

Mathematik für Studierende der Informatik und des Ingenieurwesens

Kurzskript

Wintersemester 2007/08
und Sommersemester 2008

Dieter Wolke

orientiert an dem Buch von K. Meyberg und P. Vachenauer:
Höhere Mathematik Bd.1, Springer-Verlag

Als Leitfaden kann auch empfohlen werden: Merziger, Wirth:
Repertitorium der höheren Mathematik, Binomi-Verlag.

Inhaltsverzeichnis

1	Mathematische Grundbegriffe	5
1.1	Mengen und Abbildungen	5
1.2	Reelle Zahlen	8
1.3	Die Ebene	12
1.4	Der Raum	15
1.5	Produkte	17
1.6	Komplexe Zahlen	20
2	Funktionen, Grenzwerte, Stetigkeit	23
2.1	Grundbegriffe	23
2.2	Polynome und rationale Funktionen	24
2.3	Trigonometrische Funktionen	28
2.4	Zahlenfolgen und Grenzwerte	32
2.5	Rechenregeln für Grenzwerte und Konvergenzkriterien	35
2.6	Funktionengrenzwerte, Stetigkeit	36
3	Differentiation	39
3.1	Die Ableitung	39
3.2	Anwendungen der Differentiation	42
3.3	Umkehrfunktionen	45
3.4	Exponential- und Logarithmusfunktion	47
4	Integration	53
4.1	Das bestimmte Integral	53
4.2	Integrationsregeln	57
4.3	Integration rationaler Funktionen	59
4.4	Uneigentliche Integrale	61
4.5	Kurven-, Längen- und Flächenmessung	64
5	Potenzreihen	69
5.1	Unendliche Reihen	69
5.2	Reihen von Funktionen	71
5.3	Potenzreihen	73
5.4	Taylor-Reihen	77
5.5	Fourier-Reihen	79

6	Lineare Algebra	81
6.1	Lineare Gleichungssysteme und Matrizen	81
6.2	Matrizenmultiplikation	88
6.3	Vektorräume	92
6.4	Elementarmatrizen	97
6.5	Determinanten	100
6.6	Lineare Abbildungen und Eigenwerte	104
6.7	Symmetrische Matrizen	117
7	Differentiation von Funktionen mehrerer Variablen	121
7.1	Kurven im \mathbb{R}^n	121
7.2	Reellwertige Funktionen mehrerer Veränderlicher	124
7.3	Anwendungen	130
7.4	Vektorwertige Funktionen	134
8	Integration von Funktionen mehrerer Variablen	137
8.1	Parameterintegrale	137
8.2	Integrale auf rechteckigen Bereichen	139
8.3	Integrale auf Normalbereichen	140
8.4	Kurven- und Oberflächenintegrale	143
8.5	Die Integralsätze	146

Kapitel 1

Mathematische Grundbegriffe

Ein wesentliches Ziel dieses Kapitels ist es, die wichtigsten Begriffe und Bezeichnungen, die im Weiteren benötigt werden, einzuführen. Vieles ist dabei bereits aus der Schule bekannt. Insbesondere soll eine Klarstellung und Vereinheitlichung der Notation erreicht werden.

1.1 Mengen und Abbildungen

Definition 1.1.1 *Eine Menge ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. (Georg Cantor, 1845–1918. Begründer der Mengenlehre, ca. 1875)*

Bezeichnungen:

- Objekte der Mengen heißen **Elemente**.

$$\begin{aligned} a \in A, & \quad a \text{ ist Element von } A, \\ a \notin A, & \quad a \text{ ist kein Element von } A. \end{aligned}$$

- Beschreibung einer Menge durch Aufzählen ihrer Elemente.
Bsp. $M = \{1, 2, 3, 4\}$.
Hinweis: Innerhalb der Mengenklammern soll kein Element mehrfach auftreten.
- Beschreibung einer Menge X durch Angeben einer definierenden Eigenschaft

$$X := \{x \in M; x \text{ hat die Eigenschaft } E\}.$$

Dabei ist M eine (umfassende) „Grundmenge“.

Bsp.

$$\begin{aligned} X &= \{x \text{ ist natürliche Zahl; } x \text{ ist durch } 3 \text{ teilbar}\} \\ &= \{3, 6, 9, \dots\}. \end{aligned}$$

- Die **leere Menge** \emptyset enthält kein Element.
- B heißt **Teilmenge** von A , wenn jedes Element von B auch ein Element von A ist

$$B \subseteq A \Leftrightarrow (b \in B \Rightarrow b \in A).$$

- B heißt **echte Teilmenge** von A , in Zeichen $B \subsetneq A$, wenn $B \subseteq A$, und es ein $x \in A$ gibt mit $x \notin B$.

Definition 1.1.2 Eine **Aussage** ist ein sinnvolles sprachliches Gebilde, das entweder wahr oder falsch ist.

- Aussagen kann man miteinander verknüpfen. Der Wahrheitsgehalt der verknüpften Aussagen wird durch die Wahrheitstabellen gegeben. Es gibt:
Negation $\neg A$, **Konjunktion** $A \wedge B$, **Alternative** $A \vee B$,
Implikation $A \Rightarrow B$, **Äquivalenz** $A \Leftrightarrow B$

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
W	W	F	W	W	W	W
W	F		F	W	F	F
F	W	W	F	W	W	F
F	F		F	F	W	W

- **Existenzquantor** \exists

$$\exists x \in A : E(x)$$

heißt: es existiert ein x aus A mit der Eigenschaft E .

- **Generalisations- oder Allquantor** \forall

$$\forall x \in A : E(x)$$

heißt: alle x aus A haben die Eigenschaft E .

1.1.3 (Mengenoperationen) Seien $A, B \subseteq M$ zwei Teilmengen von M . Dann definieren wir folgende Verknüpfungen:

$$\begin{aligned} \text{Durchschnitt} \quad A \cap B &:= \{x \in M; (x \in A) \wedge (x \in B)\}, \\ \text{Vereinigung} \quad A \cup B &:= