} heißt *Vektor*. Zwei Vektoren heißen *gleich*, wenn sie die gleiche Länge und Richtung haben.

Bemerkung. Vektoren stellen Parallelverschiebungen dar. Zwei Vektoren sind gleich, wenn sie die gleiche Verschiebung beschreiben.

Rechnen mit Vektoren. Addition: Seien \vec{v}, \vec{w} zwei Vektoren. Dann bezeichnet $\vec{v} + \vec{w}$ die Hintereinanderausführung der zugehörigen Verschiebungen. Es gilt $\vec{v} + \vec{w} = \vec{w} + \vec{v}$.

Skalare Vielfache: Sei $v \in \mathbb{R}$, dann bezeichnet $r\vec{v}$ eine Verschiebung in Richtung \vec{v} um die r -fache Länge.

Es sei $\vec{v} = \vec{PQ}$, dann ist $-\vec{v} = \vec{QP}$ der *entgegengesetzte Vektor*.

Der Vektor $\vec{0} = \vec{PP} = \vec{QQ}$ heißt *Nullvektor*.

Rechenregeln. Es seien $\vec{u}, \vec{v}, \vec{w}$ Vektoren und $r, s \in \mathbb{R}$.

$$\begin{array}{ll}
 (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}) & \text{Assoziativit\u00e4t} \\
 \vec{v} + \vec{0} = \vec{v} & \text{Nullvektor} \\
 \vec{v} + (-\vec{v}) = \vec{0} & \text{entgegengesetzter Vektor} \\
 \vec{v} + \vec{w} = \vec{w} + \vec{v} & \text{Kommutativit\u00e4t} \\
 r(s\vec{v}) = (rs)\vec{v} & \text{Vertr\u00e4glichkeit der Multiplikation} \\
 r(\vec{v} + \vec{w}) = r\vec{v} + r\vec{w} & 1. \text{ Distributivit\u00e4tsgesetz} \\
 (r + s)\vec{v} = r\vec{v} + s\vec{v} & 2. \text{ Distributivit\u00e4tsgesetz} \\
 1\vec{v} = \vec{v} & \text{Wirkung der Eins}
 \end{array}$$

Folgerung. Es gilt

$$0 \cdot \vec{v} = \vec{0} \quad \text{und} \quad (-1) \cdot \vec{v} = -\vec{v}.$$

1.3.3. Vektoren in Koordinaten. Es sei $P = (x, y)$, $Q = (z, w)$, $R = (a, b)$ und $S = (c, d)$ in kartesischen Koordinaten. Dann gilt $\vec{PQ} = \vec{RS}$ genau dann, wenn $z - x = c - a$ und $w - y = d - b$.

Definition. Der Vektor $\vec{v} = \vec{PQ}$ hat die *Koordinaten* $\begin{pmatrix} z-x \\ w-y \end{pmatrix}$. Der Vektor \vec{OP} mit den Koordinaten $\begin{pmatrix} x \\ y \end{pmatrix}$ heisst *Ortsvektor* von P .

Folgerung. Seien $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ Vektoren und $v \in \mathbb{R}$. Dann gilt:

$$\begin{aligned}
 \vec{v} + \vec{w} &= \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}, \quad -\vec{v} = \begin{pmatrix} -v_1 \\ -v_2 \end{pmatrix}, \quad \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\
 r \cdot \vec{v} &= \begin{pmatrix} rv_1 \\ rv_2 \end{pmatrix}.
 \end{aligned}$$

Definition. Ein Vektor $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ hat die *L\u00e4nge* (den *Betrag*)

$$|\vec{v}| = \sqrt{v_1^2 + v_2^2}.$$

Rechenregeln. Seien \vec{v}, \vec{w} Vektoren und $r \in \mathbb{R}$. Dann gilt

$$\begin{array}{ll}
 |\vec{v}| \geq 0 \text{ und } |\vec{v}| = 0 \iff \vec{v} = \vec{0} & \text{Positivit\u00e4t} \\
 |r \cdot \vec{v}| = |r| \cdot |\vec{v}| & \text{Homogenit\u00e4t} \\
 |\vec{v} + \vec{w}| \leq |\vec{v}| + |\vec{w}| & \text{Dreiecksungleichung}
 \end{array}$$

1.3.4. Winkel und Winkelfunktionen. Ein Winkel entsteht durch Drehung eines Vektors \vec{PQ} um den Punkt P . Wir messen Winkel als L\u00e4nge des zugeh\u00f6rigen Bogens auf dem Kreis um P mit Radius 1 (*Bogenma\u00df*). Drehungen entgegen dem Uhrzeigersinn entsprechen positiven Winkeln, Drehungen im Uhrzeigersinn entsprechen negativen Winkeln. Ein Halbkreisbogen hat die L\u00e4nge $\pi = 3,14159\dots$

Folgerung.

$$\pi = 180^\circ, \quad 1^\circ = \frac{\pi}{180}, \quad \frac{\pi}{2} = 90^\circ \quad (\text{rechter Winkel}).$$

Definition. Die *Winkelfunktionen* $\cos, \sin : \mathbb{R} \rightarrow [-1, 1] \subset \mathbb{R}$ sind so definiert, dass eine Drehung des Punktes $(1, 0)$ um den Punkt $O = (0, 0)$ mit dem Winkel α den Punkt $(\cos \alpha, \sin \alpha)$ ergibt.

Drehung von $(1, 0)$ um den rechten Winkel $\frac{\pi}{2}$ liefert $(0, 1)$, also

$$\cos \frac{\pi}{2} = 0, \quad \sin \frac{\pi}{2} = 1.$$

1.3.1. Satz. Eine Drehung der Ebene um 0 mit dem Winkel α überführt den Punkt (x, y) in (x', y') , mit

$$\begin{aligned} x' &= x \cos \alpha - y \sin \alpha \\ \text{und} \quad y' &= x \sin \alpha + y \cos \alpha . \end{aligned}$$

1.3.2. Satz. Sei P ein Punkt der Ebene mit den Koordinaten (x, y) . Dreht man das Koordinatensystem um den Winkel α , so hat P im neuen Koordinatensystem die Koordinaten (x', y') , mit

$$\begin{aligned} x &= x' \cos \alpha - y' \sin \alpha , \\ y &= x' \sin \alpha + y' \cos \alpha , \\ \text{bzw.} \quad x' &= x \cos \alpha + y \sin \alpha , \\ y' &= -x \sin \alpha + y' \cos \alpha . \end{aligned}$$

1.3.5. Komplexe Zahlen.

Definition. Die Menge $\mathbb{C} = \mathbb{R}^2$ heisst Menge der *komplexen Zahlen*. Wir definieren $0 = (0, 0)$, $1 = (1, 0)$ und $i = (0, 1)$. Schreibe $(a, b) = a + bi$ für $a, b \in \mathbb{R}$. *Komplexe Addition* und *Multiplikation* sind definiert durch

$$\begin{aligned} (a + bi) + (c + di) &= (a + c) + (b + d)i , \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i . \end{aligned}$$

Real- und *Imaginäranteil* sind definiert als

$$\operatorname{Re}(a + bi) = a, \quad \operatorname{Im}(a + bi) = b.$$

Bemerkung. Addition von komplexen Zahlen ist gerade die Addition von Vektoren. Multiplikation mit $r + 0 \cdot i = r \in \mathbb{R}$ entspricht der Multiplikation eines Vektors mit einem Skalar, Multiplikation mit i liefert Drehung um einen rechten Winkel:

$$i \cdot (x + iy) = -y + ix.$$

Multiplikation mit $\cos \alpha + i \sin \alpha$ entspricht Drehung um den Winkel α , vgl. Satz 1.3.1

$$(\cos \alpha + i \sin \alpha)(x + yi) = (x \cos \alpha - y \sin \alpha) + (y \cos \alpha + x \sin \alpha)i$$

Definition. Sei $z = x + yi$ mit $x, y \in \mathbb{R}$ eine komplexe Zahl, dann heißt $\bar{z} = x - yi$ die *konjugierte* komplexe Zahl, und $|z| = \sqrt{z \cdot \bar{z}}$ heißt ihr Betrag.

Bemerkung. Es gilt

$$z \cdot \bar{z} = (x + iy)(x - iy) = x^2 + y^2 \in \mathbb{R}$$

und $z \cdot \bar{z} \geq 0$. Für $z = x + yi \neq 0$ folgt

$$z \cdot \frac{\bar{z}}{|z|^2} = \frac{z \cdot \bar{z}}{z \cdot \bar{z}} = 1 \quad \implies \quad \frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

1.3.3. Satz. \mathbb{C} ist ein Körper, das heisst, es gelten die Körperaxiome aus Abschnitt 1.2.2. Insbesondere gilt für $z, v, w \in \mathbb{C}$:

$$\begin{array}{ll} (z + v) + w = z + (v + w) & (z \cdot v) \cdot w = z \cdot (v \cdot w) \\ z + 0 = z & z \cdot 1 = z \\ z + (-z) = 0 & z \cdot \frac{1}{z} = 1 \quad \text{für } z \neq 0 \\ z + w = w + z & z \cdot w = w \cdot z \\ z \cdot (v + w) = zv + zw & 0 \neq 1 \end{array}$$

BEWEIS. z.B. Assoziativität der Multiplikation. Es seien $a, b, c, d, e, f \in \mathbb{R}$. Dann gilt

$$\begin{aligned} ((a + bi)(c + di))(e + fi) &= ((ac - bd) + (ad + bc)i)(e + fi) \\ &= (ace - bde - adf - bcf) + (acf - bdf + ade + bce)i \end{aligned}$$

Auf der anderen Seite

$$\begin{aligned} (a + bi)((c + di)(e + fi)) &= (a + bi)((ce - df) + (cf + de)i) \\ &= (ace - adf - bcf - bde) + (acf + ade + bce - bdf)i \\ &= ((a + bi)(c + di))(e + fi) \end{aligned}$$

Alle anderen Rechenregeln führt man genauso auf Rechnungen in \mathbb{R} zurück. \square

Jede Zahl $z \in \mathbb{C} \setminus \{0\}$ lässt sich schreiben als

$$z = r(\cos \varphi + i \sin \varphi) \quad \text{mit } \varphi \in [0, 2\pi) \text{ und } r = |z|.$$

Definition. Es heisst φ das *Argument* oder die *Phase* von z , $\varphi = \arg z$.

Bemerkung. Da Multiplikation mit $\cos \varphi + i \sin \varphi$ einer Drehung um den Winkel φ entspricht, folgt

$$\arg(z \cdot w) = \arg z + \arg w + 2\pi n \quad \text{mit } n \in \mathbb{Z}.$$

Folgerung. Jede Zahl $z = r(\cos \varphi - i \sin \varphi) \notin 0$ hat zwei Quadratwurzeln

$$\pm\sqrt{z} = \pm\sqrt{r} \left(\cos \frac{\varphi}{2} + i \sin \frac{\varphi}{2} \right).$$

Insbesondere hat die quadratische Gleichung $z^2 + pz + q = 0$ mit $p, q \in \mathbb{C}$ stets Lösungen

$$z = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \in \mathbb{C},$$

diese stimmen überein, falls $q = \frac{p^2}{4}$.

1.4. Dreidimensionale Geometrie

1.4.1. Kartesische Koordinaten im Raum. Im Raum wähle ein *kartesisches Koordinatensystem*, bestehend aus

- einem Punkt 0, dem *Ursprung*,
- drei paarweise aufeinander senkrecht stehenden Geraden durch 0, die *x*-, *y*- und *z*-Achse. Diese bilden ein *Rechtssystem*, wenn sie wie Daumen, Zeige- und Mittelfinger der rechten Hand angeordnet sind,
- drei Punkten $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$ auf der *x*-, *y*- bzw. *z*-Achse im gleichen Abstand von 0.

Die drei Ebenen durch je zwei der Achsen heissen *Koordinatenebenen*. Sei P ein Punkt im Raum. Man erhält die *kartesischen* Koordinaten bezüglich des eben gewählten Koordinatensystems, indem man die zu den Koordinatenebenen parallelen Ebenen durch P mit den Koordinatenachsen schneidet. Zu jedem Tripel $(x, y, z) \in \mathbb{R}^3$ existiert genau ein Punkt mit diesen Koordinaten, und umgekehrt.

1.4.2. Vektoren im Raum. Zu je zwei Punkten P, Q im Raum gibt es genau eine Parallelverschiebung \vec{PQ} , die P in Q abbildet, den *Vektor* \vec{PQ} . Hierfür gelten die gleichen Rechenregeln wie in Abschnitt 3.2.

Sei jetzt ein kartesisches Koordinatensystem gegeben. Dann heissen die Vektoren von 0 zu den Punkten $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$ die *Standardbasisvektoren*. Sie werden mit $\vec{e}_1, \vec{e}_2, \vec{e}_3$ (oder $\vec{i}, \vec{j}, \vec{k}$ oder $\vec{e}_x, \vec{e}_y, \vec{e}_z$) bezeichnet. Man nennt $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine *kartesische Basis*. Das kartesische Koordinatensystem lässt sich jetzt angeben als

$$(0, \vec{e}_1, \vec{e}_2, \vec{e}_3).$$

Jeder Vektor \vec{v} lässt sich eindeutig zerlegen als

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3,$$

schreibe

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Man nennt $v_i \vec{e}_i$ die *Komponente* von v in Richtung \vec{e}_i , und $v_i \in \mathbb{R}$ die *i-te Koordinate* von \vec{v} (bezüglich $(0, \vec{e}_1, \vec{e}_2, \vec{e}_3)$) für $i = 1, 2, 3$. Seien $P = (p_1, p_2, p_3)$, $Q = (q_1, q_2, q_3)$, dann gilt

$$\vec{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix} = \sum_{i=1}^3 (q_i - p_i) \vec{e}_i.$$

Für zwei Vektoren $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$ und einen *Skalar* $r \in \mathbb{R}$ gilt:

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix}, \quad r \cdot \vec{v} = \begin{pmatrix} rv_1 \\ rv_2 \\ rv_3 \end{pmatrix}$$

Der Nullvektor ist $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, wie in Abschnitt 3.3.

1.4.3. Winkel und Skalarprodukt. Seien $\vec{v} \neq \vec{0}$, $\vec{w} \neq \vec{0}$ zwei Vektoren. Trägt man beide am Ursprung an, so spannen sie eine Ebene E auf. In dieser Ebene bestimmen wir den *Winkel* $\angle(\vec{v}, \vec{w})$ wie in Abschnitt 3.4 als Länge des zugehörigen Bogens auf dem Einheitskreis. Dabei vereinbaren wir $0 \leq \angle(\vec{v}, \vec{w}) \leq \pi$. Beachte: Die Ebene E ist nicht orientiert, daher hat der Winkel kein Vorzeichen.

Definition. Zwei Vektoren \vec{v}, \vec{w} heißen *orthogonal*, wenn $\vec{v} = \vec{0}$, $\vec{w} = \vec{0}$ oder $\angle(\vec{v}, \vec{w}) = \frac{\pi}{2}$, schreibe $\vec{v} \perp \vec{w}$.

Beispiel. In einem kartesischen Koordinatensystem sind $\vec{0}, \vec{e}_1, \vec{e}_2, \vec{e}_3$ paarweise orthogonal.

Rechenregeln. Seien $\vec{v} \neq 0$, $\vec{w} \neq 0$ Vektoren, $r \in \mathbb{R} \setminus \{0\}$. Dann gilt

$$\begin{aligned} \angle(\vec{v}, \vec{w}) &= \angle(\vec{w}, \vec{v}), \\ \angle(\vec{v}, r\vec{v}) &= \begin{cases} 0 & \text{falls } r > 0, \\ \pi & \text{falls } r < 0, \end{cases} \\ \angle(\vec{v}, r\vec{w}) &= \begin{cases} \angle(\vec{v}, \vec{w}) & \text{falls } r > 0, \text{ und} \\ \pi - \angle(\vec{v}, \vec{w}) & \text{falls } r < 0. \end{cases} \end{aligned}$$

Definition. Das *Skalarprodukt* $\vec{v}, \vec{w} \in \mathbb{R}$ zweier Vektoren ist definiert als:

$$\vec{v}, \vec{w} = \begin{cases} |\vec{v}| |\vec{w}| \cos \angle(\vec{v}, \vec{w}) & \text{falls } \vec{v} \neq 0 \text{ und } \vec{w} \neq 0, \\ 0 & \text{sonst.} \end{cases}$$

Andere Schreibweisen: $\vec{v} \cdot \vec{w} = \vec{v}\vec{w} = \langle \vec{v}, \vec{w} \rangle = s(\vec{v}, \vec{w})$.

Beispiel. In einem kartesischen Koordinatensystem $(0, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ gilt

$$\vec{e}_1 \cdot \vec{e}_1 = \vec{e}_2 \cdot \vec{e}_2 = \vec{e}_3 \cdot \vec{e}_3 = 1 \text{ und } \vec{e}_1 \cdot \vec{e}_2 = \vec{e}_1 \cdot \vec{e}_3 = \vec{e}_2 \cdot \vec{e}_3 = 0$$

Für einen Vektor $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ gilt $v_i = \vec{v} \cdot \vec{e}_i$.

Rechenregeln. Für Vektoren $\vec{u}, \vec{v}, \vec{w}$ und $r \in \mathbb{R}$ gilt

$$\begin{aligned} \vec{v} \cdot \vec{w} &= \vec{w} \cdot \vec{v} && \text{Kommutativgesetz} \\ (r\vec{v}) \cdot \vec{w} &= \vec{v} \cdot (r\vec{w}) = r \cdot (\vec{v} \cdot \vec{w}) && \text{Homogenität} \\ (\vec{u} + \vec{v}) \cdot \vec{w} &= \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w} && \text{Distributivgesetz} \\ \vec{v} \cdot \vec{w} = 0 &\iff \vec{v} \perp \vec{w} && \text{Orthogonalitätstest} \\ |\vec{v}| &= \sqrt{\vec{v} \cdot \vec{v}} \end{aligned}$$

BEACHTEN: Es gilt im allgemeinen $\vec{u}(\vec{v} \cdot \vec{w}) \neq (\vec{u} \cdot \vec{v})\vec{w}$, denn der erste Vektor hat die Richtung von $\pm \vec{u}$, der zweite die von $\pm \vec{w}$.

In kartesischen Koordinaten seien

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

gegeben. Dann gilt

$$\begin{aligned} \vec{v} \cdot \vec{w} &= v_1 w_1 + v_2 w_2 + v_3 w_3, \\ \cos \angle(\vec{v}, \vec{w}) &= \frac{v_1 w_1 + v_2 w_2 + v_3 w_3}{\sqrt{v_1^2 + v_2^2 + v_3^2} \sqrt{w_1^2 + w_2^2 + w_3^2}}. \end{aligned}$$

Definition. Seien \vec{v}, \vec{w} Vektoren, $\vec{w} \neq 0$. Dann heisst $\vec{v}_{\vec{w}} := \frac{\vec{v} \cdot \vec{w}}{|\vec{w}|^2} \cdot \vec{w}$ die *Komponente* von \vec{v} in Richtung \vec{w} , und $\vec{v}_n = \vec{v} - \vec{v}_{\vec{w}}$ die zu \vec{w} *senkrechte Komponente* von \vec{v} .

Es gilt $|\vec{v}_{\vec{w}}| = |\vec{v}| \cos \angle(\vec{v}, \vec{w})$, $|\vec{v}_n| = |\vec{v}| \sin \angle(\vec{v}, \vec{w})$.

1.4.4. Kreuzprodukt und Spatprodukt.

Definition. Das *Kreuzprodukt* (*Vektorprodukt*) zweier Vektoren \vec{v}, \vec{w} im dreidimensionalen Raum ist der Vektor $\vec{v} \times \vec{w}$ mit den Eigenschaften

- $\vec{v} \times \vec{w} \perp \vec{v}, \vec{v} \times \vec{w} \perp \vec{w}$,
- $(\vec{v}, \vec{w}, \vec{v} \times \vec{w})$ bilden ein Rechtssystem,
- $|\vec{v} \times \vec{w}| = |\vec{v}| |\vec{w}| \sin \angle(\vec{v}, \vec{w})$ (Fläche des von \vec{v}, \vec{w} aufgespannten Parallelogramms)

Rechenregeln. Seien $\vec{u}, \vec{v}, \vec{w}$ Vektoren und $r \in \mathbb{R}$, dann gilt

$$\begin{aligned} \vec{v} \times \vec{v} &= \vec{0}, & \vec{w} \times \vec{v} &= -\vec{v} \times \vec{w} && \text{Antikommutativgesetz} \\ (r\vec{v}) \times \vec{w} &= \vec{v} \times (r\vec{w}) = r(\vec{v} \times \vec{w}) && \text{Homogenität} \\ \vec{u} \times (\vec{v} + \vec{w}) &= \vec{u} \times \vec{v} + \vec{u} \times \vec{w} && \text{Distributivgesetz} \\ (\vec{u} + \vec{v}) \times \vec{w} &= \vec{u} \times \vec{w} + \vec{v} \times \vec{w} \\ \vec{v} \times \vec{w} &= \vec{0} \iff \vec{v} \text{ ist Vielfaches von } \vec{w} \text{ oder } \vec{w} \text{ ist Vielfaches von } \vec{v} && \text{Parallelitätstest} \end{aligned}$$

$$|\vec{v} \times \vec{w}|^2 = |\vec{v}|^2 |\vec{w}|^2 - (\vec{v} \cdot \vec{w})^2 .$$

In kartesischen Koordinaten gilt

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} = \begin{vmatrix} v_1 & w_1 & \vec{e}_1 \\ v_2 & w_2 & \vec{e}_2 \\ v_3 & w_3 & \vec{e}_3 \end{vmatrix} .$$

Es gilt $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}$.

Definition. Das Spatprodukt von drei Vektoren $\vec{u}, \vec{v}, \vec{w}$ ist definiert als

$$[\vec{u}, \vec{v}, \vec{w}] = \vec{u} \cdot (\vec{v} \times \vec{w}) \in \mathbb{R} .$$

1.4.1. Satz. Das Volumen des von drei Vektoren $\vec{u}, \vec{v}, \vec{w}$ aufgespannten Spats (Parallelotops, Parallelepipeds) ist

$$V = |[\vec{u}, \vec{v}, \vec{w}]| .$$

Rechenregeln. Für Vektoren $\vec{u}, \vec{v}, \vec{w}, \vec{x}$ und $r \in \mathbb{R}$ folgt

$$\begin{aligned} [\vec{u}, \vec{v}, \vec{w}] &= -[\vec{v}, \vec{u}, \vec{w}] = [\vec{v}, \vec{w}, \vec{u}] = \dots && \text{Antisymmetrie} \\ [r\vec{u}, \vec{v}, \vec{w}] &= [\vec{u}, r\vec{v}, \vec{w}] = [\vec{u}, \vec{v}, r\vec{w}] = r[\vec{u}, \vec{v}, \vec{w}] && \text{Homogenität} \\ [\vec{u} + \vec{v}, \vec{w}, \vec{x}] &= [\vec{u}, \vec{w}, \vec{x}] + [\vec{v}, \vec{w}, \vec{x}] \\ [\vec{u}, \vec{v} + \vec{w}, \vec{x}] &= [\vec{u}, \vec{v}, \vec{x}] + [\vec{u}, \vec{w}, \vec{x}] && \text{Distributivgesetz} \\ [\vec{u}, \vec{v}, \vec{w} + \vec{x}] &= [\vec{u}, \vec{v}, \vec{w}] + [\vec{u}, \vec{v}, \vec{x}] \end{aligned}$$

Darstellung in kartesischen Koordinaten:

$$\left[\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right] = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} \\ = u_1 v_2 w_3 + u_2 v_3 w_1 + u_3 v_1 w_2 - u_1 v_3 w_2 - u_2 v_1 w_3 - u_3 v_2 w_1.$$

KAPITEL 2

Funktionen, Grenzwerte, Stetigkeit

In diesem Kapitel werden grundlegende Begriffe wie Funktionen und Grenzwerte eingeführt. Unter anderem werden Standardbeispiele für Funktionen, wie Polynome, Kreisfunktionen und die Exponentialfunktion, diskutiert. Der Grenzwertbegriff wird an Hand von Zahlenfolgen und Funktionen genauer betrachtet.

2.1. Grundbegriffe

2.1.1. Funktionen. In Abschnitt 1.1.4 wurden Funktionen für allgemeine Mengen definiert. Nun betrachten wir den Spezialfall einer reellen Funktion einer Veränderlichen.

$$f: D \subseteq \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x).$$

Beispiele. 1. Eine *lineare Funktion* ist gegeben durch

$$f(x) = ax + b$$

mit dem Definitionsbereich $D(f) = \mathbb{R}$.

2. Eine *quadratische Funktion* ist definiert als

$$f(x) = ax^2 + bx + c,$$

wobei $a \neq 0$ und der Definitionsbereich $D(f) = \mathbb{R}$ ist.

3. Die *Wurzelfunktion* ist gegeben durch

$$f(x) = \sqrt{x}$$

mit dem Definitionsbereich $D(f) = \mathbb{R}_0^+ = \{x \in \mathbb{R}, x \geq 0\}$.

Für $x > 0$ ist als \sqrt{x} immer die positive Quadratwurzel zu nehmen.

Definition. Sei $D \subseteq \mathbb{R}$ symmetrisch zum Nullpunkt, d.h. $x \in D \Rightarrow -x \in D$. Eine Funktion $f: D \rightarrow \mathbb{R}$ heisst *gerade* (bzw. *ungerade*) wenn $f(-x) = f(x)$ (bzw. $f(-x) = -f(x)$) für alle $x \in D$ gilt.

2.1.2. Monotonie.

Definition. Man nennt eine Funktion $f: D \rightarrow \mathbb{R}$

- a) *monoton fallend* (bzw. *monoton wachsend*), wenn für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ die Ungleichung $f(x_1) \geq f(x_2)$ (bzw. $f(x_1) \leq f(x_2)$) gilt.

- b) *strikt monoton fallend* (bzw. *strikt monoton wachsend*), wenn für alle $x_1 < x_2 \in D$ die strikte Ungleichung $f(x_1) > f(x_2)$ (bzw. $f(x_1) < f(x_2)$) gilt.

2.1.3. Rechnen mit Funktionen.

Definition. Seien $f, g : D \rightarrow \mathbb{R}$ zwei Funktionen. Dann definiert man die Summe, die Differenz, das Produkt und den Quotienten dieser Funktionen durch:

$$\begin{aligned}(f \pm g)(x) &:= f(x) \pm g(x) \\ (f \cdot g)(x) &:= f(x) \cdot g(x) \\ \left(\frac{f}{g}\right)(x) &:= \frac{f(x)}{g(x)}, \quad \text{falls } g(x) \neq 0\end{aligned}$$

d.h. die Operationen werden punktweise ausgeführt.

Definition. Zu Funktionen $f : I \rightarrow \mathbb{R}$, $g : D \rightarrow \mathbb{R}$ mit $g(D) \subset I$ kann man die *Komposition* $f \circ g : D \rightarrow \mathbb{R}$ bilden. Sie ist definiert durch:

$$(f \circ g)(x) := f(g(x)) .$$

2.2. Polynome und rationale Funktionen

2.2.1. Polynome. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heisst *Polynom vom Grad n* , wenn es Zahlen $a_0, \dots, a_n \in \mathbb{R}$ gibt mit $a_n \neq 0$, so dass

$$f(x) = a_0 + a_1x + \dots + a_nx^n = \sum_{i=0}^n a_ix^i .$$

- Die a_i heissen *Koeffizienten* des Polynoms, a_n wird als *Leitkoeffizient* bezeichnet.
- Das Nullpolynom $f(x) = 0$ (d.h. $a_i = 0 \forall i$) erhält keinen Grad, wird aber z.B. bei der Sprechweise „ein Polynom von Grad $\leq n$ “ mit eingeschlossen.

2.2.1. Satz. *Zwei Polynome sind genau dann gleich, wenn ihre Koeffizienten paarweise übereinstimmen, d.h.*

$$\sum_{i=1}^n a_ix^i = \sum_{i=1}^n b_ix^i \quad \iff \quad a_i = b_i, \quad i = 1, \dots, n .$$

2.2.2. Das Horner-Schema. Sei x_0 konkreter Wert. Um den Funktionswert $f(x_0)$ in obiger Darstellung naiv zu berechnen, braucht man $2n - 1$ Multiplikationen und n Additionen. Es gibt andere Darstellungen, z. B. die *Horner-Darstellung*:

$$f(x) = (\dots((a_n \cdot x + a_{n-1}) \cdot x + a_{n-2}) \cdot x + \dots + a_1) \cdot x + a_0 .$$

Hier beträgt der Aufwand n Multiplikationen und n Additionen (William George Horner, 1786–1837).

2.2.2. Satz. Sei $f(x) = \sum_{i=0}^n a_i x^i$, $a_n \neq 0$ und $b \in \mathbb{R}$. Für die Zahlen $c_n := a_n$, $c_{n-1} := c_n b + a_{n-1}$ bis $c_0 := c_1 b + a_0$ gilt

$$f(b) = c_0$$

$$f(x) = (x - b) \sum_{i=1}^n c_i x^{i-1} + c_0,$$

d.h. das Horner Schema enthält die Division von $f(x) - f(b)$ durch $x - b$.

2.2.3. Nullstellen und Faktorisierung; komplexe Polynome.

Definition. Als *Nullstelle* einer Funktion $f : D \rightarrow \mathbb{R}$ bezeichnet man jedes $x \in D$ mit $f(x) = 0$.

- Für Polynome erhält man aus Satz 2.2.2:

$$f(b) = 0 \quad \Leftrightarrow \quad f(x) = (x - b)h(x),$$

d.h. $f(x)$ enthält den Linearfaktor $(x - b)$.

- Nun kann $h(b)$ wiederum 0 sein. In diesem Fall gibt es ein Polynom h_1 mit $h(b) = (x - b)h_1(x)$. Damit gilt: $f(x) = (x - b)^2 h_1(x)$.

Definition. Man nennt $b \in \mathbb{R}$ eine *k-fache Nullstelle* von f und k die *Vielfachheit* von b , wenn gilt:

$$f(x) = (x - b)^k g(x) \quad \text{und} \quad g(b) \neq 0.$$

2.2.3. Satz.

- a) Sei $f(x) = \sum_{i=0}^n a_i x^i$ ein Polynom vom Grad größer gleich eins. Sind b_1, \dots, b_r alle (verschiedenen) reellen Nullstellen von f mit der jeweiligen Vielfachheit l_1, \dots, l_r , dann gilt

$$f(x) = \prod_{i=1}^r (x - b_i)^{l_i} q(x)$$

mit einem Polynom $q(x)$ vom Grad $n - \sum_{i=1}^r l_i$, das in \mathbb{R} keine Nullstellen hat.

- b) Jedes Polynom vom Grad n mit $n \geq 1$ hat höchstens n Nullstellen.

Betrachten wir komplexe Polynome:

$$f(z) = \sum_{j=0}^n a_j z^j, \quad a_j \in \mathbb{C}.$$

- Alle Operationen, wie Addition, Subtraktion, Multiplikation, Division, die für reelle Polynome definiert sind, werden analog für komplexe Polynome definiert.
- Der Grund für die Einführung der komplexen Zahlen war das Polynom $x^2 + 1$, welches in \mathbb{R} keine Nullstelle hat. Die Gleichung $x^2 + 1 = 0$ hat aber in \mathbb{C} die Lösungen $\pm i$.

2.2.4. Satz (Fundamentalsatz der Algebra). *Jedes komplexe Polynom vom Grad $n \geq 1$ besitzt eine Nullstelle.*

2.2.5. Satz. *Jedes komplexe Polynom f vom Grad $n \geq 1$ besitzt eine Faktorisierung über \mathbb{C} der Form*

$$f(z) = a_n(z - w_1)^{l_1} \cdots (z - w_k)^{l_k}$$

mit paarweise verschiedenen Nullstellen $w_j \in \mathbb{C}$ der Vielfachheit l_j mit $\sum_{j=1}^k l_j = n$.

Jedes reelle Polynom kann als komplexes Polynom aufgefasst werden, da $\mathbb{R} \subseteq \mathbb{C}$.

2.2.6. Lemma. *Sei $w \in \mathbb{C}$ eine Nullstelle eines Polynoms f mit reellen Koeffizienten. Dann ist auch \bar{w} eine Nullstelle von f .*

2.2.7. Satz. *Jedes reelle Polynom $f(x) = \sum_{i=0}^n a_i x^i$, $n \geq 1$, $a_i \in \mathbb{R}$, $a_n \neq 0$, besitzt die Faktorisierung über \mathbb{R}*

$$f(x) = a_n(x - b_1)^{l_1} \cdots (x - b_r)^{l_r} (x^2 + c_1x + d_1)^{k_1} \cdots (x^2 + c_sx + d_s)^{k_s},$$

mit reellen Nullstellen $b_j \in \mathbb{R}$ der Vielfachheit l_j und quadratischen Polynomen $x^2 + c_jx + d_j$, die keine reelle Nullstelle haben.

2.2.8. Satz. *Die rationalen Nullstellen eines Polynoms*

$$f(x) = \sum_{j=0}^n a_j x^j \quad , \quad a_j \in \mathbb{Z}$$

findet man unter den Brüchen $\frac{a}{b}$ ($a, b \in \mathbb{Z}$) in denen a ein Teiler von a_0 und b ein Teiler von a_n ist.

2.2.4. Das Newton Interpolationsverfahren.

2.2.9. Satz. *Zu $(n + 1)$ beliebigen Stützpunkten (x_i, y_i) , $i = 0, \dots, n$, mit $x_i \neq x_j$ für $i \neq j$, gibt es genau ein Polynom p_n vom Grad kleiner gleich n , so dass $p_n(x_i) = y_i$, $i = 0, \dots, n$.*

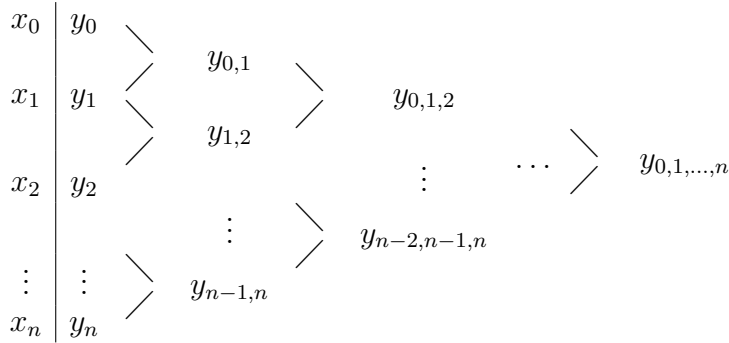
Man suche das Polynom p_n aus Satz 2.2.9 in der Form

$$p_n(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \cdots + \alpha_n(x - x_0)(x - x_1) \cdots (x - x_{n-1})$$

Die Bedingung $p_n(x_i) = y_i, i = 0, \dots, n$, liefert das Gleichungssystem:

$$\begin{aligned} y_0 &= \alpha_0 \\ y_1 &= \alpha_0 + \alpha_1(x_1 - x_0) \\ &\vdots \\ y_n &= \alpha_0 + \alpha_1(x_n - x_0) + \cdots + \alpha_n(x_n - x_0) \cdots (x_n - x_{n-1}) \end{aligned}$$

Dieses kann schrittweise von oben nach unten mit der Methode der *dividierten Differenzen* gelöst werden:



Hierbei ist

$$\begin{aligned} y_{0,1} &= \frac{y_1 - y_0}{x_1 - x_0}, & y_{1,2} &= \frac{y_2 - y_1}{x_2 - x_1}, & \dots \\ y_{0,1,2} &= \frac{y_{1,2} - y_{0,1}}{x_2 - x_0}, & y_{1,2,3} &= \frac{y_{2,3} - y_{1,2}}{x_3 - x_1}, & \dots \\ &\vdots & &\vdots & \end{aligned}$$

Man kann nachrechnen, dass

$$\alpha_i = y_{0,1,\dots,i},$$

d.h. die gesuchten Koeffizienten stehen in oberen Schrägzeile.

2.2.5. Rationale Funktionen, Polynomdivision.

Definition. Der Quotient zweier Polynome

$$f(x) = \frac{p(x)}{q(x)} = \frac{\sum_{i=0}^n a_i x^i}{\sum_{j=0}^m b_j x^j}, \quad a_n \neq 0, b_m \neq 0,$$

heißt *rationale Funktion*.

2.2.10. Satz. Jede rationale Funktion $\frac{p(x)}{q(x)}$ mit Zählergrad \geq Nennergrad lässt sich darstellen als

$$\frac{p(x)}{q(x)} = h(x) + \frac{r(x)}{q(x)}$$

mit einem Polynom h und einem Restpolynom r wobei entweder $r = 0$ oder $\text{Grad } r < \text{Grad } q$.

Definition. Seien p, d Polynome. Man nennt d eine Teiler von p , wenn es ein Polynom p_0 gibt, so dass $p(x) = d(x)p_0(x)$.

2.2.11. Lemma. Sei $\frac{p}{q}$ eine rationale Funktion und sei $\text{Grad } p \geq \text{Grad } q$. Sei ferner

$$\frac{p}{q} = h + \frac{r}{q},$$

wobei $\text{Grad } r < \text{Grad } q$. Dann ist d genau dann ein gemeinsamer Teiler von p und q , wenn d gemeinsamer Teiler von q und r ist.

Sei $f(x) = \frac{p(x)}{q(x)}$ eine rationale Funktion, wobei p und q teilerfrei sind. Dann haben $p(x)$ und $q(x)$ keine gemeinsamen Nullstellen und der maximale Definitionsbereich der rationalen Funktion f ist gegeben durch

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

Sei $b \in \mathbb{R}$ eine l -fache Nullstelle von $q(x)$, d.h.

$$q(x) = (x - b)^l q_1(x), \quad \text{mit } q_1(b) \neq 0.$$

Dann nennt man b einen l -fachen Pol von f . Die Funktion f verhält sich wie

$$f(x) \approx \frac{p(b)}{q_1(b)} \cdot \frac{1}{(x - b)^l} \quad \text{für } x \text{ nahe } b,$$

und für $|x| \rightarrow \pm\infty$ wie

$$f(x) = \frac{a_n x^n (1 + \dots + \frac{a_0}{a_n x^n})}{b_m x^m (1 + \dots + \frac{b_0}{b_m x^m})} \approx \frac{a_n}{b_m} x^{n-m}.$$

2.3. Trigonometrische Funktionen

2.3.1. Definition und einfache Eigenschaften. In Abschnitt 1.3.4 haben wir die Sinus- und die Cosinusfunktion definiert. Aus der Definition erhält man sofort folgende Eigenschaften:

$$\begin{aligned} -1 &\leq \cos x \leq 1, \\ \cos(-x) &= \cos x && \text{(gerade Funktion),} \\ \cos(x + 2k\pi) &= \cos x && \text{(} 2\pi\text{-periodisch),} \end{aligned}$$

$$\begin{aligned}
-1 &\leq \sin x \leq 1, \\
\sin(-x) &= -\sin x && \text{(ungerade Funktion),} \\
\sin(x + 2k\pi) &= \sin x && \text{(}2\pi\text{-periodisch),} \\
\cos x = 0 &\Leftrightarrow x = \pm\frac{1}{2}\pi, \pm\frac{3}{2}\pi, \pm\frac{5}{2}\pi, \dots, \\
\sin x = 0 &\Leftrightarrow x = 0, \pm\pi, \pm2\pi, \dots
\end{aligned}$$

2.3.2. Das Additionstheorem.

2.3.1. Satz (Additionstheorem). *Für alle $x, y \in \mathbb{R}$ gilt:*

$$\begin{aligned}
\cos(x + y) &= \cos x \cos y - \sin x \sin y, \\
\sin(x + y) &= \sin x \cos y + \cos x \sin y.
\end{aligned}$$

Bemerkung. Der Spezialfall $y = \frac{\pi}{2}$, bzw. $y = -\frac{\pi}{2}$ impliziert:

$$\begin{aligned}
\sin\left(x + \frac{\pi}{2}\right) &= \cos x, \\
\cos\left(x - \frac{\pi}{2}\right) &= \sin x,
\end{aligned}$$

d.h. wenn man die Sinuskurve nach links um $\frac{\pi}{2}$ verschiebt, erhält man die Cosinuskurve.

Bemerkung. In Formelsammlungen gibt es zahlreiche Identitäten, die sich aus Satz 2.3.1 herleiten lassen. Hier einige Beispiele:

$$\begin{aligned}
\cos(x - y) &= \cos x \cos y + \sin x \sin y, \\
\sin(x - y) &= \sin x \cos y - \cos x \sin y, \\
\sin x + \sin y &= 2 \sin \frac{x+y}{2} \cos \frac{x-y}{2}, \\
\cos x + \cos y &= 2 \cos \frac{x+y}{2} \cos \frac{x-y}{2}, \\
\cos(2x) &= \cos^2 x - \sin^2 x = 2 \cos^2 x - 1, \\
\sin 2x &= 2 \sin x \cos x, \\
1 + \cos x &= 2 \cos^2 \frac{x}{2}, \\
1 - \cos y &= 2 \sin^2 \frac{y}{2},
\end{aligned}$$

Achtung: $\sin^2 x := (\sin(x))^2$.

2.3.3. Die Tangens- und Cotangensfunktion.

Definition. Die *Tangens- und Cotangensfunktionen* sind definiert durch

$$\begin{aligned}
\tan x &:= \frac{\sin x}{\cos x} && \text{mit } x \neq (2k + 1)\frac{\pi}{2}, k \in \mathbb{Z}, \\
\cot x &:= \frac{\cos x}{\sin x} && \text{mit } x \neq k\pi, k \in \mathbb{Z}.
\end{aligned}$$

Wir erhalten sofort folgende Eigenschaften:

$$\begin{aligned} \tan(-x) &= -\tan x && \text{(ungerade Funktion),} \\ \cot(-x) &= -\cot x && \text{(ungerade Funktion),} \\ \tan(x + \pi) &= \tan x && \text{(\pi-periodisch),} \\ \cot(x + \pi) &= \cot x && \text{(\pi-periodisch),} \\ \tan(x + y) &= \frac{\tan x + \tan y}{1 - \tan x \tan y} && \text{für „erlaubte“ } x, y. \end{aligned}$$

2.3.4. Polardarstellung komplexer Zahlen. Die Menge der komplexen Zahlen ist $\mathbb{C} = \{z = x + iy; x, y \in \mathbb{R}\}$. Man kann jede komplexe Zahl $z \neq 0$ auch als „Zeiger“ auffassen und deshalb ist z eindeutig bestimmt durch den Winkel φ und den Betrag $|z|$.

- Man nennt φ das *Argument* von z , $\varphi =: \arg z$. Falls $-\pi < \varphi \leq \pi$, heisst φ *Hauptwert* von $\arg z$
- Wir haben zwei Möglichkeiten, eine komplexe Zahl anzugeben:

$$z = x + iy \quad \text{mit } x = \operatorname{Re} z, y = \operatorname{Im} z,$$

oder durch Angabe des Betrages $r = |z|$ und des Arguments $\varphi = \arg z$, siehe Abschnitt 1.3.5. Aber wir wissen, dass $x = r \cos \varphi$, $y = r \sin \varphi$ und somit erhalten wir

$$z = r(\cos \varphi + i \sin \varphi).$$

Diese Form der Darstellung von z heisst *Polardarstellung*.

- Die folgende abkürzende Schreibweise ist sinnvoll

$$e^{i\varphi} := \cos \varphi + i \sin \varphi, \quad \varphi \in \mathbb{R}.$$

Sie heisst *Euler-Formel* (Leonhard Euler, 1707–1783).

Definition. Die Arcus-Cosinus-Funktion $\arccos: [-1, 1] \rightarrow [0, \pi]$ ist dadurch definiert, dass

$$\begin{aligned} \arccos(\cos \varphi) &= \varphi && \text{für alle } \varphi \in [0, \pi], \text{ und} \\ \cos(\arccos t) &= t && \text{für alle } t \in [-1, 1]. \end{aligned}$$

Bemerkung. Umrechnungen.

- 1) Sei $z \in \mathbb{C}$, mit $z \neq 0$, gegeben in der Form: $z = x + iy$, $x, y \in \mathbb{R}$. Dann setzen wir

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, \\ \varphi &= \begin{cases} \arccos \frac{x}{r} & \text{falls } y \geq 0, \\ -\arccos \frac{x}{r} & \text{falls } y < 0. \end{cases} \end{aligned}$$

Wir erhalten die Polardarstellung $z = re^{i\varphi}$.

2) Sei $z \in \mathbb{C}$ in der Polardarstellung $z = re^{i\varphi}$ gegeben. Dann setzen wir

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

und erhalten die Darstellung $z = x + iy$.

2.3.5. Die de Moivre-Formeln und Anwendungen.

2.3.2. Satz. (Abraham de Moivre, 1667–1754). Für alle $\varphi, \psi \in \mathbb{R}$ gilt

$$\begin{aligned} e^{i(\psi+\varphi)} &= e^{i\varphi} e^{i\psi}, \\ e^{in\varphi} &= (e^{i\varphi})^n \quad \text{für alle } n \in \mathbb{N}, \\ e^{-i\varphi} &= \overline{e^{i\varphi}} = \frac{1}{e^{i\varphi}}. \end{aligned}$$

Die Multiplikation und Division komplexer Zahlen lässt sich besonders einfach in der Polardarstellung berechnen. Sei $z = |z|e^{i\varphi}$ und $w = |w|e^{i\psi}$, dann gilt:

$$\begin{aligned} z \cdot w &= |z||w| e^{i(\varphi+\psi)}, \\ \frac{z}{w} &= \frac{|z|}{|w|} e^{i(\varphi-\psi)}, \quad \text{falls } w \neq 0. \end{aligned}$$

2.3.6. Harmonische Schwingungen.

Definition. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt *periodisch* mit einer *Periode* $2l$, wenn

$$f(x + 2l) = f(x) \quad \forall x \in \mathbb{R}.$$

Definition. Als *Schwingung* bezeichnet man einen Vorgang, der durch eine periodische Funktion eines „Zeitparameters“ $t \in \mathbb{R}$ beschrieben wird. Eine durch

$$s(t) = A \cos(\omega t + \alpha), \quad t \in \mathbb{R}$$

mit festem $A, \omega, \alpha \in \mathbb{R}$ dargestellte Schwingung heißt *harmonisch*. A heißt *Amplitude*, $\omega t + \alpha$ die *Phase*, α die *Nullphase* und ω die *Kreisfrequenz*. Die *Periode* beträgt $T = \frac{2\pi}{\omega}$ und die *Frequenz* $\nu = \frac{1}{T} = \frac{\omega}{2\pi}$.

Die *Überlagerung* $s(t)$ zweier Schwingungen $s_1(t)$, $s_2(t)$ ist punktweise definiert, d.h. die Auslenkungen addieren sich

$$s(t) := s_1(t) + s_2(t).$$

Die Überlagerung $s(t)$ ist im Allgemeinen nicht mehr periodisch. Ist der Quotient der Kreisfrequenzen

$$\frac{\omega_1}{\omega_2} = \frac{n_1}{n_2} = \frac{\nu_1}{\nu_2} = \frac{T_2}{T_1}$$

jedoch eine rationale Zahl, so ist die Überlagerung periodisch, und zwar mit der Periode $T := \frac{2\pi n_1}{\omega_1} = \frac{2\pi n_2}{\omega_2}$, und mit der Kreisfrequenz $\frac{\omega_1}{n_1} = \frac{\omega_2}{n_2}$.

Die Behandlung von harmonischen Schwingungen $s(t)$ vereinfacht sich, wenn man *komplexe Schwingungen* $z(t)$ einführt. Sei

$$s(t) = A \cos(\omega t + \alpha),$$

dann definiert man

$$\begin{aligned} z(t) &:= A \cos(\omega t + \alpha) + iA \sin(\omega t + \alpha) \\ &= Ae^{i(\omega t + \alpha)} \\ &= ae^{i\omega t}. \end{aligned}$$

Hierbei ist a gegeben durch $a := Ae^{i\alpha}$ und heißt *komplexe Amplitude*.

2.3.3. Satz. *Besitzen zwei harmonische Schwingungen die gleiche Kreisfrequenz, d.h.*

$$s_1(t) = A_1 \cos(\omega t + \alpha_1), \quad s_2(t) = A_2 \cos(\omega t + \alpha_2),$$

so ist ihre Überlagerung (Superposition) $s(t)$ wieder eine harmonische Schwingung, die gegeben ist durch

$$s(t) := s_1(t) + s_2(t) = A \cos(\omega t + \alpha)$$

mit $A = \sqrt{u^2 + v^2}$, $\cos \alpha = \frac{u}{A}$, $\sin \alpha = \frac{v}{A}$, wobei

$$u := A_1 \cos \alpha_1 + A_2 \cos \alpha_2, \quad v := A_1 \sin \alpha_1 + A_2 \sin \alpha_2.$$

Bemerkung. • Es gilt folgende Formel:

$$\sin \delta + \sin 2\delta + \dots + \sin n\delta = \frac{\sin \frac{n+1}{2}\delta \sin \frac{n}{2}\delta}{\sin \frac{\delta}{2}}$$

- Auf Grund der Darstellungen $\omega_1 = \frac{\omega_1 + \omega_2}{2} + \frac{\omega_1 - \omega_2}{2}$ und $\omega_2 = \frac{\omega_1 + \omega_2}{2} - \frac{\omega_2 - \omega_1}{2}$ läßt sich die Überlagerung zweier komplexer Schwingungen immer als Produkt

$$a_1 e^{i\omega_1 t} + a_2 e^{i\omega_2 t} = \left(a_1 e^{i\frac{\omega_1 - \omega_2}{2} t} + a_2 e^{i\frac{\omega_2 - \omega_1}{2} t} \right) e^{i\frac{\omega_1 + \omega_2}{2} t}$$

darstellen. Man sagt dazu *modulierte Schwingung* mit *modulierter Amplitude*.

- Im Allgemeinen ist eine modulierte Schwingung nicht periodisch. Wenn das Verhältnis der Perioden (oder der Frequenzen/Kreisfrequenzen) rational ist, ist die modulierte Schwingung zwar periodisch, aber im allgemeinen nicht harmonisch.

2.4. Zahlenfolgen und Grenzwerte

2.4.1. Folgen. Das Herz der Analysis sind Grenzwerte. Wir fangen unsere Untersuchungen mit Folgen an und übertragen die Ergebnisse dann auf Funktionen.

Definition. Unter einer *Folge* reeller Zahlen versteht man eine auf \mathbb{N}_0 oder \mathbb{N} erklärte Funktion, das heisst jedem $n \in \mathbb{N}_0$ ist ein $a_n \in \mathbb{R}$ zugeordnet. Man schreibt

$$(a_n)_{n \in \mathbb{N}_0}; (a_n)_{n \geq 0}; a_0, a_1, a_2, \dots$$

Die Zahlen a_n heissen *Glieder* der Folge.

Beispiele.

- | | | |
|-----|-------------------------------|---|
| (1) | $a_n = c$ | <i>konstante Folge, c, c, c, \dots</i> |
| (2) | $a_n = n$ | <i>Folge der natürlichen Zahlen $1, 2, 3, \dots$</i> |
| (3) | $a_n = a_0 + nd$ | <i>arithmetische Folge $a_0, a_0 + d, a_0 + 2d, \dots$</i> |
| (4) | $a_n = a_0 q^n$ | <i>geometrische Folge $a_0, a_0 q, a_0 q^2, \dots$</i> |
| (5) | $a_0 = 1, a_{n+1} = (n+1)a_n$ | <i>rekursiv definierte Folge.</i> |

Definition. Eine Folge heisst *beschränkt*, wenn es Konstanten K_1, K_2 gibt, so dass

$$K_1 \leq a_n \leq K_2 \quad \text{für alle } n \in \mathbb{N}_0.$$

Im obigen Beispiel sind die Folgen in (1) und (4) für $|q| \leq 1$ beschränkt. Die Folgen in (2), (3) für $d \neq 0$, (4) für $|q| > 1$ und (5) sind unbeschränkt.

2.4.2. Konvergenz und Grenzwert.

Definition. Man sagt, die Folge $(a_n)_{n \geq 0}$ *konvergiert* gegen den *Grenzwert* $a \in \mathbb{R}$ und schreibt

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a,$$

wenn es zu jeder beliebigen kleinen Schranke $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}_0$ gibt, so dass gilt:

$$|a_n - a| < \varepsilon \quad \forall n \geq n_0,$$

d.h. alle Glieder ab einem bestimmten Index (dieser hängt im allgemeinen von ε ab!) liegen in einer ε -Umgebung von a . Falls der Grenzwert existiert, heisst die Folge *konvergent*, ansonsten *divergent*. Falls $a = 0$, heisst die Folge *Nullfolge*.

2.4.1. Satz. Für jede konvergente Folge $(a_n)_{n \geq 0}$ gilt:

- 1) Ist $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} a_n = b$, so gilt $a = b$.

2) Die Folge $(a_n)_{n \geq 0}$ ist beschränkt.

Definition. Man sagt, eine Folge $(a_n)_{n \geq 0}$ *divergiert gegen* ∞ (‘‘konvergiert’’ gegen ∞ , in Zeichen

$$\lim_{n \rightarrow \infty} a_n = \infty,$$

wenn für alle $K \in \mathbb{N}_0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $a_n > K$ für alle $n \geq n_0$. Analog definiert man: *divergiert gegen* $-\infty$.

2.4.2. Lemma. Für alle $x \in \mathbb{R}$ gilt

- 1) $\lim_{n \rightarrow \infty} x^n = 0$, falls $|x| < 1$,
- 2) $\lim_{n \rightarrow \infty} x^n = \infty$, falls $x > 1$,
- 3) Die Folge $(x^n)_{n \geq 0}$ *divergiert*, falls $x \leq -1$.

Beispiel. (*geometrische Reihe.*) Die Folge $(a_n)_{n \geq 0}$ sei gegeben durch:

$$a_n = x^n, \quad x \in \mathbb{R}.$$

Für die endliche geometrische Reihe gilt:

$$s_n = \sum_{k=0}^n a_k = 1 + x + x^2 + \cdots + x^n = \begin{cases} \frac{1-x^{n+1}}{1-x}, & \text{falls } x \neq 1, \\ n+1, & \text{falls } x = 1. \end{cases}$$

Es können folgende Fälle auftreten:

$$\begin{aligned} |x| < 1 &\Rightarrow \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \left(\frac{1}{1-x} + \frac{x^{n+1}}{1-x} \right) = \frac{1}{1-x}, \\ x \geq 1 &\Rightarrow \lim_{n \rightarrow \infty} s_n = \infty, \\ x \leq -1 &\Rightarrow s_n \text{ divergiert.} \end{aligned}$$

Statt $\lim_{n \rightarrow \infty} \sum_{k=0}^n x^k$ schreibt man $\sum_{k=0}^{\infty} x^k$, d.h.

$$\sum_{k=0}^{\infty} x^k := \lim_{N \rightarrow \infty} \sum_{k=0}^N x^k = \begin{cases} \frac{1}{1-x}, & \text{falls } |x| < 1, \\ \infty, & \text{falls } x \geq 1, \\ \text{unbestimmt,} & \text{falls } x < -1. \end{cases}$$

Beispiel. *Harmonische Reihe.* Wir betrachten die Folge $(\frac{1}{n})_{n \geq 1}$. Die zugehörigen Partialsummen s_n sind

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Es gilt:

$$\lim_{n \rightarrow \infty} s_n = \infty \quad \text{oder} \quad \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Die Reihe divergiert, aber sehr langsam. Man kann zeigen, dass

$$\sum_{k=1}^{2^{2n}} \frac{1}{k} \geq n + 1, \quad \text{aber} \quad \sum_{k=1}^{2^n - 1} \frac{1}{k} \leq n.$$

Definition. Ist $(a_n)_{n \geq 0}$ eine Folge und $n_0 < n_1 < \dots$ eine aufsteigende *Indexfolge*, dann heisst die Folge a_{n_0}, a_{n_1}, \dots *Teilfolge* der Folge $(a_n)_{n \geq 0}$.

Bemerkung. Wenn eine Folge konvergiert, dann konvergiert jede Teilfolge gegen den gleichen Grenzwert. Umgekehrt gibt es Folgen, die nicht konvergieren, aber konvergente Teilfolgen besitzen. Z.B. konvergiert $(a_n)_{n \in \mathbb{N}}$ mit $a_n = (-1)^n$ nicht, aber die Teilfolgen $(a_{2n})_{n \in \mathbb{N}}$ und $(a_{2n+1})_{n \in \mathbb{N}}$ konvergieren gegen 1 bzw. -1 .

2.5. Rechenregeln für Grenzwerte und Konvergenzkriterien

2.5.1. Rechenregeln. Seien $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$ Folgen.

2.5.1. Satz. Sind $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$ konvergente Zahlenfolgen mit $a_n \rightarrow a$ und $b_n \rightarrow b$, dann gilt:

- $\lim_{n \rightarrow \infty} (a_n \pm b_n) = a \pm b$,
- $\lim_{n \rightarrow \infty} (a_n b_n) = a \cdot b$,
- ist $a \neq 0$, dann gibt es ein $n_1 \in \mathbb{N}$ mit $a_n \neq 0$ für alle $n \geq n_1$ und für die Folgen $(a_n)_{n \geq n_1}, (b_n)_{n \geq n_1}$ gilt: $\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{1}{a}$, $\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = \frac{b}{a}$,
- $\lim_{n \rightarrow \infty} |a_n| = |a|$,
- $\lim_{n \rightarrow \infty} \sqrt{a_n} = \sqrt{a}$, falls alle $a_n \geq 0$.

2.5.2. Grenzwertbestimmung durch Abschätzungen.

2.5.2. Satz. Lassen sich für alle $n \geq n_1$ die Glieder der Folge $(a_n)_{n \geq 0}$ abschätzen durch $b_n \leq a_n \leq c_n$ mit $\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} c_n = c$, dann gilt $\lim_{n \rightarrow \infty} a_n = c$.

2.5.3. Satz. Seien $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$ konvergente Folgen mit $a_n \leq b_n$ für alle $n \geq n_1$. Dann gilt: $\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n$.

2.5.3. Monotone Folgen. Wir definieren Monotonität für Folgen wie in Abschnitt 2.1.2 für Funktionen.

Definition. Eine Folge heißt *monoton wachsend* (bzw. *monoton fallend*), wenn für alle $n \geq 0$ gilt $a_{n+1} \geq a_n$ (bzw. $a_{n+1} \leq a_n$).

2.5.4. Satz. Jede *monoton wachsende* oder *fallende*, *beschränkte* Folge ist *konvergent*.

Beispiel. Wir betrachten die Folge $(a_n)_{n \geq 0}$, wobei $a_n := \sum_{k=0}^n \frac{1}{k!}$. Diese Folge ist monoton wachsend und beschränkt. Also konvergiert sie und man definiert die *Eulersche Zahl* e durch

$$e := \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n \frac{1}{k!} \right) = 2,718 \dots$$

Die Folge $(a_n)_{n \geq 0}$ konvergiert sehr schnell. Man kann zeigen, dass die Eulersche Zahl e auch der Grenzwert der Folge $(b_n)_{n \geq 1}$ mit $b_n := \left(1 + \frac{1}{n}\right)^n$ ist, d.h.

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

Die Folge $(b_n)_{n \geq 1}$ konvergiert sehr langsam.

Zum Beweis der Monotonität benötigen wir die *Bernoullische Ungleichung*: Für alle $s > -1$ und alle $n \in \mathbb{N}_0$ gilt

$$(1 + s)^n \geq 1 + ns.$$

Sie lässt sich durch Induktion über n beweisen.

2.5.4. Die Exponentialfunktion. Die *Exponentialfunktion* $e : x \mapsto e^x$ ist für alle $x \in \mathbb{R}$ definiert durch

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

2.6. Funktionengrenzwerte, Stetigkeit

Seien $I \subseteq \mathbb{R}$ ein Intervall, $a \in I \cup \{-\infty, \infty\}$ und $f : I \setminus \{a\} \rightarrow \mathbb{R}$ eine Funktion. Uns interessiert das Verhalten von f , wenn x sich a nähert, wobei $x \neq a$.

2.6.1. Grenzwerte.

Definition. Die Funktion f hat für x gegen a den *rechtsseitigen Grenzwert* (bzw. *linksseitigen Grenzwert*) c , in Zeichen $\lim_{x \rightarrow a^+} f(x) = c$ (bzw. $\lim_{x \rightarrow a^-} f(x) = c$), wenn für jede Folge $(x_n)_{n \geq 0}$ aus I mit $x_n \rightarrow a$ und $a < x_n$ für alle n (bzw. $x_n \rightarrow a$ und $x_n < a$ für alle n) die Folge $(f(x_n))_{n \geq 0}$ den Grenzwert c hat. f hat für x gegen a den *Grenzwert* c , in Zeichen $\lim_{x \rightarrow a} f(x) = c$, wenn gilt $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = c$.

2.6.1. Satz. Aus $\lim_{x \rightarrow a} f(x) = c$, $\lim_{x \rightarrow a} g(x) = d$, mit $c, d \in \mathbb{R}$ folgt:

- a) $\lim_{x \rightarrow a} [f(x) \pm g(x)] = c \pm d$,
- b) $\lim_{x \rightarrow a} f(x) \cdot g(x) = c \cdot d$,

$$c) \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{c}{d}, \quad \text{falls } d \neq 0.$$

Dies gilt auch für $a = \pm\infty$ und einseitige Grenzwerte, $x \rightarrow a^+$, $x \rightarrow a^-$, aber nur für endliche Grenzwerte c, d .

2.6.2. Satz. Wenn $g(x) \leq f(x) \leq h(x)$ für alle x in der Nähe von a gilt (bzw. für alle hinreichend großen x) und wenn $g(x) \rightarrow c$, $h(x) \rightarrow c$ für $x \rightarrow a$ (bzw. $x \rightarrow \infty$), dann gilt auch $\lim_{x \rightarrow a} f(x) = c$ (bzw. $\lim_{x \rightarrow \infty} f(x) = c$).

Beispiel. Es gelten:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1, \\ \lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0.$$

2.6.2. Asymptoten.

- (1) Man nennt die Gerade $x = a$ eine *vertikale Asymptote* der Kurve $y = f(x)$, wenn beim Grenzübergang $x \rightarrow a^+$ oder $x \rightarrow a^-$ die Funktionswerte $f(x)$ gegen ∞ oder $-\infty$ streben.
- (2) Die Gerade $y = c$ heißt *horizontale Asymptote* der Kurve $y = f(x)$, wenn $\lim_{x \rightarrow \infty} f(x) = c$ oder $\lim_{x \rightarrow -\infty} f(x) = c$ gilt.
- (3) Als *schräge Asymptote* der Kurve $y = f(x)$ bezeichnet man die Gerade $y = px + q$, falls $p \neq 0$ und $f(x) - (px + q) \rightarrow 0$ für $x \rightarrow \infty$ oder $x \rightarrow -\infty$.

2.6.3. Stetigkeit.

Definition. Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion. Die Funktion f heißt *im Punkt* $x_0 \in I$ *stetig*, wenn $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. Falls x_0 ein Randpunkt des Intervalls ist, so ist der Grenzwert nur einseitig zu verstehen. Die Funktion f heißt *auf* I *stetig*, wenn sie in jedem Punkt $x_0 \in I$ stetig ist.

2.6.3. Satz.

- a) Sind f und g auf einem Intervall $I \subseteq \mathbb{R}$ stetig, so gilt das auch für $f \pm g$ und $f \cdot g$. Ferner ist $\frac{f}{g}$ stetig in allen $x \in I$ mit $g(x) \neq 0$.
- b) Sind $f : I \rightarrow \mathbb{R}$, $g : D \rightarrow \mathbb{R}$ stetig mit $g(D) \subseteq I$, dann ist auch die Komposition $h : D \rightarrow \mathbb{R}$, $h = f \circ g$ auf D stetig.

2.6.4. Folgerung. a) Jedes Polynom $p(x) = \sum_{i=0}^n a_i x^i$ ist auf ganz \mathbb{R} stetig.

- b) Seien p, q Polynome, die teilerfremd sind. Dann ist die rationale Funktion $f(x) = \frac{p(x)}{q(x)}$ stetig in allen Punkten $x \in \mathbb{R}$ mit $q(x) \neq 0$.

2.6.5. Satz. Für jede auf einem abgeschlossenen beschränkten Intervall $[a, b]$ stetige Funktion gilt:

- a) **Schrankensatz:** Es gibt eine Schranke K mit $|f(x)| \leq K$ für alle $x \in [a, b]$.
(Man sagt: f ist auf $[a, b]$ beschränkt.)
- b) **Satz vom Maximum und Minimum:** Es gibt Werte $x_0, x_1 \in [a, b]$, so dass $f(x_0) \leq f(x) \leq f(x_1)$ für alle $x \in [a, b]$.
(Man sagt: f nimmt auf $[a, b]$ sein Minimum und sein Maximum an.)
- c) **Zwischenwertsatz:** Zu jeder Zahl c zwischen dem Minimum $f(x_0)$ und dem Maximum $f(x_1)$ gibt es wenigstens ein $\bar{x} \in [a, b]$ mit $f(\bar{x}) = c$.
(Man sagt: f nimmt jeden Wert zwischen seinem Minimum und Maximum an.)
- d) **Die gleichmäßige Stetigkeit:** Zu jeder beliebig kleinen Zahl $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass zwei Funktionswerte sich um höchstens ε unterscheiden, sobald die Argumente weniger als δ voneinander entfernt sind, d.h.

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : (|x - x'| \leq \delta \Rightarrow |f(x) - f(x')| \leq \varepsilon).$$

2.6.6. Folgerung. Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und haben $f(a)$ und $f(b)$ verschiedene Vorzeichen, dann gibt es wenigstens eine Nullstelle $\bar{x} \in (a, b)$ von f , d.h. $f(\bar{x}) = 0$.

Bemerkung. Bisektion Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion so, dass $f(b) \cdot f(a) < 0$. Dann existiert nach Korollar 2.6.6 eine Nullstelle in $[a, b]$. Diese kann man durch sukzessives Halbieren des Intervalls genauer bestimmen. Im Falle $f\left(\frac{a+b}{2}\right) \cdot f(b) < 0$ liegt die Nullstelle im Intervall $\left[\frac{a+b}{2}, b\right]$, sonst in $\left[a, \frac{a+b}{2}\right]$. Dieses Verfahren kann beliebig oft wiederholt werden und somit kann die Nullstelle beliebig genau bestimmt werden.

KAPITEL 3

Differentiation

Die Differentialrechnung ist eines der wichtigsten Hilfsmittel in der Mathematik selbst und in der mathematischen Behandlung von Problemen aus Wissenschaft und Technik. Insbesondere wird sie in diesem Kapitel zur Kurvendiskussion und zur Extremwertbestimmung benutzt.

3.1. Die Ableitung

3.1.1. Definition der Ableitung. Die Funktion f sei auf dem Intervall $I \subseteq \mathbb{R}$ definiert und es sei $x_0 \in I$. Man sagt, f ist in x_0 differenzierbar, wenn der Grenzwert

$$(3.1.1) \quad \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert und endlich ist. Dieser Grenzwert wird mit $f'(x_0)$ bezeichnet. Man nennt

$$\frac{\Delta f(x)}{\Delta x} := \frac{f(x_0 + h) - f(x_0)}{h}$$

den *Differenzenquotienten*. Andere Bezeichnungen:

$$f'(x) = \frac{df(x)}{dx} = \frac{d}{dx} f.$$

Falls f in jedem Punkt $x_0 \in I$ differenzierbar ist, sagt man dass f auf I differenzierbar ist. Die so definierte Funktion $f' : I \rightarrow \mathbb{R} : x \mapsto f'(x)$ heisst *Ableitung von f* .

Beispiele.

$$(3.1.2) \quad \begin{array}{l|l} f(x) & f'(x) \\ \hline c & 0 \\ ax + b & a \\ x^n & nx^{n-1} \quad \text{für } n \in \mathbb{N} \\ \sqrt{x} & \frac{1}{2\sqrt{x}} \quad \text{für } x > 0 \end{array}$$

Geometrische Interpretation. Sei $y = f(x)$ eine gegebene Funktion. Wir betrachten den Graphen

$$G_f := \{(x, y) \in \mathbb{R}^2 \mid x \in I, y = f(x)\}$$

wie in Abschnitt 1.3.1. Dann ist

$$\tan \varphi = \frac{\Delta y}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0}$$

der Anstieg der Sekante durch die Punkte $(x_0, f(x_0))$, $(x, f(x))$. Der Grenzwert für $x \rightarrow x_0$ ist der Anstieg der Tangente an $y = f(x)$ im Punkt $(x_0, f(x_0))$, d.h. die Tangente des Graphen $y = f(x)$ hat im Punkt $(x_0, f(x_0))$ die Formel

$$(*) \quad y = f'(x_0)(x - x_0) + f(x_0).$$

Analytische Interpretation. Sei $y = f(x)$ eine gegebene Funktion. Wir suchen die „beste“ lineare Approximation von f in der Nähe von x_0 , d.h. eine lineare Funktion $g(x)$, so dass

$$(\dagger) \quad \lim_{x \rightarrow x_0} \frac{f(x) - g(x)}{x - x_0} = 0,$$

d.h. der Fehler $f(x) - g(x)$ strebt schneller gegen Null als $x - x_0$. Die Funktion g ist linear und somit gegeben durch $g(x) = m(x - x_0) + f(x_0)$. Aus (\dagger) erhalten wir $m = f'(x_0)$, d.h. die Tangente $(*)$ ist die „beste“ lineare Approximation von $f(x)$ nahe $(x_0, f(x_0))$.

3.1.3. Satz. *Jede in $x_0 \in I$ differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ ist in x_0 auch stetig.*

WARNUNG: Stetigkeit von f ist *nicht* hinreichend für Differenzierbarkeit!

3.1.2. Differentiationsregeln.

3.1.4. Satz. *Sind Funktionen $f, g : I \rightarrow \mathbb{R}$ im Punkt $x \in I$ differenzierbar, so gilt:*

- (a) $(f + g)'(x) = f'(x) + g'(x)$,
- (b) $(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$,
- (c) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g^2(x)}$, falls $g(x) \neq 0$,
- (d) $\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)}$, falls $g(x) \neq 0$.

Für Polynome $p(x) = \sum_{i=0}^n a_i x_i$ gilt:

$$p'(x) = \sum_{i=1}^n i a_i x^{i-1}.$$

Wir haben folgenden Spezialfall von Satz 3.1.4 (d):

$$\frac{d}{dx} \left(\frac{1}{x^n} \right) = \frac{-n}{x^{n+1}} = -n x^{-n-1} \quad \text{für alle } n \in \mathbb{N} \text{ und alle } x \neq 0.$$

3.1.5. Satz. Die Sinus- und die Cosinusfunktionen sind auf \mathbb{R} differenzierbar und es gilt:

- (a) $\sin' x = \cos x$,
- (b) $\cos' x = -\sin x$,
- (c) $\tan' x = \frac{1}{\cos^2 x}$, falls $x \neq (k + \frac{1}{2})\pi$, $k \in \mathbb{Z}$, und
- (d) $\cot' x = -\frac{1}{\sin^2 x}$, falls $x \neq k\pi$, $k \in \mathbb{Z}$.

Mit Hilfe der Eulerschen Formel aus Abschnitt 2.3.4 und Satz 3.1.5 kann man zeigen, dass

$$(3.1.6) \quad \frac{d}{dt} e^{iwt} = iw e^{iwt}, \quad t \in \mathbb{R}.$$

3.1.7. Satz (Kettenregel). Die Komposition $x \mapsto (f \circ g)(x) = f(g(x))$ zweier differenzierbarer Funktionen ist differenzierbar und es gilt

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

3.1.3. Höhere Ableitungen. Die Ableitung der Ableitung von f bezeichnen wir, falls sie existiert, mit f'' oder $\frac{d^2 f(x)}{dx^2}$. Allgemein definieren wir

$$\begin{aligned} f^{(0)}(x) &:= f(x) \\ f^{(1)}(x) &:= f'(x) \\ f^{(n+1)}(x) &:= \frac{df^{(n)}(x)}{dx} = \frac{d^{n+1} f(x)}{dx^{n+1}}. \end{aligned}$$

Hierbei heisst $f^{(n)}(x)$ die n -te Ableitung von f und f heisst n -mal differenzierbar wenn die n -te Ableitung existiert.

3.2. Anwendungen der Differentiation

3.2.1. Maxima und Minima.

Definition. Man sagt, eine auf $D \subseteq \mathbb{R}$ erklärte Funktion f hat in $a \in D$ ein *globales Maximum*, wenn $f(x) \leq f(a)$ für alle $x \in D$ gilt. Die Zahl $b \in D$ heißt *lokales Maximum von f* , wenn es eine ε -Umgebung $(b-\varepsilon, b+\varepsilon)$ von b gibt, so dass $f(x) \leq f(b)$ für alle $x \in D \cap (b-\varepsilon, b+\varepsilon)$. Analog definiert man *globales Minimum*, *lokales Minimum*. Jedes Minimum und jedes Maximum heißt *Extremum*.

3.2.1. Satz. *Ist f in einem offenen Intervall I differenzierbar, so gilt:*

$$x_0 \in I \text{ ist lokales Extremum} \implies f'(x_0) = 0.$$

Aus Satz 2.2 folgt, dass

- (1) Randpunkte von D ,
- (2) Punkte, an denen f nicht differenzierbar ist, und
- (3) *stationäre Punkte*, d.h. Punkte, an denen $f' = 0$ gilt,

Kandidaten für Extrema sind. WARNUNG: nicht jeder solche Punkt ist wirklich ein Extremum.

3.2.2. Der Mittelwertsatz.

3.2.2. Satz (Mittelwertsatz). *Ist die Funktion f auf $[a, b]$ stetig und auf (a, b) differenzierbar, dann gibt es einen Punkt $\xi \in (a, b)$ mit*

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

3.2.3. Satz. *Für eine auf dem Intervall $I \subseteq \mathbb{R}$ differenzierbare Funktion f gilt:*

- a) $f'(x) > 0$ auf $I \implies f$ ist auf I streng monoton wachsend,
- b) $f'(x) < 0$ auf $I \implies f$ ist auf I streng monoton fallend,
- c) $f'(x) \geq 0$ auf $I \iff f$ ist auf I monoton wachsend,
- d) $f'(x) \leq 0$ auf $I \iff f$ ist auf I monoton fallend,
- e) $f'(x) = 0$ auf $I \iff f$ ist konstant auf I .

3.2.4. Satz. *Eine auf (a, b) differenzierbare Funktion f hat im stationären Punkt $x_0 \in (a, b)$ ein lokales Maximum (bzw. lokales Minimum), wenn es ein $\varepsilon > 0$ gibt, so dass die Ableitung $f'(x)$ im Intervall $(x_0 - \varepsilon, x_0)$ positiv und im Intervall $(x_0, x_0 + \varepsilon)$ negativ ist (bzw. in $(x_0 - \varepsilon, x_0)$ negativ und in $(x_0, x_0 + \varepsilon)$ positiv).*

Definition. Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt *stetig differenzierbar*, wenn die Ableitung existiert und stetig ist. Sie heißt *n -mal stetig differenzierbar*, wenn die n -te Ableitung $f^{(n)}$ existiert und stetig ist.

3.2.5. Satz. Ist f auf (a, b) zweimal stetig differenzierbar und $x_0 \in (a, b)$ ein stationärer Punkt, dann gilt:

- a) $f''(x_0) < 0 \implies f$ hat in x_0 ein lokales Maximum,
 b) $f''(x_0) > 0 \implies f$ hat in x_0 ein lokales Minimum.

3.2.3. Wendepunkte.

Definition. Sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $f : I \rightarrow \mathbb{R}$ heisst *konvex*, wenn für alle $x_1, x_2 \in I$ und alle $\lambda \in (0, 1)$ gilt:

$$f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1).$$

Die Funktion heisst *konkav*, wenn $-f$ konvex ist.

3.2.6. Satz. Sei f auf einem Intervall I zweimal differenzierbar, dann gilt:

- a) $f'' \geq 0$ auf $I \implies f$ ist konvex,
 b) $f'' \leq 0$ auf $I \implies f$ ist konkav.

Definition. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion. Der Punkt $x_0 \in D$ heisst *Wendepunkt*, wenn es ein $\varepsilon > 0$ gibt, so dass f auf dem Intervall $(x_0 - \varepsilon, x_0)$ konvex (bzw. konkav) ist und auf dem Intervall $(x_0, x_0 + \varepsilon)$ konkav (bzw. konvex) ist.

3.2.7. Satz. Sei $f : I \rightarrow \mathbb{R}$ zweimal differenzierbar. Sei ferner $x_0 \in I$ ein Punkt, so dass $f''(x_0) = 0$ ist und die zweite Ableitung im Punkt x_0 ihr Vorzeichen wechselt. Dann ist x_0 ein Wendepunkt.

Folgerung. Sei $f : I \rightarrow \mathbb{R}$ dreimal stetig differenzierbar. Sei ferner $x_0 \in I$ ein Punkt, so dass $f''(x_0) = 0$ und $f'''(x_0) \neq 0$ gilt. Dann ist x_0 ein Wendepunkt.

3.2.4. Die Regeln von L'Hospital.

3.2.8. Satz (Verallgemeinerter Mittelwertsatz). Sind f, g in $[a, b]$ stetig und in (a, b) differenzierbar und ist $g'(x) \neq 0$ für alle $x \in (a, b)$, dann gibt es einen Punkt $x_0 \in (a, b)$, mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

3.2.9. Satz (Regeln von de l'Hospital). Sind f, g auf (a, b) differenzierbare Funktionen, $g'(x) \neq 0$ für alle $x \in (a, b)$, mit den Eigenschaften

- a) $f(x) \rightarrow 0, g(x) \rightarrow 0$ oder $f(x) \rightarrow \pm\infty, g(x) \rightarrow \pm\infty$ für $x \rightarrow b^-$.
 b) $\lim_{x \rightarrow b^-} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\}$, dann gilt:

$$\lim_{x \rightarrow b^-} \frac{f(x)}{g(x)} = \lim_{x \rightarrow b^-} \frac{f'(x)}{g'(x)}.$$

Entsprechendes gilt für $x \rightarrow a^+$ und $x \rightarrow \pm\infty$.

Bemerkung. Oftmals muss man die Regeln von L'Hospital mehrmals anwenden, solange bis man einen berechenbaren Grenzwert erhält.

3.2.5. Kurvendiskussion. Um eine Vorstellung von der Gestalt des Graphen $y = f(x)$ zu bekommen führt man eine Kurvendiskussion durch. Dazu muss man

- (1) Maximalen Definitions- und Wertebereich bestimmen,
- (2) Symmetrie, Periodizität testen,
- (3) Stetigkeit und Differenzierbarkeit prüfen,
- (4) Nullstellen und Vorzeichen bestimmen,
- (5) Extremwerte ermitteln,
- (6) Monotoniebereiche ermitteln,
- (7) Wendepunkte suchen,
- (8) Konvexität, Konkavität untersuchen,
- (9) Asymptoten, Grenzwerte bestimmen, $|x| \rightarrow \infty, x \rightarrow$ „kritische Stellen“,
- (10) Skizze des Graphen anfertigen.

Dabei ist es manchmal geschickter, die Punkte (5) und (6) in umgekehrter Reihenfolge zu behandeln, genau wie (7) und (8).

3.2.6. Nullstellen und Fixpunkte. Oftmals benötigt man die Nullstellen einer Funktion, vgl. Abschnitt 2.2.3. Jedoch kann man nur selten eine explizite Lösung angeben. Deshalb bestimmt man Folgen von Näherungslösungen der Gleichung $f(x) = 0$, d.h. man konstruiert eine Folge $(x_n)_{n \geq 0}$, so dass $f(x_n) \rightarrow 0$. Das Finden von Nullstellen von f ist äquivalent zum Finden von Fixpunkten von $g(x) = x + c f(x)$ für ein geeignetes $c \in \mathbb{R} \setminus \{0\}$. Sei die Funktion f auf $[a, b]$ definiert, dann heisst $x^* \in [a, b]$ *Fixpunkt* von f , wenn $f(x^*) = x^*$.

3.2.10. Satz. *Hat eine auf $[a, b]$ stetig differenzierbare Funktion f folgende Eigenschaften*

- a) $a \leq f(x) \leq b$ für alle $x \in [a, b]$,
- b) es gibt eine Konstante K mit $|f'(x)| \leq K < 1$ für alle $x \in [a, b]$,

dann gilt:

- (1) *Es gibt genau ein $x^* \in [a, b]$ mit $f(x^*) = x^*$.*
- (2) *Die Iterationsfolge*

$$(3.2.11) \quad x_{n+1} = f(x_n), \quad n \in \mathbb{N}_0$$

mit beliebigen Startwert $x_0 \in [a, b]$ konvergiert gegen den Fixpunkt x^ .*

- (3) *Es gilt die Abschätzung:*

$$|x_n - x^*| \leq \frac{K}{1-K} |x_n - x_{n-1}| \quad \text{für alle } n \in \mathbb{N}.$$

Das Newton-Verfahren. Man sucht die Nullstellen $f(x) = 0$ einer gegebenen Funktion mit Hilfe einer Fixpunktiteration für die Hilfsfunktion

$$F(x) = x - \frac{f(x)}{f'(x)},$$

unter der Voraussetzung $f'(x) \neq 0$ auf $[a, b]$. Wir erhalten die Iterationsfolge

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 \in [a, b].$$

3.2.12. Satz. Sei $x^* \in [a, b]$ eine Nullstelle der zweimal stetig differenzierbaren Funktion f . Falls $f'(x^*) \neq 0$ ist, gibt es ein kleines Intervall I , welches x^* enthält, so dass die oben definierte Iterationsfolge x_n für Startwerte x_0 aus diesem Intervall I gegen x^* konvergiert. Weiterhin gilt

$$|x_{n+1} - x^*| \leq M |x_n - x^*|^2 \quad \text{für} \quad M = \frac{\max\{|f''(x)| \mid x \in I\}}{\min\{|f'(x)| \mid x \in I\}}.$$

3.3. Umkehrfunktionen

3.3.1. Grundlagen. Umkehrabbildungen zwischen Mengen haben wir schon in Abschnitt 1.1.4 kennengelernt.

Definition. Sei $f : I \rightarrow \mathbb{R}$ eine Funktion und $D \subseteq I$. Man sagt, f ist über D umkehrbar, wenn zu jedem $y \in f(D)$ die Gleichung $y = f(x)$ genau eine Lösung $x \in D$ hat. Die Umkehrfunktion $g : f(D) \rightarrow D$ ordnet jedem y die eindeutige Lösung x von $y = f(x)$ zu, d.h.

$$x = g(y) \quad \Leftrightarrow \quad y = f(x).$$

Die Umkehrfunktion wird manchmal mit f^{-1} bezeichnet, was es schwierig macht, sie von $f^{-1} = \frac{1}{f}$ zu unterscheiden. Es gilt

$$\begin{aligned} f \text{ umkehrbar über } D &\iff f : D \rightarrow f(D) \text{ ist bijektiv} \\ &\iff f : D \rightarrow \mathbb{R} \text{ ist injektiv.} \end{aligned}$$

Sei g die Umkehrfunktion, dann folgt

$$\begin{aligned} f(x) &= f(g(y)) = y, \\ g(y) &= g(f(x)) = x. \end{aligned}$$

- 3.3.1. Satz.** (1) Jede strikt monotone Funktion $f : D \rightarrow \mathbb{R}$ ist umkehrbar. Jede über einem Intervall I stetig differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in I$ ist über I umkehrbar.
- (2) Ist f über D umkehrbar mit der Umkehrfunktion $g : f(D) \rightarrow \mathbb{R}$, dann liegen die Graphen von f und g symmetrisch zur Geraden $y = x$.

- (3) Die Umkehrfunktion $g: f(D) \rightarrow \mathbb{R}$ einer über einem Intervall $I \subseteq \mathbb{R}$ umkehrbaren und differenzierbaren Funktion f ist in allen Punkten $x \in f(I)$ mit $f'(g(x)) \neq 0$ differenzierbar und es gilt

$$g'(x) = \frac{1}{f'(g(x))}.$$

3.3.2. Wurzeln und rationale Exponenten. Für $n \in \mathbb{N}$ betrachten wir die Funktion

$$f(x) = x^n, \quad x \in \mathbb{R}.$$

- (1) Falls n gerade ist, d.h. $n = 2k$, ist f nicht auf \mathbb{R} umkehrbar, da $x^n = (-x)^n$. Aber für $x \in \mathbb{R}_0^+$ gilt $f'(x) > 0$ und somit ist x^{2k} auf \mathbb{R}_0^+ umkehrbar. Die Umkehrfunktion heisst n -te Wurzel, in Zeichen $\sqrt[n]{\cdot} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$, d.h. $x = \sqrt[n]{y}$ genau dann wenn $x^n = y$.
- (2) Falls n ungerade ist, d.h. $n = 2k + 1$, ist f auf ganz \mathbb{R} strikt monoton wachsend, denn $f'(x) = (2k + 1)x^{2k} > 0$, für $x \neq 0$. Damit ist die n -te Wurzel auf ganz \mathbb{R} definiert. Zusammenfassend haben wir:

$$y = \sqrt[n]{x} \iff y^n = x \begin{cases} \text{für alle } x \geq 0 & \text{falls } n \in \mathbb{N} \text{ gerade, bzw.} \\ \text{für alle } x \in \mathbb{R} & \text{falls } n \in \mathbb{N} \text{ ungerade.} \end{cases}$$

Falls $x \in \mathbb{R}_0^+$ und $\alpha = \frac{m}{n} \in \mathbb{Q}$, $m \in \mathbb{Z}$, $n \in \mathbb{N}$ setzen wir

$$\begin{aligned} x^{\frac{1}{n}} &:= \sqrt[n]{x}, \\ x^{\frac{m}{n}} &:= \left(x^{\frac{1}{n}}\right)^m. \end{aligned}$$

Aus der Definition ergeben sich sofort die Rechenregeln

$$\begin{aligned} x^\alpha x^\beta &= x^{\alpha+\beta}, \\ (x^\alpha)^\beta &= x^{\alpha\beta}, \\ x^\alpha y^\alpha &= (xy)^\alpha \end{aligned}$$

und die Formel für die Ableitung

$$\frac{d}{dx} (x^\alpha) = \alpha x^{\alpha-1}, \quad x > 0.$$

3.3.3. Arcusfunktionen. Wir betrachten jetzt die Umkehrfunktionen der Winkelfunktionen \sin , \cos , \tan und \cot . Dazu müssen wir geeignete Intervalle finden, auf denen diese Funktionen streng monoton sind.

Die Sinusfunktion ist auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ strikt monoton wachsend, denn $\sin'(x) = \cos x > 0$ falls $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Die Umkehrfunktion heisst *Arcussinus*, in Zeichen

\arcsin , und ist definiert durch

$$\begin{aligned} \arcsin: [-1, 1] &\longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ y = \arcsin x &\iff \sin y = x \text{ und } y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \end{aligned}$$

Für die Ableitung gilt:

$$\arcsin' x = \frac{1}{\sqrt{1-x^2}} \quad \text{für } -1 < x < 1.$$

Die Cosinusfunktion ist auf $[0, \pi]$ strikt monoton fallend. Die Umkehrfunktion heisst *Arcuscosinus*, in Zeichen \arccos , und ist definiert durch:

$$\begin{aligned} \arccos: [-1, 1] &\longrightarrow [0, \pi] \\ y = \arccos x &\iff \cos y = x \text{ und } y \in [0, \pi]. \end{aligned}$$

Aus $\cos \varphi = \sin(\frac{\pi}{2} - \varphi)$ folgt

$$\begin{aligned} \arccos x &= \frac{\pi}{2} - \arcsin x, \\ \arccos' x &= -\frac{1}{\sqrt{1-x^2}}, \quad \text{für } -1 < x < 1 \end{aligned}$$

Die Tangens- und die Cotangensfunktion sind auf dem Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ bzw. $(0, \pi)$ streng monoton. Die Umkehrfunktionen heissen *Arcustangens* und *Arcuscotangens* und sind definiert als

$$\begin{aligned} \arctan: \mathbb{R} &\longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \\ \arctan x = y &\iff \tan y = x \text{ und } y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right); \\ \operatorname{arccot}: \mathbb{R} &\longrightarrow (0, \pi), \\ \operatorname{arccot} x = y &\iff \cot y = x \text{ und } y \in (0, \pi) \end{aligned}$$

Aus $\cot \varphi = \tan(\frac{\pi}{2} - \varphi)$ folgt

$$\operatorname{arccot} x = \frac{\pi}{2} - \arctan x.$$

Für die Ableitungen ergibt sich für $x \in \mathbb{R}$:

$$\begin{aligned} \arctan' x &= \frac{1}{1+x^2}, \\ \operatorname{arccot}' x &= -\frac{1}{1+x^2}. \end{aligned}$$

3.4. Exponential- und Logarithmusfunktion

3.4.1. Die e -Funktion. Zur Erinnerung: die Exponentialfunktion $\exp x = e^x$ wurde in Abschnitt 2.5.4 als Grenzwert definiert.

3.4.1. Satz. *Die Exponentialfunktion hat folgende Eigenschaften:*

- (1) $e^0 = 1, e^x > 0$ für alle $x \in \mathbb{R}$.
 (2) Für alle $x \in \mathbb{R}$ gilt:

$$\frac{de^x}{dx} = e^x, \quad \text{kurz} \quad \exp' = \exp .$$

Somit ist die Exponentialfunktion insbesondere beliebig oft stetig differenzierbar.

- (3) Jede auf einem Intervall $I \subseteq \mathbb{R}$ differenzierbare Funktion f , welche $f'(x) = a f(x)$ für alle $x \in I$ erfüllt, ist von der Gestalt $f(x) = c e^{ax}$, wobei $c \in \mathbb{R}$ eine Konstante ist.
 (4) Es gelten:

$$e^{x+y} = e^x e^y, \quad \text{für alle } x, y \in \mathbb{R} ,$$

$$e^{-x} = \frac{1}{e^x}, \quad \text{für alle } x \in \mathbb{R} .$$

Wir haben für e -Funktion die Potenzschreibweise benutzt. Dies ist noch zu rechtfertigen. Es ist also zu zeigen dass

$$\exp(x) := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$$

die x -te Potenz von

$$e = \exp(1) = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

ist. Dazu benötigen wir

$$\exp(rx) = (\exp(x))^r \quad \forall x \in \mathbb{R}, r \in \mathbb{Q}.$$

Da die e -Funktion stetig ist, können wir Potenzen von e nun für alle reellen Zahlen $x \in \mathbb{R}$ definieren als:

$$e^x := \exp(x) = \lim_{n \rightarrow \infty} \exp(r_n) = \lim_{n \rightarrow \infty} e^{r_n},$$

wobei $r_n \rightarrow x, r_n \in \mathbb{Q}$.

3.4.2. Exponentielle Prozesse. In vielen Wachstums- und Zerfallsprozessen ist die Wachstums- oder Zerfallsgeschwindigkeit proportional zur vorhandenen Menge der im Prozess beschriebenen Größe $u(t)$, d.h. $u(t)$ erfüllt eine Differentialgleichung der Form

$$\dot{u}(t) = au(t) \quad \text{mit} \quad a \in \mathbb{R}.$$

Also gilt nach Satz 3.4.1 (3), dass

$$u(t) = ce^{at}, \quad \text{mit } c = u(0).$$

3.4.2. Satz. Die e -Funktion ist strikt monoton wachsend und konvex. Ferner gilt:

$$(1) \quad \lim_{x \rightarrow \infty} e^x = \infty, \quad \lim_{x \rightarrow -\infty} e^x = 0,$$

$$(2) \quad \lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \infty, \quad \lim_{x \rightarrow -\infty} x^n e^x = 0, \quad \text{für alle } n \in \mathbb{N}.$$

3.4.3. Der natürliche Logarithmus. Die e -Funktion wächst strikt monoton, also existiert nach Satz 3.3.1 (1) auf $(0, \infty)$ eine Umkehrfunktion. Diese heißt der *natürliche Logarithmus* und ist definiert durch:

$$\ln: (0, \infty) \longrightarrow \mathbb{R}$$

$$y = \ln x \quad \Longleftrightarrow \quad e^y = x.$$

Insbesondere gilt:

$$\ln(e^x) = x, \quad x \in \mathbb{R},$$

$$e^{\ln x} = x, \quad x > 0.$$

$$\ln e = 1,$$

$$\ln 1 = 0,$$

$$x \in (0, 1) \implies \ln x < 0,$$

$$x \in (1, \infty) \implies \ln x > 0,$$

$$\lim_{x \rightarrow 0^+} \ln x = -\infty,$$

$$\lim_{x \rightarrow \infty} \ln x = \infty.$$

3.4.3. Satz. Der natürliche Logarithmus ist strikt konkav und differenzierbar. Für $x, y > 0$ gilt

$$\ln' x = \frac{1}{x},$$

$$\ln(xy) = \ln x + \ln y,$$

$$\ln \frac{x}{y} = \ln x - \ln y.$$

Der natürliche Logarithmus wächst langsamer als jede Wurzelfunktion.

3.4.4. Satz. Für alle $n \in \mathbb{N}$ gilt

$$\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[n]{x}} = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} \sqrt[n]{x} \ln x = 0.$$

3.4.4. Allgemeine Exponentialfunktionen und Logarithmen. Sei $a > 0$. In Abschnitt 3.4.1 wurde gezeigt, dass für alle $r \in \mathbb{Q}$ und alle $x \in \mathbb{R}$ gilt

$$(e^x)^r = e^{rx}.$$

Die Wahl $x = \ln a$ liefert also

$$a^r = e^{r \ln a}.$$

Auf Grund der Stetigkeit der e -Funktion definiert man in Analogie zu (4.5)

$$a^x := e^{x \ln a} \quad \text{für alle } x \in \mathbb{R}, a > 0$$

Man nennt $x \mapsto a^x$ die *Exponentialfunktion zur Basis a* .

3.4.5. Satz. Für die Exponentialfunktion zur Basis a , mit $a > 0$, gilt für alle $x, y \in \mathbb{R}$ und $b > 0$

$$\begin{aligned} a^x a^y &= a^{x+y}, \\ (ab)^x &= a^x b^x, \\ (a^x)^y &= a^{xy}, \\ \ln(a^x) &= x \ln a, \\ \frac{d}{dx} a^x &= a^x \ln a. \end{aligned}$$

Wir können wir auch *Potenzfunktionen* von x für beliebige $\alpha \in \mathbb{R}$ definieren. Für alle $x > 0$ und $\alpha \in \mathbb{R}$ setzen wir wieder

$$x^\alpha = e^{\alpha \ln x}.$$

Man berechnet sofort

$$\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}.$$

Für alle $a > 0, a \neq 1$, ist die Funktion a^x umkehrbar, denn die Gleichung $y = e^{x \ln a}$ hat für $y > 0$ die eindeutige Lösung $x = \frac{\ln y}{\ln a}$. Die Umkehrfunktion heisst *Logarithmus zur Basis a* und ist definiert durch:

$$\begin{aligned} \log_a : (0, \infty) &\longrightarrow \mathbb{R} \\ y = \log_a x &\iff a^y = x. \end{aligned}$$

Es gelten:

$$\begin{aligned} \log_a(xy) &= \log_a(x) + \log_a(y), \\ \log'_a(x) &= \frac{1}{x \ln a}. \end{aligned}$$

3.4.5. Hyperbel- und Flächenfunktionen. Die Hyperbelfunktionen *sinus hyperbolicus*, *cosinus hyperbolicus*, *tangens hyperbolicus* und *cotangens hyperbolicus* sind auf \mathbb{R} definiert durch

$$\begin{aligned}\sinh x &:= \frac{e^x - e^{-x}}{2}, \\ \cosh x &:= \frac{e^x + e^{-x}}{2}, \\ \tanh x &:= \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \\ \coth x &:= \frac{\cosh x}{\sinh x} = \frac{e^x + e^{-x}}{e^x - e^{-x}} \quad \text{für } x \neq 0.\end{aligned}$$

Aus der Definition 3.4.5 und den Eigenschaften der e -Funktion erhält man folgende Rechenregeln, vgl. Abschnitt 2.3:

$$\begin{aligned}\sinh(-x) &= -\sinh x \\ \cosh(-x) &= \cosh x \\ \sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y, \\ \cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y, \\ 1 &= \cosh^2 x - \sinh^2 x.\end{aligned}$$

Für die Ableitungen gilt:

$$\begin{aligned}\sinh' x &= \cosh x, & \cosh' x &= \sinh x, \\ \tanh' x &= \frac{1}{\cosh^2 x}, & \coth' x &= -\frac{1}{\sinh^2 x}.\end{aligned}$$

Die Funktion \sinh ist auf \mathbb{R} strikt monoton wachsend, also existiert eine Umkehrfunktion. Diese heisst *area sinus hyperbolicus* und wird mit $\operatorname{ar} \sinh x$ bezeichnet. Die Funktion \cosh ist auf \mathbb{R}_0^+ strikt monoton wachsend mit Bild $[1, \infty)$. Die Umkehrfunktion heisst *area cosinus hyperbolicus* und wird mit $\operatorname{ar} \cosh$ bezeichnet. Es gilt:

$$\begin{aligned}\operatorname{ar} \sinh x &= \ln(x + \sqrt{x^2 + 1}), \quad x \in \mathbb{R}, \\ \operatorname{ar} \cosh x &= \ln(x + \sqrt{x^2 - 1}), \quad x \geq 1.\end{aligned}$$

Die Kettenregel liefert

$$\begin{aligned}\operatorname{ar} \sinh' x &= \frac{1}{\sqrt{x^2 + 1}}, \quad x \in \mathbb{R}, \\ \operatorname{ar} \cosh' x &= \frac{1}{\sqrt{x^2 - 1}}, \quad x > 1.\end{aligned}$$

KAPITEL 4

Integration

Die Integration ist die Umkehroperation zur Differentiation. Es wird also das Problem behandelt, wie man aus der Kenntnis der Ableitung einer Funktion die Funktion selbst wiederherstellt. Eine entscheidende Rolle bei der Lösung dieser Aufgabe spielt der Mittelwertsatz. Es wird sich zeigen, dass das Integral das geeignete Mittel zur Berechnung von Flächen- und Rauminhalten ist.

4.1. Das bestimmte Integral

Motivation: Sei $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ eine Zerlegung von $[a, b]$, dann existieren nach dem Mittelwertsatz 3.2.2 Zahlen $\xi_i \in (x_{i-1}, x_i)$ mit

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1}).$$

Demzufolge gilt:

$$(*) \quad f(b) - f(a) = \sum_{i=1}^n f'(\xi_i)(x_i - x_{i-1}).$$

Wenn die Zerlegung fein genug und f stetig ist, kann man ξ_i durch einen beliebigen Punkt in Intervall $[x_{i-1}, x_i]$ ersetzen und erhält eine gute Approximation von $f(b) - f(a)$, d.h. die rechte Seite von (*) ist eine gute Approximation des Integrals von $f'(x)$.

4.1.1. Die Definition des bestimmten Integrals. Sei f eine auf dem Intervall $[a, b]$ definierte, beschränkte Funktion, die an höchstens endlich vielen Stellen nicht stetig ist. Solche Funktionen nennt man *stückweise stetig*. Sei durch

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$$

eine Zerlegung von $[a, b]$ in n Teilintervalle $[x_{i-1}, x_i]$ gegeben und seien $\xi_i \in [x_{i-1}, x_i]$ beliebige Zwischenpunkte. Dann heisst

$$Z_n := \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

die zugehörige *Riemannsche Summe von f* . Der Grenzwert der Folge $(Z_n)_{n \geq 1}$ existiert (mit neuer Auswahl der Punkte x_0, \dots, x_n für jedes n), sofern die maximale Länge der Teilintervalle $[x_{i-1}, x_i]$ mit $n \rightarrow \infty$ gegen Null strebt. Weiterhin

ist der Grenzwert unabhängig von der gewählten Zerlegung des Intervalls $[a, b]$ und der Zwischenpunkte ξ_i . Der Grenzwert der Folge $(Z_n)_{n \geq 1}$ heisst das *bestimmte Integral von f über $[a, b]$* und wird mit $\int_a^b f(x) dx$ bezeichnet, d.h.

$$\int_a^b f(x) dx := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1}).$$

Um Fallunterscheidungen zu vermeiden setzen wir

$$\int_a^a f(x) dx := 0,$$

$$\int_b^a f(x) dx := - \int_a^b f(x) dx, \quad b > a.$$

4.1.2. Geometrische Interpretation. Die Riemannsche Summe einer positiven Funktion f ist eine Summe von Rechteckflächen, die die Fläche I unter dem Graph von f immer genauer approximiert. Also gilt

$$I = \int_a^b f(x) dx.$$

Genau genommen wird erst durch das Integral der Flächeninhalt definiert.

4.1.3. Elementare Integrationsregeln und der Mittelwertsatz.

4.1.1. Satz. Seien f, g stückweise stetige Funktion auf einem Intervall I , seien $a, b, c \in I$ und $\alpha, \beta \in \mathbb{R}$. Es gelten:

$$(a) \quad \int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx,$$

$$(b) \quad \int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx,$$

$$(c) \quad f(x) \leq g(x) \implies \int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

4.1.2. Satz. Sei f auf $[a, b]$ stückweise stetig.

a) Aus $m \leq f(x) \leq M$ für alle $x \in [a, b]$ folgt

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a).$$

b) Es gilt die Ungleichung

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

4.1.3. Satz. (Mittelwertsatz der Integralrechnung) Sei f auf $[a, b]$ stetig und g auf $[a, b]$ nichtnegativ und stückweise stetig. Dann gibt es ein $\xi \in [a, b]$ mit

$$\int_a^b f(x)g(x) dx = f(\xi) \int_a^b g(x) dx.$$

Der Spezialfall $g(x) = 1$ zeigt, dass es ein $\xi \in [a, b]$ gibt, so dass

$$\int_a^b f(x) dx = f(\xi)(b-a).$$

4.1.4. Differentiation und Integration.

Definition. Man nennt eine auf dem Intervall I differenzierbare Funktion F eine *Stammfunktion* von f , wenn $F'(x) = f(x)$ für alle $x \in I$ gilt.

4.1.4. Satz (Hauptsatz der Differential- und Integralrechnung). *Ist f eine auf dem Intervall I stetige Funktion, dann gilt:*

a) Die durch

$$F_a(x) := \int_a^x f(t) dt \quad a, x \in I$$

definierte Funktion ist eine Stammfunktion von f , d.h.

$$\frac{d}{dx} \left(\int_a^x f(t) dt \right) = f(x).$$

b) Jede andere Stammfunktion F von f hat die Form $F(x) = F_a(x) + c$ für ein $c \in \mathbb{R}$.

c) Mit einer beliebigen Stammfunktion F von f gilt:

$$\int_a^b f(x) dx = F(x) \Big|_a^b := F(b) - F(a).$$

Berechnung des bestimmten Integrals:

- (1) Finde eine Stammfunktion F von f , d.h. $F' = f$.
- (2) $\int_a^b f(x) dx = F(b) - F(a)$

Definition. Die Menge aller Stammfunktionen von f wird mit $\int f dx$ bezeichnet und heisst *unbestimmtes Integral* von f .

- Nach Satz 4.1.4 a) gilt

$$\int f(x) dx = F + c,$$

wobei $c \in \mathbb{R}$ eine Konstante ist und F eine feste Stammfunktion von f . Somit haben wir

$$\int f(x) dx = F + c \Leftrightarrow F' = f.$$

Aus den Differentiationsformeln folgen somit folgende Integrationsformeln:

$F(x)$	$f(x) = F'(x)$	Bemerkungen
$\frac{1}{n+1} x^{n+1}$	x^n	$n \in \mathbb{Z} \setminus \{-1\}$, $x \neq 0$ falls $n < 0$
$\frac{1}{a+1} x^{a+1}$	x^a	$a \in \mathbb{R} \setminus \{-1\}$, $x > 0$
$\ln x $	x^{-1}	$x \neq 0$
$-\cos x$	$\sin x$	
$\sin x$	$\cos x$	
$\tan x$	$\frac{1}{\cos^2 x}$	$x \neq (k + \frac{1}{2})\pi, k \in \mathbb{Z}$
$\cot x$	$\frac{-1}{\sin^2 x}$	$x \neq k\pi, k \in \mathbb{Z}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$ x \leq 1$
$\arctan x$	$\frac{1}{1+x^2}$	
$\frac{1}{a} e^{ax}$	e^{ax}	$a \neq 0$
$\cosh x$	$\sinh x$	
$\sinh x$	$\cosh x$	
$\operatorname{ar} \tanh x = \frac{1}{2} \ln \frac{1+x}{1-x}$	$\frac{1}{1-x^2}$	$ x < 1$
$\operatorname{ar} \sinh x = \ln(x + \sqrt{1+x^2})$	$\frac{1}{\sqrt{1+x^2}}$	
$\operatorname{ar} \cosh x = \ln(x + \sqrt{x^2-1})$	$\frac{1}{\sqrt{x^2-1}}$	$x \geq 1$

4.2. Integrationsregeln

4.2.1. Linearität. Aus $F' = f, G' = g$ folgt $af + bg = aF' + bG'$. Somit gilt für das unbestimmte Integral

$$\int (af(x) + bg(x)) dx = a \int f(x) dx + b \int g(x) dx.$$

4.2.2. Partielle Integration. Die Produktregel (Satz 3.1.4 (b)) liefert, dass uv eine Stammfunktion von $u'v + uv'$ ist.

4.2.1. Satz (Partielle Integration). *Seien $u, v: I \rightarrow \mathbb{R}$ differenzierbar, dann gilt*

$$\int u'(x)v(x) dx = uv - \int u(x)v'(x) dx .$$

Für das bestimmte Integral erhalten wir

$$\int_a^b u'(x)v(x) dx = u(x)v(x) \Big|_a^b - \int_a^b u(x)v'(x) dx .$$

Beispiel. Aus Satz 4.2.1 folgt mit $u' = 1$

$$\int v(x) dx = xv - \int xv'(x) dx,$$

insbesondere also:

$$\int \ln x dx = x \ln x - \int x \frac{1}{x} dx = x(\ln x - 1) + c .$$

Beispiel. Für Integrale der Form

$$\begin{array}{ll} S_n := \int_a^b (\sin x)^n dx & C_n := \int_a^b (\cos x)^n dx \\ A_n := \int_a^b x^n \sin x dx & B_n := \int_a^b x^n \cos x dx \\ E_n := \int_a^b x^n e^x dx & L_n := \int_a^b (\ln x)^n dx \end{array}$$

kann man durch wiederholtes Anwenden von partieller Integration eine Rekursionsformel herleiten.

4.2.3. Substitutionsmethode. Aus der Kettenregel (Satz 3.1.7) folgt

4.2.2. Satz. *Es sei $f: I \rightarrow \mathbb{R}$ stetig mit Stammfunktion F und $g: I \rightarrow \mathbb{R}$ differenzierbar, dann gilt*

$$\int f(g(x)) g'(x) dx = F(g(x)) + c,$$

und für das bestimmte Integral

$$\int_a^b f(g(x)) g'(x) dx = F(g(x)) \Big|_a^b = \int_{g(a)}^{g(b)} f(t) dt.$$

Es gibt zwei Möglichkeiten, Satz 4.2.2 anzuwenden.

- (1) Berechnung von $\int f(g(x)) g'(x) dx$.
 - Substitution $g(x) = t$, $g'(x) dx = dt$,
 - Berechnung von $\int f(t) dt = F(t) + c$,
 - Resubstitution $t = g(x)$ liefert $\int f(g(x)) g'(x) dx = F(g(x)) + c$.
- (2) Berechnung von $\int f(x) dx$.
 - Substitution $x = g(t)$, $dx = g'(t) dt$ für eine geeignete umkehrbare Funktion g mit Umkehrfunktion h ,
 - Berechnung von $\int f(g(t)) g'(t) dt = H(t) + c$,
 - Resubstitution $t = h(x)$ liefert $\int f(x) dx = H(h(x)) + c$.

Beispiel. Man nennt $\frac{d}{dx} \ln(f(x)) = \frac{f'(x)}{f(x)}$ die *logarithmische Ableitung* der Funktion f . Es gilt

$$\int \frac{f'(x)}{f(x)} dx = \ln(f(x)).$$

Speziell

$$\int \tan x dx = - \int \frac{\cos' x}{\cos x} dx = - \ln(\cos x).$$

Beispiel. Sei $f: I \rightarrow J \subset \mathbb{R}$ eine umkehrbare Funktion mit Umkehrfunktion $g: J \rightarrow I$, und sei F eine Stammfunktion von f . Mit der Substitution $x = f(t)$ und $dx = f'(t) dt$ erhält man

$$\int g(x) dx = \int t f'(t) dt = t f(t) - \int f(t) dt = g(x) x - F(g(x)).$$

Beispiel. Für $k \in \mathbb{Z} \setminus \{0\}$ und $\varphi \in \mathbb{R}$ gilt

$$\int_0^{2\pi} \sin(kx + \varphi) dx = -\frac{1}{k} \cos(kx + \varphi) \Big|_0^{2\pi}.$$

Somit erhält man für $\varphi = 0$ bzw. $\varphi = -\pi$ die Formeln

$$\int_0^{2\pi} \sin(kx) dx = 0 ,$$

$$\int_0^{2\pi} \cos(kx) dx = 0 .$$

falls $k \in \mathbb{Z} - \{0\}$. Mit Hilfe des Additions-Theorems 2.3.1 erhalten wir

$$\begin{aligned} \sin mx \sin nx &= \frac{1}{2} (\cos(m-n)x - \cos(m+n)x) , \\ \sin mx \cos nx &= \frac{1}{2} (\sin(m+n)x + \sin(m-n)x) , \\ \cos mx \cos nx &= \frac{1}{2} (\cos(m+n)x + \cos(m-n)x) . \end{aligned}$$

Und somit folgende *Orthogonalitätsbeziehungen*

$$\int_0^{2\pi} \sin mx \sin nx dx = \begin{cases} \pi , & \text{falls } m = n \neq 0 , \\ -\pi , & \text{falls } m = -n \neq 0 , \\ 0 , & \text{sonst,} \end{cases}$$

$$\int_0^{2\pi} \cos mx \cos nx dx = \begin{cases} 2\pi , & \text{falls } m = n = 0 , \\ \pi , & \text{falls } m = \pm n \neq 0 , \\ 0 , & \text{sonst, und} \end{cases}$$

$$\int_0^{2\pi} \cos mx \sin nx dx = 0 .$$

4.2.4. Substitution für bestimmte Integrale. Bei der Berechnung von bestimmten Integralen muss für die Gültigkeit der Formel

$$\int_a^b f(g(x))g'(x) dx = \int_{g(a)}^{g(b)} f(t) dt$$

überprüft werden ob:

- a) $f : I \rightarrow \mathbb{R}$ stetig und beschränkt ist, wobei I ein abgeschlossenes Intervall ist und
- b) $g : [a, b] \rightarrow I$ stetig differenzierbar ist.

4.2.5. Symmetrien. Integrale lassen sich oft leichter berechnen, wenn man Symmetrien des Integranden und des Integrationsbereiches ausnutzt. Für eine gerade Funktion, d.h. $f(-x) = f(x)$ gilt

$$\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx$$

und für eine ungerade Funktion, d.h. $f(-x) = -f(x)$, gilt

$$\int_{-a}^a f(x) dx = 0.$$

4.3. Integration rationaler Funktionen

4.3.1. Partialbruchzerlegung. Eine *Partialbruchzerlegung* einer rationalen Funktion besteht aus folgenden Schritten:

- (1) Polynomdivision

$$f(x) = h(x) + \frac{r(x)}{q(x)} \quad \text{mit Grad } r < \text{Grad } q,$$

vergleiche Satz 2.2.10. In den folgenden Schritten betrachten wir daher nur rationale Funktionen $f = \frac{p}{q}$ mit $\text{Grad } p < \text{Grad } q$.

- (2) Faktorisieren von $q(x)$ wie in Satz 2.2.7, also

$$q(x) = c(x - b_1)^{k_1} \cdots (x - b_r)^{k_r} q_1(x)^{l_1} \cdots q_s(x)^{l_s}$$

mit paarweise verschiedenen reellen Nullstellen b_j der Vielfachheit k_i und verschiedenen quadratischen Polynomen q_j , die keine reellen Nullstellen besitzen.

- (3) Es gelte $\text{Grad } p < \text{Grad } q$ wie in (1), und q sei wie in (2) faktorisiert. Wir machen den Ansatz

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{A_{1,1}}{x - b_1} + \cdots + \frac{A_{1,k_1}}{(x - b_1)^{k_1}} + \cdots + \frac{A_{r,k_r}}{(x - b_r)^{k_r}} \\ &+ \frac{B_{1,1}x + C_{1,1}}{q_1(x)} + \cdots + \frac{B_{1,l_1}x + C_{1,l_1}}{q_1(x)^{l_1}} + \cdots + \frac{B_{s,l_s}x + C_{s,l_s}}{q_s(x)^{l_s}} \end{aligned}$$

mit $A_{i,j}, B_{i,j}, C_{i,j} \in \mathbb{R}$. Die einzelnen Summanden heissen *Partialbrüche*.

- (4) Um die Koeffizienten $A_{i,j}, B_{i,j}, C_{i,j}$ zu bestimmen, multipliziert man die Gleichung in (3) auf beiden Seiten mit dem Nenner $q(x)$ und erhält eine Gleichung von Polynomen in x vom Grad $\text{Grad } q - 1$. Anschließend kann man entweder

- Grad q verschiedene spezielle Werte für x einsetzen, z.B. $x = b_1, \dots, b_r$ oder $x = 0, 1, \dots$, oder

- Beide Seiten ausmultiplizieren und die Koeffizienten vor $x^0, \dots, x^{\text{Grad } q-1}$ auf beiden Seiten vergleichen, siehe Satz 2.2.1.

In beiden Fällen erhält man lineare Gleichungssysteme in den Unbekannten $A_{i,j}, B_{i,j}, C_{i,j}$. Ein Lösungsverfahren dafür lernen wir in Abschnitt 6.1.

4.3.2. Integration der Partialbrüche. Die dabei auftretenden Integrale berechnen sich wie folgt:

$$(1) \quad \int \frac{dx}{x \pm a} = \ln|x \pm a| + c,$$

$$(2) \quad \int \frac{dx}{(x \pm a)^k} = \frac{1}{k-1}(x \pm a)^{1-k} + c, \quad k \in \mathbb{Z} \setminus \{1\},$$

und falls $4q - p^2 > 0$, gilt

$$(3) \quad \int \frac{dx}{x^2 + px + q} = \frac{2}{\sqrt{4q - p^2}} \arctan \frac{2x + p}{\sqrt{4q - p^2}} + c,$$

$$(4) \quad \int \frac{ax + b}{x^2 + px + q} dx = \frac{a}{2} \ln(x^2 + px + q) + \left(b - \frac{ap}{2}\right) \int \frac{dx}{x^2 + px + q},$$

und für $k > 1$ benutzen wir

$$(5) \quad \int \frac{dx}{(x^2 + px + q)^k} = \frac{2x + p}{(k-1)(4q - p^2)(x^2 + px + q)^{k-1}} + \frac{2(2k-3)}{(k-1)(4q - p^2)} \int \frac{dx}{(x^2 + px + q)^{k-1}},$$

$$(6) \quad \int \frac{ax + b}{(x^2 + px + q)^k} dx = -\frac{a}{2(k-1)(x^2 + px + q)^{k-1}} + \left(b - \frac{ap}{2}\right) \int \frac{dx}{(x^2 + px + q)^k}.$$

Gleichung (5) ist gegebenenfalls mehrfach rekursiv anzuwenden.

4.3.3. Integration von $R(e^{ax})$. Sei R eine rationale Funktion, dann kann man das Integral

$$\int R(e^{ax}) dx$$

durch die Substitution $t = e^{ax}, dx = \frac{1}{at} dt$ auf das Integral

$$\int R(t) \frac{1}{at} dt$$

zurückführen und somit berechnen, denn $\frac{R(t)}{at}$ ist wiederum eine rationale Funktion. Rationale Ausdrücke in Hyperbelfunktionen lassen sich als rationale Ausdrücke in e^x umschreiben und genauso behandeln.

4.3.4. Integration von $R\left(x, \sqrt[k]{\frac{ax+b}{cx+g}}\right)$. Sei $R(x, y)$ ein rationaler Ausdruck in x und y , d.h. $R(x, y)$ entsteht aus x, y und Konstanten allein durch die vier Grundrechenarten. Seien $a, b, c, d \in \mathbb{R}$ mit $ag - bc \neq 0$. Ein Integral vom Typ

$$\int R\left(x, \sqrt[k]{\frac{ax+b}{cx+g}}\right) dx$$

wird mit der Substitution

$$t = \sqrt[k]{\frac{ax+b}{cx+g}}, \quad \text{d.h.,} \quad x = \frac{gt^k - b}{a - ct^k}, \quad dx = k(ag - bc) \frac{t^{k-1}}{(a - ct^k)^2} dt$$

in ein Integral einer rationalen Funktion überführt.

4.3.5. Integration von $R(\sin x, \cos x)$. Sei $R(x, y)$ ein rationaler Ausdruck. Durch die Substitution

$$x = 2 \arctan t, \quad dx = \frac{2}{1+t^2} dt$$

und somit

$$\cos x = \frac{1-t^2}{1+t^2}, \quad \sin x = \frac{2t}{1+t^2}$$

überführt man das Integral $\int R(\sin x, \cos x) dx$ in ein Integral über eine rationale Funktion in t .

4.3.6. Integration von $R(x, \sqrt{ax^2 + bx + c})$. Seien $a, b, c \in \mathbb{R}$ mit $a \neq 0$ und $4ac - b^2 \neq 0$; falls $a < 0$ gelte $4ac - b^2 < 0$. Durch die Substitution

$$x = \frac{\sqrt{|4ac - b^2|}}{2a} u - \frac{b}{2a}, \quad dx = \frac{\sqrt{|4ac - b^2|}}{2a} du$$

geht das Integral

$$\int R(x, \sqrt{ax^2 + bx + c}) dx$$

in ein Integral von einem der folgenden Typen über.

- $\int R(u, \sqrt{u^2 + 1}) dx$. Substituiere $u = \sinh t$ und fahre fort wie in 4.3.3.
- $\int R(u, \sqrt{u^2 - 1}) dx$. Substituiere $u = \cosh t$ und fahre fort wie in 4.3.3.
- $\int R(u, \sqrt{1 - u^2}) dx$. Substituiere $u = \sin t$ und fahre fort wie in 4.3.5.

4.4. Uneigentliche Integrale

Wir wollen den Integralbegriff erweitern auf unbeschränkte Integrationsintervalle $[a, b)$ und unbeschränkte Funktionen.

4.4.1. Definition der uneigentlichen Integrale. Sei $b \in \mathbb{R} \cup \{\infty\}$. Die Funktion f sei auf dem Intervall $[a, b)$ definiert und auf jedem abgeschlossenen Teilintervall $[a, c]$ für $c \in (a, b)$ stückweise stetig. Falls der Grenzwert

$$\lim_{c \rightarrow b^-} \int_a^c f(x) dx$$

existiert, wird das *uneigentliche Integral* $\int_a^b f(x) dx$ definiert durch

$$\int_a^b f(x) dx := \lim_{c \rightarrow b^-} \int_a^c f(x) dx.$$

In diesem Fall sagt man, dass das uneigentliche Integral *konvergiert*. Anderenfalls sagt man, dass es *divergiert*.

Analog definiert man in der entsprechenden Situation für ein rechts halboffenes Intervall $(a, b]$

$$\int_a^b f(x) dx := \lim_{c \rightarrow a^+} \int_a^b f(x) dx.$$

Beispiel. Für $\alpha \in \mathbb{R}$ gilt

$$\int_1^c x^\alpha dx = \begin{cases} \log c & \text{falls } \alpha = -1, \text{ und} \\ \frac{c^{\alpha+1}-1}{\alpha+1} & \text{sonst.} \end{cases}$$

Also gilt

$$\int_1^\infty x^\alpha dx = \begin{cases} -\frac{1}{\alpha+1} & \text{falls } \alpha < -1, \text{ und} \\ \text{divergiert} & \text{sonst.} \end{cases}$$

Genauso sieht man

$$\int_0^1 x^\alpha dx = \begin{cases} \frac{1}{\alpha+1} & \text{falls } \alpha > -1, \text{ und} \\ \text{divergiert} & \text{sonst.} \end{cases}$$

4.4.2. Ein Konvergenz-Test.

4.4.1. Satz. Sind $a, b \in (0, \infty)$, ist f auf $[a, \infty)$ und g auf $(0, b]$ stückweise stetig und sind α und $K \in \mathbb{R}$, dann gilt:

- (1) Falls $\alpha < -1$ und $|f(x)| \leq Kx^\alpha$ für alle $x \in [a, \infty)$, konvergiert das Integral

$$\int_a^\infty f(x) dx,$$

(2) Falls $\alpha > -1$ und $|f(x)| \leq Kx^\alpha$ für alle $x \in (0, b]$, konvergiert das Integral

$$\int_0^b f(x) dx .$$

4.4.3. Integrale mit mehreren Ausnahmestellen. Sei $k \geq 1$, seien $a = a_0 < a_1 < \dots < a_n = b$ gegeben, und sei $f: (a, b) \setminus \{a_1, \dots, a_{n-1}\} \rightarrow \mathbb{R}$ auf jedem abgeschlossenen Teilintervall $[c, d]$ des Definitionsbereiches stückweise stetig und beschränkt. Seien $a_0 < c_1 < a_1 < \dots < c_n < a_n$ Zwischenstellen. Dann konvergiert das uneigentliche Integral

$$\int_a^b f(x) dx = \sum_{i=1}^n \int_{a_{i-1}}^{c_i} f(x) dx + \sum_{i=1}^n \int_{c_i}^{a_i} f(x) dx$$

falls jedes einzelne Integral auf der rechten Seite konvergiert, ansonsten *divergiert* es.

4.4.4. Der Cauchy-Hauptwert. Sei $f: [a, b] \setminus \{c\} \rightarrow \mathbb{R}$ gegeben. Es kann passieren, dass die beiden uneigentlichen Integrale

$$\int_a^c f(x) dx \quad \text{und} \quad \int_c^b f(x) dx$$

divergieren, aber dass der Grenzwert

$$\text{CHW} \int_a^b f(x) dx = \lim_{\varepsilon \rightarrow 0^+} \left(\int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right)$$

existiert. Er heißt der *Cauchysche Hauptwert* des uneigentlichen Integrals.

4.5. Längen, Flächen und Volumina

In diesem Abschnitt interessieren wir uns für die Geometrie von Kurven in der Ebene, und von (Rotations-) Körpern im Raum \mathbb{R}^3 .

4.5.1. Parameterdarstellungen von Kurven. Unter einer *Parameterdarstellung* einer *Kurve* versteht man eine differenzierbare Abbildung

$$\vec{r}: [a, b] \rightarrow \mathbb{R}^2 \quad \text{mit} \quad t \mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

eines Intervalls $I = [a, b]$ in die Ebene, dargestellt in kartesischen Koordinaten. Man nennt t den *Parameter* und $[a, b]$ das *Parameterintervall*.

Die obige Darstellung ist äquivalent zu den Gleichungen

$$x = x(t), \quad y = y(t), \quad t \in [a, b] .$$

Jede Kurve besitzt unendlich viele Parameterdarstellungen: Sei $\vec{r}: I \rightarrow \mathbb{R}^2$ eine Kurve und sei $f: J \rightarrow I$ eine differenzierbare und differenzierbar umkehrbare Funktion, dann ist

$$\vec{\gamma} = \vec{r} \circ f: J \rightarrow \mathbb{R}^2 \quad \text{mit} \quad \vec{\gamma}(t) = \begin{pmatrix} x(f(t)) \\ y(f(t)) \end{pmatrix}$$

eine andere Parameterdarstellung derselben Kurve. Man nennt $\vec{\gamma}$ eine *Umparametrisierung* von \vec{r} und f eine *Parametertransformation*. Als Kurve K bezeichnet man die Menge aller Umparametrisierungen einer gegebenen Parameterdarstellung.

Beispielsweise haben alle Parameterdarstellungen von K das gleiche Bild in \mathbb{R}^2 . Man sagt: "das Bild einer Kurve hängt nur von der Kurve, nicht von der Parameterdarstellung ab."

4.5.2. Tangente und Normale. Zu jeder Parameterdarstellung $\vec{r}: I \rightarrow \mathbb{R}^2$ einer Kurve K definiert man den *Tangentenvektor*

$$\dot{\vec{r}}(t) := \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} : I \rightarrow \mathbb{R}^2,$$

wobei der Punkt die Ableitung nach t bezeichnet. Der Vektor

$$\vec{n}(t) := \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}$$

heißt *Normalenvektor* zur Zeit t .

- *Physikalische Interpretation:* Wenn $\vec{r}(t)$ die Bewegung eines Massenpunktes auf einer Kurve beschreibt, dann beschreibt $\dot{\vec{r}}(t)$ die Geschwindigkeit des Punktes zum Zeitpunkt t .
- *Geometrische Interpretation:* Sei in einem Kurvenpunkt $\vec{r}(t_0)$ der Tangentialvektor $\dot{\vec{r}}(t_0) \neq \vec{0}$. Dann haben zum Zeitpunkt t_0 die Tangente bzw. Normale an die Kurve K die (Parameter-) Darstellungen

$$s \mapsto \vec{r}(t_0) + s \dot{\vec{r}}(t_0) = \begin{pmatrix} x(t_0) + s \dot{x}(t_0) \\ y(t_0) + s \dot{y}(t_0) \end{pmatrix} \quad (\text{Tangente}), \text{ bzw.}$$

$$s \mapsto \vec{r}(t_0) + s \vec{n}(t_0) = \begin{pmatrix} x(t_0) - s \dot{y}(t_0) \\ y(t_0) + s \dot{x}(t_0) \end{pmatrix} \quad (\text{Normale}),$$

wobei $s \in \mathbb{R}$ der Geradenparameter ist.

4.5.3. Die Bogenlänge. Sei $a = t_0 < t_1 < \dots < t_n = b$ eine äquidistante Zerlegung von $[a, b]$, d.h. $t_i - t_{i-1} = \Delta t$, $\forall i = 1, \dots, n$. Die Kurve K wird durch die Sekanten, die $(x(t_i), y(t_i))$ und $(x(t_{i+1}), y(t_{i+1}))$ verbinden, angenähert. Die Länge dieser Approximation ist

$$P_n := \sum_{i=1}^n \|\vec{r}(t_i) - \vec{r}(t_{i-1})\|.$$

Falls der Grenzwert $n \rightarrow \infty$ der Folge $(P_n)_{n \geq 1}$ existiert wird er *Länge der Kurve* K genannt.

4.5.1. Satz. Sei $\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}: [a, b] \rightarrow \mathbb{R}^2$ eine Parameterdarstellung einer Kurve K , so dass \dot{x} und \dot{y} beschränkt und stückweise stetig sind. Dann hat die Kurve K die Länge

$$L = \int_a^b \|\dot{\vec{r}}(t)\| dt = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt.$$

4.5.2. Folgerung. Der Graph $y = f(x)$ einer stetig differenzierbaren Funktion $f: [a, b] \rightarrow \mathbb{R}$ hat die Länge

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Bemerkung. Die Länge einer Kurve hängt nicht von der Parameterdarstellung ab. Sei etwa $\vec{r}: [a, b] \rightarrow \mathbb{R}^2$ eine Parameterdarstellung und sei $\vec{\gamma} = \vec{r} \circ f: [c, d] \rightarrow \mathbb{R}^2$ eine Umparametrisierung zur Parametertransformation $f: [c, d] \rightarrow [a, b]$. Dann folgt aus der Substitutionsregel aus Satz 4.2.2, dass

$$\int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt = \int_c^d \sqrt{\left(\frac{d(x(f(s)))}{ds}\right)^2 + \left(\frac{d(y(f(s)))}{ds}\right)^2} ds.$$

4.5.4. Die Krümmung einer Kurve.

Definition. Eine Parameterdarstellung $x = x(t)$, $y = y(t)$, $t \in [a, b]$ einer Kurve heißt *regulär*, wenn die Funktionen $t \mapsto x(t)$, $t \mapsto y(t)$ stetig differenzierbar sind und $\dot{x}^2(t) + \dot{y}^2(t) \neq 0$ für alle $t \in [a, b]$ gilt (dabei sind die Ableitungen in den Endpunkten einseitige Ableitungen).

Es sei $\vec{r}(t) = (x(t), y(t))^T$ eine reguläre Parameterdarstellung einer Kurve K . Mit $\varphi(t)$ bezeichnen wir den positiv gemessenen Winkel zwischen der positiven x -Achse und dem Tangentialvektor $\dot{\vec{r}}$. Sei $s(t) := \int_a^t \sqrt{\dot{x}^2(\tau) + \dot{y}^2(\tau)} d\tau$ die Länge des Kurvenstücks über dem Parameterintervall $[a, t]$. Die Änderung $\Delta\varphi$ bezogen auf die Änderung der Länge Δs ist ein Maß für die durchschnittliche Krümmung der Kurve. Demzufolge definiert man die *Krümmung* der Kurve im Punkt $P = \vec{r}(t)$ als

$$\kappa(t) := \lim_{\Delta t \rightarrow 0} \frac{\Delta\varphi(t)}{\Delta s(t)} = \frac{\dot{\varphi}(t)}{\dot{s}(t)}.$$

4.5.3. Satz. Die Krümmung einer Kurve mit regulärer, zweimal differenzierbarer Parameterdarstellung $\vec{r}: [a, b] \rightarrow \mathbb{R}^2$ beträgt im Kurvenpunkt $P(t) = \vec{r}(t)$ gerade

$$\kappa(t) = \frac{\langle \ddot{\vec{r}}(t), \vec{n}(t) \rangle}{\|\dot{\vec{r}}(t)\|^3} = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}^2(t) + \dot{y}^2(t))^{\frac{3}{2}}}.$$

4.5.4. Folgerung. Die Krümmung des Graphen $y = f(x)$ einer zweimal differenzierbaren Funktion $f: [a, b] \rightarrow \mathbb{R}$ im Punkt $(x, f(x))$ beträgt

$$\kappa(x) = \frac{f''(x)}{(\sqrt{1 + f'^2(x)})^3}.$$

Einen durch den Kurvenpunkt $P = \vec{r}(t)$ gehenden Kreis nennt man *Krümmungskreis* der Kurve in P , wenn er dieselbe Krümmung und denselben Tangentialvektor wie die Kurve besitzt. Der Radius r des Krümmungskreises heisst *Krümmungsradius* in P und ist gegeben durch:

$$r = \frac{1}{|\kappa(t)|}.$$

Bemerkung. Die Krümmung einer Kurve hängt nicht von der Parameterdarstellung ab. Sei etwa $\vec{r}: I \rightarrow \mathbb{R}^2$ eine Parameterdarstellung und sei $\vec{\gamma} = \vec{r} \circ f: J \rightarrow \mathbb{R}^2$ eine Umparametrisierung zur Parametertransformation $f: J \rightarrow I$ mit $f'(s) > 0$ für alle s . Sei $t = f(s) \in I$ für ein $s \in J$, dann hat nach der Kettenregel aus Satz 3.1.7 die Kurve $\vec{\gamma}$ zur Zeit s die gleiche Krümmung wie die Kurve \vec{r} zur Zeit t . Radius und Mittelpunkt des Krümmungskreises hängen auch nicht von der Parameterdarstellung ab; selbst dann, wenn $f'(s) < 0$.

4.5.5. Polardarstellung einer Kurve. Analog zu den komplexen Zahlen kann man für eine mit einem kartesischen Koordinatensystem versehene Ebene Polarkoordinaten einführen. Sei $P = (x, y)$ ein Punkt in der Ebene, dann sind der Abstand r des Punktes vom Ursprung und der Drehwinkel φ , der den Punkt $(r, 0)$ in (x, y) überführt, die *Polarkoordinaten* von P . Wir haben folgende Beziehungen:

$$(1) \quad x = r \cos \varphi, \quad y = r \sin \varphi,$$

$$(2) \quad r = \sqrt{x^2 + y^2}, \quad \varphi = \begin{cases} \arccos \frac{x}{r}, & \text{falls } y \geq 0, \\ 2\pi - \arccos \frac{x}{r}, & \text{falls } y < 0, \\ \text{unbestimmt}, & \text{falls } r = 0. \end{cases}$$

Wenn ein Zeiger mit Fußpunkt im Ursprung, der seine Länge ändert, sich um den Ursprung bewegt, erhalten wir eine Kurve. Die Parameterdarstellung

$$r = r(\varphi), \quad \alpha \leq \varphi \leq \beta$$

der Kurve mit Parameter φ heisst *Polardarstellung*, wobei der Winkel φ von der positiven x -Achse aus gemessen wird. Aus der Polardarstellung $r: [\alpha, \beta] \rightarrow \mathbb{R}$

erhält man folgende Parameterdarstellung

$$\vec{r}(\varphi) = \begin{pmatrix} x(\varphi) \\ y(\varphi) \end{pmatrix} = r(\varphi) \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

mit dem Polarwinkel φ als Parameter. Manchmal kann es dabei sinnvoll sein, auch negative Radien zuzulassen.

Folgerung (aus Satz 4.5.1). *Eine Kurve mit Polardarstellung $r: [\alpha, \beta] \rightarrow \mathbb{R}$ hat die Länge*

$$L = \int_{\alpha}^{\beta} \sqrt{r^2(\varphi) + r'^2(\varphi)} d\varphi.$$

Bemerkung. Manchmal ist es günstig, einen speziellen Parameter für eine Parameterdarstellung zu wählen.

- (1) Bei der Polardarstellung ist der Polarwinkel φ der Parameter:

$$\varphi \mapsto (r(\varphi), \varphi) = r(\varphi) \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}.$$

- (2) Beim Graphen einer Funktion f ist die x -Koordinate der Parameter:

$$x \mapsto \begin{pmatrix} x \\ f(x) \end{pmatrix}.$$

- (3) Eine Parameterdarstellung $\vec{r}: [a, b] \rightarrow \mathbb{R}^2$ einer Kurve heißt *Bogenlängenparametrisierung*, wenn $\|\dot{\vec{r}}(s)\| = 1$ für alle $s \in [a, b]$, denn für die Länge des Teilstücks $\vec{r}|_{[a,s]}$ gilt dann

$$L(\vec{r}|_{[a,s]}) = \int_a^s \|\dot{\vec{r}}(t)\| dt = s - a.$$

4.5.6. Flächeninhalte.

4.5.5. Satz. *Sind $f, g: [a, b] \rightarrow \mathbb{R}$ stetig, $a < b$, dann beträgt der Inhalt der von den vier Kurven $y = f(x)$, $y = g(x)$, $x = a$, $x = b$ (in kartesischen Koordinaten) berandeten Fläche*

$$F = \int_a^b |f(x) - g(x)| dx.$$

4.5.6. Satz. *Ist $r: [\alpha, \beta] \rightarrow \mathbb{R}_0^+$ stetig, $\alpha < \beta$, dann beträgt der Inhalt der von den drei Kurven $r = r(\varphi)$, $\varphi = \alpha$, $\varphi = \beta$ (in Polarkoordinaten) berandeten Sektorfläche*

$$F = \frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi$$

4.5.7. Satz. Ist $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ eine stückweise stetig differenzierbare Parameterdarstellung einer Kurve K , die von jedem Ursprungsstrahl höchstens einmal getroffen wird, dann beträgt der Inhalt der durch K begrenzten Sektorfläche

$$F = \frac{1}{2} \int_a^b |\langle \vec{r}(t), n(t) \rangle| dt = \frac{1}{2} \int_a^b |x(t) \dot{y}(t) - y(t) \dot{x}(t)| dt .$$

Definition. Eine Kurve K mit Parameterdarstellung $\vec{r}: [a, b] \rightarrow \mathbb{R}^2$ heißt *geschlossen*, wenn $\vec{r}(a) = \vec{r}(b)$ gilt. Sie heißt *einfach geschlossen* oder *überschneidungsfrei*, wenn außerdem $\vec{r}(t_0) = \vec{r}(t_1)$ nur für $t_0 = t_1$ oder $t_0, t_1 \in \{a, b\}$ gilt.

4.5.8. Satz. Sei $\vec{r}: [a, b] \rightarrow \mathbb{R}^2$ Parameterdarstellung einer einfach geschlossenen Kurve K . Dann umschließt K ein Gebiet der Fläche

$$F = \frac{1}{2} \left| \int_a^b \langle \vec{r}(t), n(t) \rangle dt \right| = \frac{1}{2} \left| \int_a^b x(t) \dot{y}(t) - y(t) \dot{x}(t) dt \right| .$$

Bemerkung. Die Formeln in den Sätzen 4.5.7 und 4.5.8 gelten unabhängig von der Parameterdarstellung der Kurve.

Da die Kurven in Satz 4.5.8 geschlossen sind, vereinfacht sich die Formel weiter zu

$$F = \left| \int_a^b x(t) \dot{y}(t) dt \right| = \left| \int_a^b y(t) \dot{x}(t) dt \right| .$$

4.5.7. Volumen und Oberfläche von Rotationskörpern. Sei R ein Rotationskörper, der durch Rotation einer Kurve K um die z -Achse entsteht. Es ist oft vorteilhaft, die Koordinaten von K als "Radius" $r(t)$ und "Höhe" $h(t)$ zu bezeichnen:

$$t \mapsto \begin{pmatrix} r(t) \\ h(t) \end{pmatrix} \in \mathbb{R}^2 .$$

Dann erhält man

$$R = \left\{ \begin{pmatrix} r(t) \cos \varphi \\ r(t) \sin \varphi \\ h(t) \end{pmatrix} \middle| t \in [a, b], \varphi \in [0, 2\pi] \right\} .$$

4.5.9. Satz. Ein durch Drehung der Kurve $\vec{r} = \begin{pmatrix} r \\ h \end{pmatrix}$ um die z -Achse erzeugter Rotationskörper hat das Volumen

$$V = \pi \left| \int_a^b r^2(t) \dot{h}(t) dt \right| .$$

4.5.10. Satz. Ein durch Drehung der Kurve $\vec{r} = \begin{pmatrix} r \\ h \end{pmatrix}$ um die z -Achse erzeugter Rotationskörper hat die Oberfläche

$$F = 2\pi \int_a^b r(t) \sqrt{\dot{r}(t)^2 + \dot{h}(t)^2} dt .$$

4.6. Numerische Integration

Die näherungsweise Berechnung eines bestimmten Integrals mit Hilfe endlich vieler Funktionswerte $y_i = f(\zeta_i)$, $i = 1, \dots, n$, durch

$$\int_a^b f(x) dx \approx \sum_{i=1}^n g_i f(\zeta_i)$$

mit *Gewichtsfaktoren* g_i nennt man *Quadratur*. Die Wahl $g_i = x_i - x_{i-1}$, d.h. man approximiert das Integral mit Riemannschen Summen, spielt in der Praxis keine Rolle, da die Folge zu langsam konvergiert.

Eine bessere Approximation liefert die *Trapez - Regel*. Sei $a = x_0 < x_1 < \dots < x_n = b$ eine äquidistante Zerlegung des Intervalls $[a, b]$ der *Schrittweite* $h := \frac{b-a}{n}$ mit Stützstellen $x_i = a + ih$, $i = 0, \dots, n$. Die Kurve $y = f(x)$ wird im Intervall $[x_i, x_{i+1}]$ durch die $(x_i, f(x_i))$ und $(x_{i+1}, f(x_{i+1}))$ verbindende Sehne ersetzt. Mit $y_i = f(x_i)$ und der Flächenformel für das Trapez ergibt sich

$$\int_a^b f(x) dx \approx h \left(\frac{y_0}{2} + y_1 + \dots + y_{n-1} + \frac{y_n}{2} \right).$$

In der Praxis ist eine Halbierung der Teilintervalle sinnvoll. Für $h_n := \frac{b-a}{2^n}$ erhält man

$$T_0 := h_0 \left(\frac{f(a)}{2} + \frac{f(b)}{2} \right),$$

$$T_n := h_n \left(\frac{f(a)}{2} + f(a + h_n) + \dots + f(a + (2^n - 1)h_n) + \frac{f(b)}{2} \right).$$

Um von T_n nach T_{n+1} zu kommen braucht man nur die neu hinzugekommenen Teilpunkte zu berücksichtigen, d.h.

$$T_{n+1} = \frac{1}{2} (T_n + M_n),$$

$$M_n = h_n \left(f\left(a + \frac{h_n}{2}\right) + f\left(a + \frac{3h_n}{2}\right) + \dots + f\left(a + (2^n - 1) \frac{h_n}{2}\right) \right).$$

Verbreitet ist auch die *Simpson-Regel* (vgl. *Keplersche Fassregel*). Man ersetzt über dem Teilintervall $[x_i, x_{i+2}]$ für $i = 0, 2, \dots, 2n - 2$, den Integranden $f(x)$ durch die Parabel durch die Punkte (x_i, y_i) , (x_{i+1}, y_{i+1}) und (x_{i+2}, y_{i+2}) . Man

erhält

$$\begin{aligned}
 S_0 &:= \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) , \\
 S_n &:= \frac{h_n}{3} \left(f(a) + f(b) + 4f(a+h_n) \right. \\
 &\quad \left. + 2 \sum_{i=1}^{2^{n-1}-1} \left(f(a+2ih_n) + 2f(a+(2i+1)h_n) \right) \right) \\
 &= \frac{T_{n-1} + 2M_{n-1}}{3} = \frac{4T_n - T_{n-1}}{4-1} ,
 \end{aligned}$$

wobei $h := \frac{b-a}{2^n}$. Eine weitere Verfeinerung dieses Verfahrens erhält man beim *Romberg-Schema*...

KAPITEL 5

Potenzreihen

Eine effektive Methode bei der Lösung vieler Probleme ist die Darstellung einer Funktion $f(x)$ als eine unendliche Reihe

$$f(x) = \sum_{k=0}^{\infty} f_k(x)$$

mit „einfachen“ Funktionen $f_k(x)$. Solche Reihen heißen im Falle $f_k(x) = a_k x^k$ Potenzreihen und im Falle $f_k(x) = a_k \cos kx + b_k \sin kx$ Fourier - Reihen. Bevor man sich mit Reihen von Funktionen beschäftigt muss man sich genauer mit Reihen reeller Zahlen befassen.

5.1. Unendliche Reihen

5.1.1. Grundbegriffe. Sei $(a_k)_{k \geq 0}$ eine Folge reeller Zahlen. Durch die Festsetzung

$$S_n := \sum_{k=0}^n a_k$$

wird eine Zahlenfolge $(S_n)_{n \geq 0}$ definiert, die die zu $(a_k)_{k \geq 0}$ gehörende *unendliche Reihe* heisst und mit $\sum_{k=0}^{\infty} a_k$ bezeichnet wird. Die Zahlen a_k heissen *Glieder* der Reihe, die Summen S_n heissen *Partialsummen* oder *Zwischensummen*. Wenn $(S_n)_{n \geq 0}$ konvergiert (bzw. divergiert) sagt man, die Reihe *konvergiert* (bzw. *divergiert*). Falls

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k = S$$

mit $S \in \mathbb{R} \cup \{\pm\infty\}$ nennt man S die *Summe* oder den *Grenzwert* der Reihe und schreibt

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

Das Symbol $\sum_{k=0}^{\infty} a_k$ ist doppeldeutig:

- $\sum_{k=0}^{\infty} a_k$ bezeichnet die Reihe als solche, d.h. die Folge $(S_n)_{n \geq 0}$,

- $S = \sum_{k=0}^{\infty} a_k$ bedeutet, dass $\lim_{n \rightarrow \infty} S_n = S$, d.h. die Folge $(S_n)_{n \geq 0}$ hat den Grenzwert S .

Diese Doppeldeutigkeit ist nützlich, da man bei der Untersuchung von Reihen am Anfang nicht weiss ob sie konvergieren. Dadurch vermeidet man Fallunterscheidungen.

5.1.1. Satz (Cauchy-Kriterium für Folgen). *Die Folge $(a_k)_{k \in \mathbb{N}_0}$ ist genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ einen Index N_ε gibt, so dass für alle $n, m \geq N_\varepsilon$ gilt:*

$$|a_n - a_m| < \varepsilon .$$

5.1.2. Satz (Cauchy-Kriterium für Reihen). *Die Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ einen Index N_ε gibt, so dass für alle $n, m \geq N_\varepsilon$*

$$|S_n - S_m| = |a_{m+1} + \cdots + a_n| < \varepsilon$$

gilt.

Folgerung. *Die Glieder einer konvergenten Reihe bilden eine Nullfolge.*

Bemerkung. Dieses Kriterium ist nicht hinreichend. Zum Beispiel divergiert die harmonische Reihe in Abschnitt 2.4.2, obwohl ihre Glieder eine Nullfolge bilden.

Definition. Eine Reihe heisst *alternierend*, wenn aufeinander folgende Reihenglieder unterschiedliche Vorzeichen haben, d.h. $a_n \cdot a_{n+1} < 0$ für alle $n \in \mathbb{N}_0$ gilt.

5.1.3. Satz (Leibnitz-Kriterium). *Für jede monoton fallende Nullfolge $(a_n)_{n \geq 0}$ konvergiert die alternierende Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$.*

5.1.4. Satz. *Sei $c \in \mathbb{R}$ und seien $\sum_{k=0}^{\infty} a_k = a$, $\sum_{k=0}^{\infty} b_k = b$ konvergente Reihen. Dann gilt*

$$\sum_{k=0}^{\infty} (a_k \pm b_k) = a \pm b; \quad \sum_{k=0}^{\infty} (ca'_k) = ca' .$$

Bemerkung. Elementare Operationen wie Klammern Setzen oder Weglassen und Umordnen der Glieder, die für endliche Reihen erlaubt sind, sind im Allgemeinen für unendliche Reihen nicht zulässig. Einige Ausnahmen besprechen wir im Folgenden.

5.1.5. Satz. *In einer konvergenten Reihe darf man beliebig Klammern setzen, d.h.*

$$S = \sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots = (a_1 + \cdots + a_{k_1}) + (a_{k_1+1} + \cdots + a_{k_2}) + \cdots .$$

5.1.2. Absolute Konvergenz. Die Reihe $\sum_{k=0}^{\infty} a_k$ heisst *absolut konvergent*, wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert. Reihen die konvergieren, aber nicht absolut konvergieren heissen *bedingt konvergent*.

5.1.6. Satz. *Jede absolut konvergente Reihe ist konvergent.*

5.1.7. Satz. *Die Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann absolut konvergent, wenn die Folge der Partialsummen $S_n = \sum_{k=0}^n |a_k|$ beschränkt ist.*

Beispiel. Die Reihe

$$\sum_{k=1}^{\infty} k^{\alpha} \quad \begin{cases} \text{konvergiert} & \text{falls } \alpha < -1, \\ \text{divergiert gegen } \infty & \text{falls } \alpha \geq -1. \end{cases}$$

5.1.8. Satz. *Ist die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent mit der Summe S , dann konvergiert jede Umordnung dieser Reihe ebenfalls gegen S .*

Genauer: Es sei $\sigma : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ eine bijektive Abbildung, dann konvergiert die Reihe $\sum_{k=0}^{\infty} a_{\sigma(k)}$ genau dann, wenn $\sum_{k=0}^{\infty} a_k$ konvergiert, und beide Reihen haben den gleichen Grenzwert.

5.1.9. Satz (Cauchy-Produktformel). *Es seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergent, dann gilt*

$$\left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right).$$

5.1.10. Satz (Vergleichskriterium). *Besteht für die Reihenglieder die Abschätzung*

$$0 \leq |a_k| \leq b_k \quad \text{für } k \geq k_0$$

dann gilt:

$$\begin{aligned} \text{a) } \sum_{k=0}^{\infty} b_k \text{ konvergent} & \Rightarrow \sum_{k=0}^{\infty} a_k \text{ absolut konvergent,} \\ \text{b) } \sum_{k=0}^{\infty} |a_k| = \infty & \Rightarrow \sum_{k=0}^{\infty} b_k = \infty. \end{aligned}$$

5.1.11. Satz (Quotientenkriterium). *Sei $a_k \neq 0$ für alle $k \geq k_0$ und sei $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = a$. Dann gilt:*

$$\begin{aligned} (1) \quad a < 1 & \Rightarrow \sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent,} \\ (2) \quad a > 1 & \Rightarrow \sum_{k=0}^{\infty} a_k \text{ ist divergent.} \end{aligned}$$

5.1.12. Satz (Wurzelkriterium). *Es sei $\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = a$. Dann gilt*

$$(1) \ a < 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent,}$$

$$(2) \ a > 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist divergent.}$$

Bemerkung. Die Sätze 5.1.11 und 5.1.12 machen im Fall $a = 1$ keine Aussage. Betrachte etwa die Reihe $\sum_{k=1}^{\infty} k^c$, die für $c < -1$ konvergiert und für $c \geq -1$ divergiert. Für alle c erhält man sowohl im Quotienten- als auch im Wurzelkriterium den Grenzwert $a = 1$.

5.2. Reihen von Funktionen

Sei $(f_n)_{n \geq 0}$ eine auf dem Intervall $I \subseteq \mathbb{R}$ definierte Funktionenfolge. Falls für alle $x \in I$ die Zahlenfolge $(f_n(x))_{n \geq 0}$ den Grenzwert $f(x)$ besitzt, sagt man, dass die Funktionenfolge $(f_n)_{n \geq 0}$ *punktweise gegen f konvergiert*.

Bemerkung. Die Konvergenzgeschwindigkeit kann vom Punkt $x \in I$ abhängen, und damit können manche Eigenschaften der Funktionen f_n für die Grenzfunktion verloren gehen.

5.2.1. Gleichmäßige Konvergenz. Eine Funktionenfolge $(f_n)_{n \geq 0}$ *konvergiert gleichmäßig* auf I gegen die Funktion $f : I \rightarrow \mathbb{R}$, wenn es für alle $\varepsilon > 0$ einen für alle $x \in I$ gemeinsamen Index N_ε gibt, so dass gilt:

$$n \geq N_\varepsilon \quad \Rightarrow \quad \forall x \in I : |f(x) - f_n(x)| < \varepsilon.$$

Anders ausgedrückt konvergiert (f_n) genau dann gleichmäßig gegen f auf I , wenn

$$\sup_{x \in I} \|f_n(x) - f(x)\|$$

für $n \rightarrow \infty$ gegen 0 konvergiert.

5.2.1. Satz. *Sind alle Funktionen f_n , $n \geq 0$, auf dem Intervall I stetig und konvergiert die Folge $(f_n)_{n \geq 0}$ auf I gleichmäßig gegen f , dann ist auch die Grenzfunktion f stetig.*

Beispiel. Die Folge $f_n(x) = x^n$, $x \in [0, 1]$ konvergiert punktweise (aber nicht gleichmäßig) gegen die Funktion

$$f(x) = \begin{cases} 0, & \text{falls } x \in [0, 1), \\ 1, & \text{falls } x = 1. \end{cases}$$

Die Folgenglieder f_n sind stetig, aber die Grenzfunktion nicht.

5.2.2. Satz. *Konvergiert die Folge stetiger Funktionen f_n , $n \geq 0$, auf dem Intervall I gleichmäßig gegen $f : I \rightarrow \mathbb{R}$, dann gilt für alle $a, b \in I$*

$$\int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Beispiel. Sei

$$h_n(x) = \begin{cases} n^2 x, & \text{falls } x \in [0, \frac{1}{n}], \\ 2n - n^2 x, & \text{falls } x \in [\frac{1}{n}, \frac{2}{n}], \\ 0, & \text{falls } x \in [\frac{2}{n}, 1]. \end{cases}$$

Dann konvergiert $(h_n)_{n \geq 0}$ punktweise (aber nicht gleichmäßig) gegen 0. Aber es gilt

$$1 = \lim_{n \rightarrow \infty} \int_{n=0}^1 h_n(x) dx \neq \int_{n=0}^1 \lim_{n \rightarrow \infty} h_n(x) dx = \int_{n=0}^1 0 dx = 0.$$

5.2.3. Satz. *Sind alle Funktionen f_n , $n \geq 0$, auf I stetig differenzierbar, konvergiert $(f_n)_{n \geq 0}$ punktweise gegen f und konvergiert $(f'_n)_{n \geq 0}$ auf I gleichmäßig, dann ist auch die Grenzfunktion f differenzierbar und es gilt*

$$f'(x) = \left(\lim_{n \rightarrow \infty} f_n(x) \right)' = \lim_{n \rightarrow \infty} f'_n(x).$$

Beispiel. Die Folge $g_n(x) = \frac{2x}{\pi} \arctan(nx)$, $x \in [-1, 1]$ besteht aus differenzierbaren Funktionen die punktweise gegen die Funktion $x \mapsto |x|$ konvergieren. Die Grenzfunktion ist in $x = 0$ nicht differenzierbar.

5.2.2. Gleichmäßig konvergente Reihen. Alle obigen Überlegungen lassen sich leicht auf Reihen übertragen.

Sei eine Funktion f auf I als unendliche Reihe dargestellt, d.h. für alle $x \in I$ gilt $f(x) = \sum_{k=0}^{\infty} f_k(x)$. Man sagt die Reihe $\sum_{k=0}^{\infty} f_k(x)$ konvergiert *punktweise* (bzw. *gleichmäßig*) gegen $f(x)$, wenn die Folge der Partialsummen $S_n(x) := \sum_{k=0}^n f_k(x)$ punktweise (bzw. gleichmäßig) gegen f konvergiert.

5.2.4. Satz. *Konvergiert die Reihe $\sum_{k=0}^{\infty} f_k(x)$ stetiger Funktionen f_k auf I gleichmäßig gegen f , dann ist*

$$f(x) = \sum_{k=0}^{\infty} f_k(x)$$

stetig und für alle $a, b \in I$ gilt:

$$\int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx = \int_a^b f(x) dx = \sum_{k=0}^{\infty} \left(\int_a^b f_k(x) dx \right).$$

5.2.5. Satz. Sind alle Funktionen f_k , $k \geq 0$, auf I stetig differenzierbar, konvergiert die Reihe $\sum_{k=0}^{\infty} f_k(x)$ auf I punktweise gegen $f(x)$ und konvergiert die Reihe $\sum_{k=0}^{\infty} f'_k(x)$ auf I gleichmäßig, dann gilt:

$$f'(x) = \left(\sum_{k=0}^{\infty} f_k(x) \right)' = \sum_{k=0}^{\infty} f'_k(x).$$

5.2.6. Satz (“M-Test”). Gilt für jede Funktion f_k der auf dem Intervall $I \subseteq \mathbb{R}$ definierten Funktionenfolge $(f_k)_{k \geq 0}$ eine Abschätzung

$$|f_k(x)| \leq M_k$$

und gilt für die Zahlenreihe $(M_k)_{k \geq 0}$

$$\sum_{k=0}^{\infty} M_k < \infty,$$

dann ist die Funktionenreihe $\sum_{k=0}^{\infty} f_k(x)$ auf I gleichmäßig und absolut konvergent.

5.3. Potenzreihen

Eine Funktionenreihe $\sum_{k=0}^{\infty} f_k(x)$ mit den speziellen Funktionen $f_k(x) = a_k x^k$ heißt **Potenzreihe**. Also übertragen sich alle Ergebnisse aus dem Paragraphen 5.2 auf Potenzreihen $\sum_{k=0}^{\infty} a_k x^k$.

5.3.1. Der Konvergenzradius. Sei $\sum_{k=0}^{\infty} a_k x^k$, $x \in \mathbb{R}$, eine Potenzreihe und sei

$$M := \left\{ x \in \mathbb{R} \mid \sum_{k=0}^{\infty} a_k x^k \text{ konvergiert} \right\}.$$

Der *Konvergenzradius* R der Potenzreihe ist definiert durch

$$R := \begin{cases} \sup\{|x|; x \in M\} & \text{falls } M \text{ beschränkt,} \\ \infty & \text{falls } M \text{ unbeschränkt.} \end{cases}$$

Für R gibt es drei Möglichkeiten:

$$R = 0, \quad 0 < R < \infty, \quad R = \infty.$$

5.3.1. Satz. Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius R . Dann gilt:

- (1) $R = 0 \iff$ Reihe konvergiert nur für $x = 0$.
- (2) Ist $R > 0$ und $\varrho \in (0, R)$, dann konvergieren die Reihen $\sum_{k=0}^{\infty} a_k x^k$ und $\sum_{k=1}^{\infty} k a_k x^{k-1}$ absolut und gleichmäßig auf $[-\varrho, \varrho]$.
- (3) Für alle x mit $|x| > R$ ist die Reihe $\sum_{k=0}^{\infty} a_k x^k$ divergent.

Der Satz sagt nichts über $x = \pm R$ aus; diese Punkte müssen für jede Reihe neu untersucht werden.

5.3.2. Berechnung des Konvergenzradius. Der Konvergenzradius kann immer durch folgende Formel berechnet werden:

$$R = \sup \{ r \geq 0; \text{ die Folge } (|a_n| r^n)_{n \geq 0} \text{ ist beschränkt} \},$$

wobei $R = \infty$ falls die Menge unbeschränkt ist. Oft ist es einfacher, folgende Formeln zu benutzen:

(1) Sei $a_k \neq 0$ für $k \geq k_0$, und sei

$$a = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

Dann ist der Konvergenzradius $R = \frac{1}{a} \in [0, \infty]$.

(2) Sei

$$a = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|},$$

dann ist $R = \frac{1}{a} \in [0, \infty]$.

Das geht natürlich nur dann, wenn der jeweilige Grenzwert in $[0, \infty]$ existiert.

Definition. Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$ und Summe $f(x)$ für $|x| < R$. Man sagt, f wird auf $(-R, R)$ durch die Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ dargestellt.

5.3.3. Differentiation und Integration von Potenzreihen.

5.3.2. Satz. Eine durch eine Potenzreihe dargestellte Funktion f ist im offenen Konvergenzintervall $(-R, R)$, $R > 0$, beliebig oft differenzierbar. Die Ableitung erhält man durch gliedweise Differentiation:

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1},$$

$$f''(x) = \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2},$$

etc. Die abgeleiteten Reihen haben alle den Konvergenzradius R .

Beispiel.

$$f(x) = \frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad |x| < 1,$$

$$f'(x) = \frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1} \quad |x| < 1,$$

$$f''(x) = \frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k(k-1)x^{k-2} \quad |x| < 1.$$

5.3.3. Satz. Für alle a, b aus dem offenen Konvergenzintervall $(-R, R)$ der Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k x^k$ gilt:

$$\int_a^b f(x) dx = \sum_{k=0}^{\infty} \int_a^b a_k x^k dx = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (b^{k+1} - a^{k+1}).$$

Insbesondere ist

$$F(x) := \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$$

eine Stammfunktion von f auf $(-R, R)$ mit Konvergenzradius R .

5.3.4. Potenzreihendarstellungen spezieller Funktionen.

5.3.4. Satz. Wir haben folgende Potenzreihendarstellungen:

$$\begin{aligned} a) \quad e^x &= \sum_{k=0}^{\infty} \frac{1}{k!} x^k, & x \in \mathbb{R}, \\ b) \quad \sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}, & x \in \mathbb{R}, \\ c) \quad \cos x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, & x \in \mathbb{R}, \\ d) \quad \ln(1+x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}, & |x| < 1, \\ e) \quad \arctan x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}, & |x| < 1. \end{aligned}$$

5.3.5. Satz (Binomialreihe). Für alle x mit $|x| < 1$ und alle $\alpha \in \mathbb{R}$ gilt:

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,$$

wobei

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

Bemerkung. (1) Für $\alpha = n \in \mathbb{N}$ gilt

$$\begin{aligned} \binom{\alpha}{k} &= \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = 0 \quad \text{falls } n < k \\ \implies (1+x)^n &= \sum_{k=0}^n \binom{n}{k} x^k \quad (\text{klassische binomische Formel}) \end{aligned}$$

(2) Für $\alpha = -\frac{1}{2}$ ist

$$\begin{aligned} \binom{-\frac{1}{2}}{k} &= \frac{\left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdots \left(-\frac{2k-1}{2}\right)}{k!} \\ &= (-1)^k \frac{1 \cdot 3 \cdots (2k-1)}{2^k k!} = (-1)^k \frac{(2k)!}{2^{2k} (k!)^2}, \end{aligned}$$

also

$$\frac{1}{\sqrt{1+x}} = \sum_{k=0}^{\infty} (-1)^k \frac{(2k)!}{2^{2k} (k!)^2} x^k.$$

Einsetzen von $\mp x^2$ und Integration liefert Potenzreihendarstellungen für \arcsin und arsinh , siehe Tabelle am Ende von Abschnitt 4.1.4.

5.3.5. Potenzreihen mit Zentrum x_0 . Eine unendliche Reihe der Form

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

heißt *Potenzreihe mit Zentrum x_0* (oder kurz *Potenzreihe um x_0*).

Für theoretische Überlegungen reicht es $x_0 = 0$ zu betrachten, denn durch Substitution $z = x - x_0$ geht die Reihe $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ in die Reihe $\sum_{k=0}^{\infty} a_k z^k$ mit Zentrum in 0 über.

5.3.6. Lemma. Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius R , und sei $x_0 \in \mathbb{R}$. Dann gilt:

$$\begin{aligned} x \in (x_0 - R, x_0 + R) &\implies \sum_{k=0}^{\infty} a_k (x - x_0)^k \text{ konvergiert,} \\ x \notin [x_0 - R, x_0 + R] &\implies \sum_{k=0}^{\infty} a_k (x - x_0)^k \text{ divergiert.} \end{aligned}$$

5.3.6. Koeffizientenvergleich. Sei f auf Intervall $(x_0 - R, x_0 + R)$ als Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ um x_0 dargestellt. Nach Satz 5.3.2 gilt für die n -te Ableitung

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k \cdot (k-1) \cdots (k-n+1) a_k (x - x_0)^{k-n}.$$

Setzt man $x = x_0$, so erhält man $f^{(n)}(x_0) = n! a_n$.

5.3.7. Satz (Eindeutigkeit von Potenzreihen). Sei $R > 0$ und gelte für alle $x \in (x_0 - R, x_0 + R)$:

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k = \sum_{k=0}^{\infty} b_k (x - x_0)^k,$$

dann folgt

$$a_k = b_k = \frac{f^{(k)}(x_0)}{k!} \quad \text{für } k = 0, 1, \dots$$

5.4. Taylorreihen

5.4.1. Die Taylorformel. Wir benutzen obige Formel, um differenzierbare Funktionen durch Potenzreihen zu approximieren.

5.4.1. Satz (Taylorentwicklung mit Restglied). Für jede auf dem offenen Intervall $I \subseteq \mathbb{R}$ $(n+1)$ -mal stetig differenzierbare Funktion f und $x, x_0 \in I$ gilt:

$$f(x) = \underbrace{\sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k}_{=: T_n(x, x_0)} + R_{n+1}(x, x_0),$$

wobei $T_n(x, x_0)$ das Taylor-Polynom heißt, und das Restglied $R_{n+1}(x, x_0)$ die Darstellungen

$$\begin{aligned} R_{n+1}(x, x_0) &= \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt && \text{(Cauchy), bzw.} \\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} && \text{für ein } \xi \in (x, x_0) \text{ (Lagrange) hat.} \end{aligned}$$

5.4.2. Satz (Extremwert-Test). Ist die Funktion f auf dem Intervall I n -mal stetig differenzierbar und $x_0 \in I$ mit

$$f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0 \quad \text{und} \quad f^{(n)}(x_0) \neq 0,$$

dann gilt:

- (1) x_0 ist lokale Extremstelle $\iff n$ gerade,
 (2) x_0 ist lokales Minimum $\iff n$ gerade und $f^{(n)}(x_0) > 0$,
 (3) x_0 ist lokales Maximum $\iff n$ gerade und $f^{(n)}(x_0) < 0$.

Definition. Sei f auf dem offenen Intervall I beliebig oft differenzierbar und sei $x_0 \in I$. Die unendliche Reihe

$$T_f(x, x_0) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

heißt *Taylor-Reihe* mit Zentrum x_0 . Gilt für alle $x \in (x_0 - R, x_0 + R)$ die Gleichung $f(x) = T_f(x, x_0)$, so sagt man, dass sich f um x_0 als *Taylor-Reihe entwickeln* lässt.

Bemerkung. Man darf nicht erwarten, dass $R_n(x, x_0) \rightarrow 0$ für $n \rightarrow \infty$ für jede beliebig oft differenzierbare Funktion f gilt. Zum Beispiel ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \text{ und} \\ 0 & \text{für } x \leq 0 \end{cases}$$

beliebig oft stetig differenzierbar mit

$$f^{(n)}(0) = \lim_{x \rightarrow 0^+} f^{(n)}(x) = \lim_{x \rightarrow 0^+} P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} = 0,$$

wobei P_n ein Polynom vom Grad $2n$ ist. Somit hat f um 0 die Taylorentwicklung

$$f(x) = \sum_{k=0}^n \underbrace{\frac{f^{(k)}(0)}{k!}}_{=0} x^k + R_{n+1}(x, 0) = R_{n+1}(x, 0).$$

Für alle n gilt also $R_{n+1}(x, 0) = f(x)$.

5.4.3. Lemma. Sei f auf dem Intervall I beliebig oft differenzierbar und sei $x_0 \in I$ und $R > 0$. Dann konvergiert die Taylor-Reihe $T_f(x, x_0)$ genau für diejenigen $x \in I \cap (x_0 - R, x_0 + R)$ gegen $f(x)$, für die das Restglied $R_n(x, x_0)$ mit $n \rightarrow \infty$ gegen 0 strebt. Eine hinreichende Bedingung dafür ist, dass es eine Konstante A gibt mit

$$|f^{(n)}(x)| \leq \frac{n! A}{R^n} \quad \forall x \in I, \forall n \in \mathbb{N}.$$

5.4.4. Satz. Sei f auf dem Intervall I beliebig oft differenzierbar und sei $x_0 \in I$. Dann lässt sich f auf I als die Taylorreihe um x_0 entwickeln, wenn es Konstanten A, B gibt mit

$$|f^{(n)}(x)| \leq A B^n \quad \forall x \in I, \forall n \in \mathbb{N}.$$

5.4.2. Methoden der Reihenentwicklung. Es gibt verschiedene Möglichkeiten, die Taylorentwicklung einer Funktion zu bestimmen.

- (1) Die *Taylor-Formel* aus Satz 5.4.1 zusammen mit dem Nachweis, dass $R_n(x, a) \rightarrow 0$ (z.B. mit Lemma 5.4.3 oder Satz 5.4.4), liefert, dass die Taylor-Reihe gegen $f(x)$ konvergiert. Da man dazu alle Ableitungen von f bestimmen muss, funktioniert das nur selten.

Beispiele: e^x , \sin , \cos .

- (2) Bekannte *Reihen differenzieren oder integrieren*. Die Sätze 5.3.2 und 5.3.3 ermöglichen es, durch gliedweise Differentiation und Integration bekannter Reihen neue Reihenentwicklung zu erhalten (siehe Satz 5.3.4).

Beispiele: \log , \arctan , \arcsin , arsinh etc.

- (3) *Summe und Produkt* bekannter Reihen liefern neue Reihenentwicklungen.

Beispiel: $\cosh x = \frac{e^x + e^{-x}}{2}$, $\sinh x = \frac{e^x - e^{-x}}{2}$.

Die Cauchy-Produktformel aus Satz 5.1.9 liefert für Potenzreihen, dass

$$\left(\sum_{n=0}^{\infty} a_n x^n \right) \cdot \left(\sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^n a_i b_{n-i} \right) x^n .$$

Seien c_n die Koeffizienten der Produktreihe. Auflösen nach a_n liefert auch eine induktive Formel für Quotienten (falls $b_0 \neq 0$):

$$\left(\sum_{n=0}^{\infty} c_n x^n \right) / \left(\sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} a_n x^n \quad \text{mit} \quad a_n = \frac{1}{b_0} \left(c_n - \sum_{i=0}^{n-1} a_i b_{n-i} \right) .$$

- (4) Andere Verfahren: Potenzreihenansatz zur Lösung von Differentialgleichungen; Potenzreihen ineinander einsetzen (mit dem Cauchyprodukt die Potenzen der Potenzreihe von g bestimmen und in die Potenzreihe von f einsetzen), etc.

Für weitere Details zur Taylorentwicklung wird aus Zeitgründen auf die Literatur (z.B. Meyberg-Vachenauer) verwiesen.

KAPITEL 6

Lineare Algebra

Bei der Lösung verschiedenster Probleme wird man letztendlich mit der Lösung von Gleichungssystemen konfrontiert. Die damit eng verbundenen Problematiken werden in diesem Kapitel behandelt.

6.1. Lineare Gleichungssysteme und Matrizen

6.1.1. Matrizen. Ein Rechteckiges Zahlenschema der Form

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ heisst $m \times n$ *Matrix*. Die Zahlen a_{ij} heissen *Einträge* (*Elemente*) der Matrix \mathbf{A} . Man schreibt abkürzend

$$\mathbf{A} = (a_{ij})$$

Insbesondere heissen $m \times 1$ -Matrizen *Spaltenvektoren* und $1 \times n$ -Matrizen *Zeilenvektoren*. Sie haben die Form:

$$\mathbf{s} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}, \quad \mathbf{z} = (a_1, \dots, a_n).$$

Eine Matrix $\mathbf{A} = (a_{ij})$ vom Typ $m \times n$ besteht aus m Zeilenvektoren und n Spaltenvektoren. Seien

$$\begin{aligned} \mathbf{z}_i &:= (a_{i1}, \dots, a_{in}), & i &= 1, \dots, m, \\ \mathbf{s}_j &:= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, & j &= 1, \dots, n, \end{aligned}$$

dann schreibt man die Matrix \mathbf{A} in *Zeilen-* bzw. *Spaltendarstellung*

$$\mathbf{A} = \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_m \end{pmatrix}, \quad \mathbf{A} = (\mathbf{s}_1, \dots, \mathbf{s}_n).$$

Zwei Matrizen $\mathbf{A} = (a_{ij})$ und $\mathbf{B} = (b_{ij})$ sind genau dann gleich, in Zeichen $\mathbf{A} = \mathbf{B}$, wenn \mathbf{A} und \mathbf{B} vom Typ $m \times n$ sind und $a_{ij} = b_{ij}$ für alle $i = 1, \dots, m$, $j = 1, \dots, n$ gilt, d.h. wenn alle Einträge gleich sind.

Die Menge aller $m \times n$ Matrizen mit Elementen aus \mathbb{R} bezeichnen wir mit $\mathbb{R}^{m \times n}$. Insbesondere schreiben wir $\mathbb{R}^n := \mathbb{R}^{n \times 1}$ (Spaltenvektoren) und $\mathbb{R}_m := \mathbb{R}^{1 \times m}$ (Zeilenvektoren).

6.1.2. Addition, Subtraktion und Multiplikation mit einer Zahl. Für Matrizen $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$ aus $\mathbb{R}^{m \times n}$ und $\lambda \in \mathbb{R}$ ist die *Summe* $\mathbf{A} + \mathbf{B}$ und das *skalare Vielfache* $\lambda \mathbf{A}$ komponentenweise definiert, d.h.

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= (a_{ij} + b_{ij}), \\ \lambda \mathbf{A} &= (\lambda a_{ij}).\end{aligned}$$

Für Spalten- und Zeilenvektoren definiert man Summen und skalare Vielfache analog, z.B.

$$\begin{aligned}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} &= \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \\ \lambda \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &= \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}.\end{aligned}$$

Konvention: Falls ein Spaltenvektor in einer Zeile geschrieben steht, dann schreiben wir

$$(a_1, \dots, a_n)^T := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Wir haben

$$\begin{aligned}\mathbf{a} &= \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ a_n \end{pmatrix} \\ &= a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + a_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},\end{aligned}$$

und somit lässt sich jeder Spaltenvektor $\mathbf{a} \in \mathbb{R}^n$ eindeutig als Summe

$$\mathbf{a} = a_1 \mathbf{e}_1 + \cdots + a_n \mathbf{e}_n$$

mit den *Standardbasis-* (spalten-) *vektoren*

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

darstellen.

Definition. Das System $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ heisst *Standardbasis* des \mathbb{R}^n .

Analog lässt sich jeder Vektor $\mathbf{b} \in \mathbb{R}_n$ als Summe der *Standardzeilenbasisvektoren* $\mathbf{e}'_i, i = 1, \dots, n$, darstellen, wobei

$$\mathbf{e}'_1 := (1, 0, \dots, 0), \dots, \mathbf{e}'_n := (0, \dots, 0, 1).$$

Das System $(\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ heisst *Standardbasis* des \mathbb{R}_n .

Auf Grund der Darstellung von Matrizen mit Hilfe ihrer Spalten- bzw. Zeilenvektoren und der Definition der Summe und des skalaren Vielfachen von Spalten- bzw. Zeilenvektoren, kann man Summe und Skalarmultiplikation auch spaltenweise ausführen, d.h. sei $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$, $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n) \in \mathbb{R}^{m \times n}$, dann gilt:

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= (\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n), \\ \lambda \mathbf{A} &= (\lambda \mathbf{a}_1, \dots, \lambda \mathbf{a}_n). \end{aligned}$$

Für jede Matrix $\mathbf{A} = (a_{ij})$ bezeichnet man die Matrix $(-1) \mathbf{A} = (-a_{ij})$ mit $-\mathbf{A}$. Die *Differenz* zweier Matrizen ist durch

$$\mathbf{A} - \mathbf{B} =: \mathbf{A} + (-\mathbf{B})$$

definiert. Die Matrix deren sämtlichen Elemente 0 Null sind heisst *Nullmatrix*, in Zeichen $\underline{0}$. Die Nullmatrix $\underline{0} \in \mathbb{R}^n$ bzw. $\underline{0} \in \mathbb{R}_n$ heisst *Nullvektor*.

Die Rechenregeln reeller Zahlen übertragen sich auf Grund der komponentenweisen Definition der Addition und der Skalaren Multiplikation auf Matrizen. Es

gilt für alle $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m \times n}$, $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ (\mathbf{A} + \mathbf{B}) + \mathbf{C} &= \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\ \mathbf{A} + \mathbf{0} &= \mathbf{A} \\ \mathbf{A} + (-\mathbf{A}) &= \mathbf{0} \\ (\lambda\mu)\mathbf{A} &= \lambda(\mu\mathbf{A}) \\ 1\mathbf{A} &= \mathbf{A} \\ (\lambda + \mu)\mathbf{A} &= \lambda\mathbf{A} + \mu\mathbf{A} \\ \lambda(\mathbf{A} + \mathbf{B}) &= \lambda\mathbf{A} + \lambda\mathbf{B} \end{aligned}$$

6.1.3. Lineare Gleichungssysteme. Matrizen treten in natürlicher Weise bei Gleichungssystemen auf.

Definition. Ein *lineares Gleichungssystem* mit m Gleichungen in n Unbekannten x_1, \dots, x_n hat die Form

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m, \end{aligned}$$

wobei a_{ij} die *Koeffizienten* und b_i die *Absolutglieder* sind. Falls $b_i = 0$ für alle $i = 1, \dots, m$, heisst das System *homogen*, ansonsten *inhomogen*. Ein Spaltenvektor $\mathbf{c} = (c_i)$ heisst *Lösung* des linearen Gleichungssystems, wenn für $x_1 = c_1, \dots, x_n = c_n$ alle Gleichungen erfüllt sind, und die Menge aller solcher Lösungen heisst die *Lösungsmenge* des Systems.

Obiges Gleichungssystem lässt sich auch als

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, m$$

schreiben, oder noch kürzer als $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, siehe Abschnitt 6.2 unten.

Bemerkung. Jedes homogene Gleichungssystem besitzt mindestens die *triviale Lösung* $c_1 = \dots = c_n = 0$.

Definition. Sei $\sum_{j=1}^n a_{ij}x_j = b_i$, $i = 1, \dots, m$ ein lineares Gleichungssystem. Die Matrix

$$\mathbf{A} := (a_{ij}) \in \mathbb{R}^{m \times n}$$

heisst *Koeffizientenmatrix*, der Spaltenvektor

$$\mathbf{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

heisst *rechte Seite* und die $m \times (n+1)$ -Matrix bestehend aus den Spaltenvektoren \mathbf{a}_i der Koeffizientenmatrix \mathbf{A} und der rechten Seite \mathbf{b}

$$(\mathbf{A}|\mathbf{b}) := (\mathbf{a}_1, \dots, \mathbf{a}_n|\mathbf{b}) \in \mathbb{R}^{m \times (n+1)}$$

heisst *erweiterte Koeffizientenmatrix*.

Der senkrechte Strich in der Notation ist nützlich, um die Koeffizienten von der rechten Seite zu trennen.

6.1.1. Satz (Zeilenumformungen). *Folgende Umformungen eines Gleichungssystems verändern die Lösungen nicht:*

- a) Vertauschen zweier Gleichungen,
- b) Multiplikation einer Gleichung mit einer Zahl $\alpha \neq 0$,
- c) Addition eines Vielfachen einer Gleichung zu einer anderen.

Die Umformungen aus Satz 6.1.1 haben Entsprechungen für die erweiterte Koeffizientenmatrix.

Definition. Sei $(\mathbf{A}|\mathbf{b})$ eine erweiterte Koeffizientenmatrix eines linearen Gleichungssystems. Die folgenden Umformungen der Matrix $(\mathbf{A}|\mathbf{b})$ heissen *elementare Zeilenumformungen*:

- a) Vertauschen zweier Zeilen,
- b) Multiplikation einer Zeile mit $\alpha \neq 0$,
- c) Addition des α -fachen einer Zeile zu einer anderen.

6.1.2. Folgerung. *Entsteht $(\mathbf{B}|\mathbf{c})$ aus $(\mathbf{A}|\mathbf{b})$ durch endlich viele elementare Zeilenumformungen, so haben die zugehörigen Gleichungssysteme $\sum_{j=1}^n b_{ij}x_j = c_i$ und $\sum_{j=1}^n a_{ij}x_j = b_i$ dieselben Lösungen.*

6.1.4. Das Gauß'sche Eliminationsverfahren. Ziel des Gauß-Verfahrens ist es, ein Gleichungssystem durch elementare Zeilenumformungen auf die sogenannte "strenge Zeilenstufenform" zu bringen. Dann lässt sich die Lösungsmenge sehr einfach angeben.

Definition. Eine $m \times n$ -Matrix $\mathbf{A} = (a_{ij})$ ist *in strenger Zeilenstufenform*, wenn es $r \leq \min(m, n)$ und Indizes $1 \leq j_1 < \dots < j_r \leq n$ gibt, so dass gilt:

- (1) alle Zeilen unterhalb der r -ten verschwinden, d.h., $a_{ij} = 0$ für alle $i > r$ und alle j ;
- (2) sei $i \leq r$, dann beginnt die i -te Zeile an der j_i -ten Stelle mit einer 1, d.h., $a_{ij} = 0$ für $j < j_i$ und $a_{ij_i} = 1$;
- (3) für alle $i \leq r$ ist die j_i -te Spalte der Einheitsvektor \mathbf{e}_i , insbesondere $a_{kj_i} = 0$ für alle $k \neq i$.

Die Zahl r heißt der *Rang* von A . Eine erweiterte Koeffizientenmatrix ist in strenger Zeilenstufenform, wenn der Teil links vom Strich in strenger Zeilenstufenform ist.

Beispiel. (1) Die einfachsten Matrizen in strenger Zeilenstufenform sind die $n \times n$ -Einheitsmatrizen \mathbf{E}_n mit $r = n$ und $j_1 = 1, \dots, j_n = n$:

$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

(2) Es sei $m = 4, n = 8$. Die folgende erweiterte Koeffizientenmatrix ist in strenger Zeilenstufenform mit $r = 3, j_1 = 3, j_2 = 5, j_3 = 6$ (Sternchen bedeuten beliebige Einträge):

$$\left(\begin{array}{cccccccc|c} 0 & 0 & 1 & * & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{array} \right).$$

Sei A in strenger Zeilenstufenform vom Rang r . Dann hat das Gleichungssystem $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ die folgende Lösungsmenge.

- (1) Falls $b_i \neq 0$ für ein $i > r$, so ist die Lösungsmenge leer, denn die i -te Gleichung $0 = b_i$ hat dann keine Lösung.
- (2) Ansonsten steht in den Zeilen $r + 1$ bis n die Gleichung $0 = 0$, die immer erfüllt ist. Wir lösen die erste Gleichung nach x_{j_1} , die zweite nach x_{j_2} usw. und die r -te nach x_{j_r} auf. Auf der rechten Seite bleibt ein linearer Ausdruck in den verbleibenden Variablen stehen. Wir erhalten zu jeder Wahl der $n - r$ Einträge $c_1, \dots, c_{j_1-1}, c_{j_1+1}, \dots, c_{j_2-1}, \dots, c_{j_r+1}, \dots, c_n$ genau eine Lösung des Gleichungssystems. Die Lösungsmenge lautet also

$$L = \left\{ \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n \left| \begin{array}{l} c_{i_1} = b_1 - a_{1,i_1+1} c_{i_1+1} - \dots - a_{1n} c_n, \\ \vdots \\ c_{i_r} = b_r - a_{r,i_r+1} c_{i_r+1} - \dots - a_{rn} c_n. \end{array} \right. \right\}.$$

Das *Gauß'sche Eliminationsverfahren* besteht aus folgenden Schritten, die an der erweiterten Koeffizientenmatrix $(\mathbf{A}|\mathbf{b})$ eines linearen Gleichungssystems ausgeführt werden.

- (0) Beginne mit $i = j = 1$.
- (1) (*Suche nächste interessante Spalte*) Falls $a_{ij} = \dots = a_{mj} = 0$, erhöhe j um 1 und mache weiter mit (1), es sei denn $j > n$, dann fertig (Am Ende dieses Schrittes ist $a_{kj} \neq 0$ für ein $i \leq k \leq m$). Setze $r = i$ und $j_i = j$.

- (2) (*Vertausche zwei Zeilen wenn nötig*) Falls $a_{ij} = 0$, aber $a_{kj} \neq 0$ für ein $k > i$, vertausche die i -te und die k -te Zeile (Danach ist $a_{ij} \neq 0$).
- (3) (*Normieren*) Dividiere die gesamte i -te Zeile durch $a_{ij} \neq 0$ (Danach ist $a_{ij} = 1$).
- (4) (*Ausräumen der j -ten Spalte*). Für alle $k \neq i$ subtrahiere das a_{kj} -fache der i -ten Zeile von der k -ten Zeile (Danach ist $a_{kj} = 0$ für alle $k \neq i$).
- (5) Erhöhe i und j jeweils um 1. Falls $i > m$ oder $j > n$, sind wir fertig. Sonst geht es weiter bei (1).

Danach bleibt eine erweiterte Koeffizientenmatrix $(\mathbf{A}|\mathbf{b})$ in strenger Zeilenstufenform. Da wir nur elementare Zeilenumformungen angewandt haben, hat sich nach Satz 6.1.3 die Lösungsmenge nicht geändert.

Bemerkung. Später wird der Rang für beliebige Matrizen definiert. Wendet man das Gauß-Verfahren auf eine Matrix A an, so bleibt der Rang unverändert. Unsere vorläufige Definition steht also nicht im Widerspruch zur üblichen Bedeutung des Begriffs "Rang".

Wir fassen die obigen Überlegungen noch einmal zusammen.

6.1.3. Satz (Lösbarkeit). Sei $(\mathbf{A}|\mathbf{b})$ eine erweiterte Koeffizientenmatrix in strenger Zeilenstufenform. Dann besitzt das zugehörige lineare Gleichungssystem genau dann Lösungen, wenn alle Absolutglieder b_i mit Index $i > r$ Null sind.

6.1.4. Satz (Konstruktion aller Lösungen). Sei $(\mathbf{A}|\mathbf{b})$ eine erweiterte Koeffizientenmatrix in strenger Zeilenstufenform. Ist n die Anzahl der Unbekannten und ist das System lösbar, so werden die Lösungen durch $n - r$ Parameter beschrieben, wobei r der Rang der Matrix \mathbf{A} ist.

Definition. Die oben in Abhängigkeit von den Parametern c_j für $j \notin \{j_1, \dots, j_r\}$ konstruierte Lösung heisst *allgemeine Lösung*. Jede spezielle Wahl der Parameter liefert eine *spezielle Lösung*.

6.1.5. Satz (Eindeutigkeit). Das System ist genau dann eindeutig lösbar wenn es lösbar ist und $n = r$, d.h. $j_1 = 1, \dots, j_n = n$.

6.1.6. Folgerung. Das homogene System, d.h. $b_i = 0$ für alle $i = 1, \dots, m$, besitzt genau dann nur die triviale Lösung, wenn $r = n$.

6.1.7. Satz (Struktur der Lösungsmenge).

- (1) Seien $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ Lösungen des homogenen linearen Gleichungssystems mit der Koeffizientenmatrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. Dann sind auch $\mathbf{c} + \mathbf{d}$ und $\lambda \mathbf{c}$, $\lambda \in \mathbb{R}$, Lösungen.
- (2) Sei das lineare Gleichungssystem mit der erweiterten Koeffizientenmatrix $(\mathbf{A}|\mathbf{b})$ lösbar. Dann läßt sich eine allgemeine Lösung $\mathbf{v} \in \mathbb{R}^n$ darstellen in der Form

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{u}$$

mit einer speziellen Lösung \mathbf{v}_0 des inhomogenen und einer allgemeinen Lösung \mathbf{u} des zugehörigen homogenen Systems.

Teil (1) besagt, dass die Lösungsmenge eines homogenen Systems immer ein *Untervektorraum* ist. Teil (2) besagt, dass die Lösungsmenge eines inhomogenen Systems entweder leer oder ein *affiner* Unterraum ist.

6.2. Matrizenmultiplikation

6.2.1. Produkte von Zeilen mit Spalten. Das *Produkt* $\mathbf{a} \cdot \mathbf{b}$ eines Zeilenvektors $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}_n$ und eines Spaltenvektors $\mathbf{b} = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ ist definiert durch

$$\mathbf{a} \cdot \mathbf{b} := \sum_{i=1}^n a_i b_i.$$

Eine lineare Gleichung kann man jetzt noch kompakter schreiben als

$$\mathbf{a} \cdot \mathbf{x} = \mathbf{b}$$

mit $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{x} = (x_1, \dots, x_n)^T$.

Es gilt für alle $\alpha \in \mathbb{R}$, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a} \in \mathbb{R}_n$ und $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b} \in \mathbb{R}^n$

$$(\mathbf{a}_1 + \mathbf{a}_2) \cdot \mathbf{b} = \mathbf{a}_1 \cdot \mathbf{b} + \mathbf{a}_2 \cdot \mathbf{b},$$

$$\mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2,$$

$$\alpha (\mathbf{a} \cdot \mathbf{b}) = (\alpha \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha \mathbf{b}).$$

6.2.2. Produkte von Matrizen. Für $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$ und $\mathbf{B} = (b_{ij}) \in \mathbb{R}^{n \times r}$ hat das *Matrizenprodukt* $\mathbf{A} \cdot \mathbf{B}$ die Einträge

$$c_{ij} := \sum_{k=1}^n a_{ik} b_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, r,$$

d.h. c_{ij} ist das Produkt der i -ten Zeile von \mathbf{A} mit der j -ten Spalte von \mathbf{B} .

Achtung! Wenn die Spaltenanzahl von \mathbf{A} ungleich der Zeilenanzahl von \mathbf{B} ist, ist das Produkt nicht definiert!

Ein $m \times n$ Gleichungssystem $\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, \dots, m$ kann man kompakter als $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ schreiben, mit $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$, $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \dots, b_m)^T \in \mathbb{R}^m$, siehe oben.

6.2.3. Rechenregeln. Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$, seien $\mathbf{a}_i \in \mathbb{R}^m$ für $i = 1, \dots, n$ die Spaltenvektoren von \mathbf{A} und sei $\mathbf{x} \in \mathbb{R}^n$. Dann gilt:

$$\mathbf{A} \cdot \mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n.$$

Setzt man nun $\mathbf{x} = \mathbf{e}_i, i = 1, \dots, n$, so erhält man

$$\mathbf{A} \cdot \mathbf{e}_i = \mathbf{a}_i,$$

d.h. Multiplikation einer Matrix \mathbf{A} mit dem i -ten Basisspaltenvektor \mathbf{e}_i liefert i -ten Spaltenvektor von \mathbf{A} . Analog gilt:

$$\mathbf{e}'_j \cdot \mathbf{A} = \mathbf{z}_j, \quad j = 1, \dots, m,$$

d.h. Multiplikation des j -ten Basiszeilenvektors \mathbf{e}'_j mit \mathbf{A} liefert die j -te Zeile von \mathbf{A} .

Definition. Die Matrix

$$\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

heißt $n \times n$ Einheitsmatrix.

Die Einheitsmatrix hat die Spaltendarstellung $\mathbf{E}_n = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ mit den Standardbasisspaltenvektoren $\mathbf{e}_i \in \mathbb{R}^n$.

6.2.1. Satz. Für alle Matrizen $\mathbf{A}, \mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{m \times n}, \mathbf{B}, \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{n \times r}, \mathbf{C} \in \mathbb{R}^{r \times s}$ und $\alpha \in \mathbb{R}$ gilt

- (1) Assoziativgesetz: $\alpha(\mathbf{A} \cdot \mathbf{B}) = (\alpha\mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\alpha\mathbf{B})$,
- (2) 1. Distributivgesetz: $(\mathbf{A}_1 + \mathbf{A}_2) \cdot \mathbf{B} = \mathbf{A}_1 \cdot \mathbf{B} + \mathbf{A}_2 \cdot \mathbf{B}$,
- (3) 2. Distributivgesetz: $\mathbf{A} \cdot (\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A} \cdot \mathbf{B}_1 + \mathbf{A} \cdot \mathbf{B}_2$,
- (4) Homogenität: $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$,
- (5) Einheitsmatrizen sind neutrale Elemente: $\mathbf{E}_m \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{E}_n = \mathbf{A}$.

Achtung! Im Allgemeinen gilt:

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A},$$

selbst wenn beide Produkte definiert sind.

Beispiel.

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & -3 \\ 14 & 19 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 14 & 11 \end{pmatrix}$$

Bemerkung. (1) Es gibt Matrizen $\mathbf{A}, \mathbf{B} \neq \mathbf{0}$ so daß $\mathbf{A} \cdot \mathbf{B} = \mathbf{0}$. Dies ist bei reellen Zahlen nicht möglich, denn für $a, b \in \mathbb{R}$ gilt: $a \cdot b = 0 \Rightarrow (a = 0) \vee (b = 0)$.

(2) Aber es gilt (wenn die Produkte definiert sind):

$$\mathbf{0} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{0} = \mathbf{0}$$

(3) *Potenzen* von quadratischen Matrizen $\mathbf{A} \in \mathbb{R}^{m \times n}$ sind rekursiv definiert:

$$\mathbf{A}^0 = \mathbf{E}_n, \quad \mathbf{A}^k := \mathbf{A}^{k-1} \cdot \mathbf{A} = \underbrace{\mathbf{A} \cdots \mathbf{A}}_{k\text{-mal}}$$

6.2.4. Transponierte Matrizen. Zu jeder $m \times n$ Matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$ gibt es eine *transponierte Matrix* $\mathbf{A}^T = (a_{ji}) \in \mathbb{R}^{n \times m}$, deren i -te Zeile aus den Koeffizienten der i -ten Spalte von \mathbf{A} bestehen, d.h. für

$$\mathbf{A} = \begin{pmatrix} \boxed{a_{11}} & \cdots & \boxed{a_{1n}} \\ \vdots & & \vdots \\ \boxed{a_{m1}} & \cdots & \boxed{a_{mn}} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

ist die transponierte Matrix gegeben durch

$$\mathbf{A}^T = \begin{pmatrix} \boxed{a_{11}} & \cdots & \boxed{a_{m1}} \\ \vdots & & \vdots \\ \boxed{a_{1n}} & \cdots & \boxed{a_{mn}} \end{pmatrix} \in \mathbb{R}^{n \times m}.$$

Die Bezeichnung ist konsistent mit unserer Schreibweise $(a_1, \dots, a_n)^T$ für Spaltenvektoren.

Es gelten folgende Rechenregeln für alle $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R}$:

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T,$$

$$(\alpha \mathbf{A})^T = \alpha \mathbf{A}^T,$$

$$(\mathbf{A}^T)^T = \mathbf{A},$$

$$(\mathbf{A} \cdot \mathbf{B})^T = \mathbf{B}^T \cdot \mathbf{A}^T.$$

Definition. Eine $n \times n$ Matrix \mathbf{A} heisst *symmetrisch*, wenn $\mathbf{A}^T = \mathbf{A}$ gilt; sie heisst *schief-symmetrisch*, falls $\mathbf{A}^T = -\mathbf{A}$ gilt.

Für symmetrische Matrizen gilt

$$\mathbf{A} = \mathbf{A}^T \Leftrightarrow a_{ij} = a_{ji} \quad \forall i, j = 1, \dots, n,$$

und für schief-symmetrische Matrizen gilt analog

$$\mathbf{A} = -\mathbf{A}^T \Leftrightarrow a_{ij} = -a_{ji} \quad \forall i, j = 1, \dots, n.$$

Insbesondere gilt für die Diagonalelemente einer schiefssymmetrischen Matrix $a_{ii} = 0$.

6.2.5. Inverse Matrizen. Eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ heisst *invertierbar*, wenn es eine Matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ gibt so, dass $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{E}_n$ gilt. Diese Matrix \mathbf{B} ist eindeutig bestimmt, wird mit \mathbf{A}^{-1} bezeichnet und heisst *inverse Matrix* von \mathbf{A} .

6.2.2. Satz. Wenn es zu einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ zwei Matrizen $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$ gibt mit $\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{C} = \mathbf{E}$, dann ist \mathbf{A} invertierbar und es gilt:

$$\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}.$$

Beispiel. Für $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc \neq 0$ gilt:

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

6.2.3. Satz.

a) Die inverse Matrix einer invertierbaren Matrix \mathbf{A} ist invertierbar und es gilt:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}.$$

b) Das Produkt zweier invertierbarer Matrizen ist invertierbar und es gilt:

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}.$$

c) Die Transponierte \mathbf{A}^T einer Matrix ist genau dann invertierbar, wenn \mathbf{A} invertierbar ist. In diesem Fall gilt:

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

Für das Produkt mehrerer invertierbarer Matrizen $\mathbf{A}_i, i = 1, \dots, n$, gilt:

$$(\mathbf{A}_1 \cdot \dots \cdot \mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1} \cdot \mathbf{A}_{n-1}^{-1} \cdot \dots \cdot \mathbf{A}_1^{-1}.$$

Invertierbare Matrizen sind bedeutend für die Lösung von Gleichungssystemen. Aus der Gleichung

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

folgt

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b},$$

falls die Matrix \mathbf{A} invertierbar ist. Dieses so berechnete \mathbf{x} ist die eindeutige Lösung des obigen linearen Gleichungssystems.

Umgekehrt erhält man das Inverse \mathbf{B} einer quadratischen $n \times n$ -Matrix \mathbf{A} durch das simultane Lösen von n linearen Gleichungssystemen. Sei $1 \leq j \leq n$, und sei \mathbf{b}_j die j -te Spalte von \mathbf{B} . Aus $\mathbf{A} \cdot \mathbf{B} = \mathbf{E}_n$ folgt

$$\mathbf{A} \cdot \mathbf{b}_j = \mathbf{e}_j \quad \text{für alle } j = 1, \dots, n.$$

Wir müssen also das Gleichungssystem $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ mit den rechten Seiten $\mathbf{e}_1, \dots, \mathbf{e}_n$ lösen.

Dazu schreiben wir alle rechten Seiten hinter den Querstrich und erhalten die erweiterte Matrix $(\mathbf{A}|\mathbf{E}_n)$. Mit dem Gauß-Verfahren aus Abschnitt 6.1.4 bringen wir die Matrix links vom Strich in strenge Zeilenstufenform. Falls der Rang $r = n$ ist, erhalten wir $(\mathbf{E}_n|\mathbf{B})$, und \mathbf{B} ist die inverse Matrix. Falls der Rang $r < n$ ist, erhalten wir für (mindestens) eine rechte Seite eine Gleichung der Form $0 = * \neq 0$, die nicht lösbar ist. In diesem Fall ist \mathbf{A} nicht invertierbar.

Da \mathbf{A} nicht invertierbar ist falls $r < n$, dürfen wir Schritt (1) im Gauß-Verfahren wie folgt modifizieren:

(1') Falls $a_{ij} = \dots = a_{mj} = 0$ brich ab: die Matrix \mathbf{A} ist nicht invertierbar.

6.2.6. Diagonal- und Dreiecksmatrizen. Sei $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Die Elemente $a_{ii}, i = 1, \dots, n$, nennt man *Diagonalelemente*. Eine Matrix die nur auf der Diagonalen nichttriviale Einträge hat heisst *Diagonalmatrix*. Wir schreiben

$$\text{Diag}(a_1, \dots, a_n) := \begin{pmatrix} a_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_n \end{pmatrix}.$$

Eine Matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times n}$, für die $a_{ij} = 0$ für alle $1 \leq j < i \leq n$ gilt, d.h. alle Einträge unterhalb der Diagonalen sind Null, heisst *obere Dreiecksmatrix*. Analog heisst eine Matrix *untere Dreiecksmatrix*, wenn $a_{ij} = 0$ für alle $1 \leq i < j \leq n$ gilt.

6.2.4. Satz. *Eine obere (bzw. untere) Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonalelemente von Null verschieden sind.*

Obere Dreiecksmatrizen lassen sich mit dem Gauß-Verfahren besonders einfach invertieren. Für untere Dreiecksmatrizen \mathbf{A} benutzen wir, dass

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} \implies \mathbf{A}^{-1} = ((\mathbf{A}^T)^{-1})^T$$

nach Satz 6.2.3, und invertieren zunächst die transponierte Matrix \mathbf{A}^T , und transponieren danach das Inverse noch einmal.

Im Spezialfall $\mathbf{A} = \text{Diag}(a_1, \dots, a_n)$ mit $a_i \neq 0, i = 1, \dots, n$, gilt:

$$\mathbf{A}^{-1} = \text{Diag}(a_1^{-1}, \dots, a_n^{-1})$$

6.3. Vektorräume

Vektorräume sind oft Mengen, deren Elemente man addieren und mit Faktoren multiplizieren kann. z.B.

- Vektoren im Anschauungsraum, Ebene,
- Funktionen $f : I \rightarrow \mathbb{R}$, stetige / differenzierbare Funktionen ...
- Matrizen, Zeilen- und Spaltenvektoren.

Diese Objekte verhalten sich unter diesem Gesichtspunkt sehr ähnlich. Um gemeinsame Eigenschaften besser zu verstehen und eine einheitliche Theorie zu schaffen, führen wir einen neuen Begriff und die zugehörigen Theorie ein.

6.3.1. Abstrakte Vektorräume. Eine nichtleere Menge V , in der man zu je zwei Elementen $\mathbf{a}, \mathbf{b} \in V$ eine Summe $\mathbf{a} + \mathbf{b} \in V$ und zu jedem Element $\mathbf{a} \in V$ und jedem Skalar $\lambda \in \mathbb{R}$ das λ -fache $\lambda \mathbf{a} \in V$ bilden kann, heisst \mathbb{R} -Vektorraum, wenn folgende Axiome erfüllt sind:

(V.1) Die Addition ist assoziativ, d.h. für alle $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$ gilt:

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) .$$

(V.2) Es gibt ein Element $\mathbf{0} \in V$, Nullvektor genannt, mit

$$\mathbf{a} + \mathbf{0} = \mathbf{a}$$

für alle $\mathbf{a} \in V$.

(V.3) Zu jedem $\mathbf{a} \in V$ gibt es genau ein mit $-\mathbf{a}$ bezeichnetes Element in V mit

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0} .$$

(V.4) Die Addition ist kommutativ, d.h. für alle $\mathbf{a}, \mathbf{b} \in V$ gilt:

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} .$$

(V.5) $\lambda(\mu \mathbf{a}) = (\lambda\mu) \mathbf{a}$ für alle $\lambda, \mu \in \mathbb{R}, \mathbf{a} \in V$.

(V.6) $\lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$ für alle $\lambda \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in V$.

(V.7) $(\mu + \lambda) \mathbf{a} = \mu \mathbf{a} + \lambda \mathbf{a}$ für alle $\lambda, \mu \in \mathbb{R}, \mathbf{a} \in V$.

(V.8) $1 \cdot \mathbf{a} = \mathbf{a}$ für alle $\mathbf{a} \in V$.

Die Elemente des Vektorraums V nennt man *Vektoren*.

Die obigen Rechenregeln haben wir bereits in Abschnitt 1.3.2 für Vektoren in der Ebene kennengelernt.

Beispiel. (1) $V = \mathbb{R}^{m \times n}$ ist ein Vektorraum mit den komponentenweise definierten Operationen Addition und skalarer Multiplikation (siehe Abschnitt 6.1.2).

(2) Insbesondere sind

$$\begin{aligned}\mathbb{R}_n &= \{(a_1, \dots, a_n) \mid a_i \in \mathbb{R}\} \\ \mathbb{R}^n &= \{(a_1, \dots, a_n)^T \mid a_i \in \mathbb{R}\}\end{aligned}$$

Vektorräume.

(3) Die Menge

$$P_n := \{a_0 + a_1x + \dots + a_nx^n \mid a_i \in \mathbb{R}, i = 0, \dots, n\}$$

der Polynome vom Grad $\leq n$ mit punktweise definierter Addition und skalarer Multiplikation ist ein \mathbb{R} -Vektorraum. Genauso bilden die Polynome von beliebigem Grad einen \mathbb{R} -Vektorraum.

(4) Die Menge $C^0(I) := \{f: I \rightarrow \mathbb{R} \mid f \text{ stetig}\}$ der stetigen Funktionen mit den Operationen

$$\begin{aligned}(f + g)(x) &:= f(x) + g(x) \\ (\lambda f)(x) &:= \lambda f(x)\end{aligned}$$

bildet den Vektorraum der stetigen Funktionen.

6.3.2. Unterräume und Linearkombinationen. Im folgenden sei V ein \mathbb{R} -Vektorraum. Eine nichtleere Menge $U \subseteq V$ heisst *Unterraum* von V , wenn zu je zwei Elementen $\mathbf{u}, \mathbf{v} \in U$ auch deren Summe $\mathbf{u} + \mathbf{v}$ in U liegt und wenn mit jedem $\mathbf{u} \in U$ und $\lambda \in \mathbb{R}$ auch $\lambda\mathbf{u}$ in U liegt, d.h.:

- (U.1) U ist abgeschlossen unter Addition: $\mathbf{u}, \mathbf{v} \in U \implies \mathbf{u} + \mathbf{v} \in U$,
 (U.2) U ist abgeschlossen unter skalarer Multiplikation: $\mathbf{u} \in U, \lambda \in \mathbb{R} \implies \lambda\mathbf{u} \in U$,
 (U.3) U ist nichtleer: $\mathbf{0} \in U$.

Aus den Axiomen (V.1)–(V.8) folgt somit sofort, dass dann auch U wieder ein Vektorraum ist.

Beispiel. (1) V besitzt die *trivialen Unterräume* $U = \{\mathbf{0}\}, U = V$.

(2) Sei $\mathbf{v} \in V$. Dann ist

$$U := \{\lambda\mathbf{v}; \lambda \in \mathbb{R}\}$$

ein Untervektorraum von V . Für $V = \mathbb{R}^3$ sind dies alle zu \mathbf{v} parallelen Vektoren.

(3) Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$. Dann ist

$$\text{Ker } \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^n; \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

ein Unterraum von \mathbb{R}^n . *Beachte:* $\text{Ker } \mathbf{A}$ ist genau die Lösungsmenge des homogenen linearen Gleichungssystemen $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$.

Definition. Jede aus endlich vielen Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ gebildete Summe der Form

$$\sum_{i=1}^k \alpha_i \mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k$$

mit Koeffizienten $\alpha_i \in \mathbb{R}$ heißt *Linearkombination* der \mathbf{v}_i . Eine solche Linearkombination heißt *trivial*, wenn $\alpha_i = 0$ für $i = 1, \dots, k$ gilt. Die Menge aller Linearkombinationen der \mathbf{v}_i heisst *lineare Hülle der \mathbf{v}_i* und wird mit

$$\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) := \left\{ \sum_{i=1}^k \alpha_i \mathbf{v}_i \mid \alpha_i \in \mathbb{R}, i = 1, \dots, k \right\}$$

bezeichnet.

6.3.1. Lemma. Die lineare Hülle der Vektoren $\mathbf{v}_i \in V, i = 1, \dots, k$, ist ein Unterraum von V .

Definition. Man sagt, ein Unterraum U von V wird von den Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ erzeugt oder auch, $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ ist ein *Erzeugendensystem* von U , wenn

$$U = \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) .$$

Beispiel. (1) \mathbb{R}^n wird von $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ erzeugt, denn

$$\mathbf{a} = (a_1, \dots, a_n)^T = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_n \mathbf{e}_n .$$

(2) Der \mathbb{R}^2 wird von $(\mathbf{e}_1, \mathbf{e}_2)$ erzeugt, aber auch $(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2)$ erzeugt \mathbb{R}^2 . Dies entspricht einer Drehung des Koordinatensystems.

Definition. Endlich viele Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ heißen *linear abhängig*, wenn es Zahlen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ gibt, die nicht sämtlich gleich Null sind, so, dass

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

gilt. Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ heissen *linear unabhängig*, wenn sie nicht linear abhängig sind, d.h.

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0} \implies \alpha_1 = \alpha_2 = \dots = \alpha_k = 0 .$$

Um zu überprüfen, ob $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ linear unabhängig sind, muss man das lineare Gleichungssystem in x_1, \dots, x_k

$$(*) \quad x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k = \mathbf{V} \cdot \mathbf{x} = \mathbf{0}$$

betrachten und überprüfen, ob es nur triviale Lösungen gibt. Hierbei ist \mathbf{V} die Matrix mit den Spalten \mathbf{v}_i . Es gilt:

- (1) $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear abhängig \iff (*) besitzt eine nichttriviale Lösung.
- (2) $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear unabhängig \iff (*) besitzt nur triviale Lösung.

6.3.2. Lemma. Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ sind genau dann linear abhängig, wenn sich einer von ihnen als Linearkombination der anderen darstellen lässt.

6.3.3. Lemma.

- (1) Jedes endliche System von Vektoren, das linear abhängige Vektoren enthält, ist linear abhängig.
- (2) Jedes endliche System von Vektoren, das den Nullvektor enthält, ist linear abhängig.
- (3) Jedes Teilsystem linear unabhängiger Vektoren ist linear unabhängig.

6.3.4. Satz. In einer Matrix in strenger Zeilenstufenform sind die nichttrivialen Zeilenvektoren linear unabhängig.

6.3.5. Satz. Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Dann sind die folgenden Aussagen äquivalent:

- (1) \mathbf{A} ist invertierbar.
- (2) Die Spalten von \mathbf{A} sind linear unabhängig.
- (3) Die Zeilen von \mathbf{A} sind linear unabhängig.

6.4. Determinanten

6.4.1. Einführung. Gegeben seien n Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$. Diese Vektoren spannen ein *Parallelotop*

$$P(\mathbf{v}_1, \dots, \mathbf{v}_n) = \{ \lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n \mid 0 \leq \lambda_1, \dots, \lambda_n \leq 1 \}$$

auf. Das n -dimensionale Volumen dieses Polytops hat drei charakteristische Eigenschaften.

- (1) *Homogenität.* Für alle $\lambda \in \mathbb{R}$ und $1 \leq i \leq n$ gilt

$$\text{vol}(P(\mathbf{v}_1, \dots, \lambda \mathbf{v}_i, \dots, \mathbf{v}_n)) = |\lambda| \cdot \text{vol}(P(\mathbf{v}_1, \dots, \mathbf{v}_n)),$$

- (2) *Scherungsinvarianz.* Für alle $\lambda \in \mathbb{R}$ und $1 \leq i \neq j \leq n$ gilt

$$\text{vol}(P(\mathbf{v}_1, \dots, \underbrace{\mathbf{v}_j + \lambda \mathbf{v}_i}_{j\text{-te Stelle}}, \dots, \mathbf{v}_n)) = \text{vol}(P(\mathbf{v}_1, \dots, \mathbf{v}_n)).$$

- (3) *Normierung.* Das Volumen des n -dimensionalen Standardwürfels ist 1, d.h., es gilt $\text{vol}(P(\mathbf{e}_1, \dots, \mathbf{e}_n)) = 1$.

Durch diese Eigenschaften ist das Volumen bereits vollständig festgelegt.

Beispiel. (1) $n = 1$: das Parallelotop ist eine Strecke der Länge $|\mathbf{v}_1|$.

(2) $n = 2$: seien $\mathbf{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix}$ und $\mathbf{v}_2 = \begin{pmatrix} c \\ d \end{pmatrix}$, dann hat das aufgespannte Parallelogramm die Fläche $|ad - bc|$.

(3) $n = 3$: das Volumen eines Spats $P(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ im \mathbb{R}^3 wird durch den Betrag $|\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle|$ des Spatprodukts gegeben, siehe Satz 1.4.1.

6.4.2. Definition der Determinante. Die *Determinante* von n Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ ist eine Zahl $\det(\mathbf{v}_1, \dots, \mathbf{v}_n) \in \mathbb{R}$ mit folgenden Eigenschaften.

(1) *Homogenität.* Für alle $\lambda \in \mathbb{R}$ und $1 \leq i \leq n$ gilt

$$\det(\mathbf{v}_1, \dots, \lambda \mathbf{v}_i, \dots, \mathbf{v}_n) = \lambda \cdot \det(\mathbf{v}_1, \dots, \mathbf{v}_n),$$

(2) *Scherungsinvarianz.* Für alle $\lambda \in \mathbb{R}$ und $1 \leq i \neq j \leq n$ gilt

$$\det(\mathbf{v}_1, \dots, \underbrace{\mathbf{v}_j + \lambda \mathbf{v}_i}_{j\text{-te Stelle}}, \dots, \mathbf{v}_n) = \det(\mathbf{v}_1, \dots, \mathbf{v}_n).$$

(3) *Normierung.* Es gilt $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$.

Durch diese Eigenschaften ist die Determinante bereits vollständig festgelegt. Die *Determinante* $\det \mathbf{A}$ einer quadratischen Matrix ist die Determinante der Spaltenvektoren. Man schreibt manchmal auch

$$\det \mathbf{A} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix},$$

aber Vorsicht — bei 1×1 -Matrizen sollte man $\det(a)$ nicht mit dem Absolutbetrag $|a|$ verwechseln.

Beispiel. (1) $n = 1$: Es gilt $\det(a) = a$.

(2) $n = 2$: Es gilt

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

(3) $n = 3$: Die Determinante einer 3×3 -Matrix ist das Spatprodukt der Spaltenvektoren. Sie lässt sich mit der *Sarrusschen Regel* berechnen, siehe dazu die Formel am Ende von Abschnitt 1.4.4.

Bemerkung. Es gilt

$$\text{vol}(P(\mathbf{v}_1, \dots, \mathbf{v}_n)) = |\det(\mathbf{v}_1, \dots, \mathbf{v}_n)|.$$

Sei $\mathbf{A} = (a_{ij})$ eine $m \times n$ -Matrix, dann bezeichnen wir mit \mathbf{A}_{ij} die $(m-1) \times (n-1)$ -Matrix, bei der die i -te Zeile und die j -te Spalte von \mathbf{A} fehlen:

$$\mathbf{A}_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m1} & \dots & a_{m,j-1} & a_{m,j+1} & \dots & a_{mn} \end{pmatrix}.$$

6.4.1. Satz (Laplace-Entwicklung). *Sei \mathbf{A} eine $n \times n$ -Matrix. Dann gilt für alle $1 \leq i, j \leq n$, dass*

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{i+k} a_{ik} \cdot \det \mathbf{A}_{ik} \quad (\text{Entwicklung nach der } i\text{-ten Zeile}),$$

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{k+j} a_{kj} \cdot \det \mathbf{A}_{kj} \quad (\text{Entwicklung nach der } j\text{-ten Spalte}).$$

Damit lässt sich die Berechnung von $n \times n$ -Determinanten rekursiv auf die Berechnung von $(n-1) \times (n-1)$ -Determinanten zurückführen. Dieses Verfahren ist sehr rechenaufwendig: man benötigt bei naivem Vorgehen über $n!$ Multiplikationen. Speichert man alle Determinanten, die als Zwischenergebnisse auftreten, so braucht man immer noch bis zu $n(2^{n-1} - 1)$ Multiplikationen (und jede Menge Speicherverwaltung). Das Verfahren lohnt sich dann, wenn sehr viele der Matrixeinträge von \mathbf{A} verschwinden.

6.4.3. Rechenregeln für Determinanten. Aus den obigen Eigenschaften (1)–(3) ergeben sich folgende Rechenregeln.

6.4.2. Satz. *Seien $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w} \in \mathbb{R}^n$ Vektoren, und sei $\lambda \in \mathbb{R}$. Dann gilt*

(1) *Es gilt $\det(\mathbf{v}_1, \dots, \mathbf{v}_n) = 0$ genau dann, wenn $\mathbf{v}_1, \dots, \mathbf{v}_n$ linear abhängig sind.*

(2) *Vertauschen zweier Vektoren ändert das Vorzeichen: für alle $i \neq j$ gilt*

$$\det(\mathbf{v}_1, \dots, \underbrace{\mathbf{v}_j}_{i\text{-te Stelle}}, \dots, \underbrace{\mathbf{v}_i}_{j\text{-te Stelle}}, \dots, \mathbf{v}_n) = -\det(\mathbf{v}_1, \dots, \mathbf{v}_n).$$

(3) *Die Determinante ist in jedem Argument linear, d.h.,*

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{w}, \dots, \mathbf{v}_n) = \det(\mathbf{v}_1, \dots, \mathbf{v}_n) + \det(\mathbf{v}_1, \dots, \underbrace{\mathbf{w}}_{i\text{-te Stelle}}, \dots, \mathbf{v}_n),$$

$$\det(\mathbf{v}_1, \dots, \lambda \mathbf{v}_i, \dots, \mathbf{v}_n) = \lambda \cdot \det(\mathbf{v}_1, \dots, \mathbf{v}_n).$$

(4) *Es gilt $\det \mathbf{A}^T = \det \mathbf{A}$.*

(5) *Sei \mathbf{A} obere oder untere Dreiecksmatrix, dann ist $\det \mathbf{A}$ das Produkt der Diagonalelemente.*

(6) *Die Determinante ist multiplikativ, d.h., es gilt $\det(\mathbf{A} \cdot \mathbf{B}) = \det \mathbf{A} \cdot \det \mathbf{B}$.*

Die Punkte (1)–(3) sagen etwas über Operationen mit Spalten einer Matrix aus. Wegen (4) gelten (1)–(3) analog für Operationen mit den Zeilen. Wir können daher Determinanten nach dem Gauß-Verfahren berechnen.

6.4.4. Berechnung mit dem Gauß-Verfahren. Wir bringen eine quadratische Matrix \mathbf{A} auf Zeilenstufenform durch elementare Zeilenumformungen. Da sich der Wert der Determinante bei einigen dieser Umformungen ändert, und da wir \mathbf{A} wegen Satz 6.4.2 (5) nur auf obere Dreiecksgestalt bringen müssen, modifizieren wir den Gauß-Algorithmus wie folgt.

Zunächst schreibe die Matrix \mathbf{A} zwischen senkrechte Striche wie oben.

- (0) Beginne mit $i = 1$.
- (1') Falls $a_{ii} = \dots = a_{ni} = 0$, brich ab: die Determinante von \mathbf{A} ist 0.
- (2') Falls $a_{ii} = 0$, aber $a_{ki} \neq 0$ für ein $k > i$, vertausche die i -te und die k -te Zeile und schreibe einen zusätzlichen Faktor “-” vor die Determinante.
- (3') Dividiere die gesamte i -te Zeile durch $a_{ii} \neq 0$ und schreibe einen zusätzlichen Faktor a_{ii} vor die Determinante.
- (4') Für alle $i < k \leq n$ subtrahiere das a_{ki} -fache der i -ten Zeile von der k -ten Zeile.
- Erhöhe i um 1. Falls $i > n$, sind wir fertig. Sonst geht es weiter bei (1).

Am Ende bleibt eine obere Dreiecksmatrix \mathbf{D} mit $d_{11} = \dots = d_{nn} = 1$ stehen. Die Determinante ist dann einfach der Vorfaktor.

Beispiel. Wir berechnen

$$\det \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & 3 \end{vmatrix} = - \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & \frac{3}{2} \\ 0 & 1 \end{vmatrix} = -2.$$

Beim Gauß-Verfahren braucht man ca. $\frac{n^3}{3}$ Multiplikationen und Divisionen. Für $n \times n$ -Matrizen mit wenig Nullen und $n \geq 4$ ist das Gauß-Verfahren bereits schneller als die Laplace-Entwicklung.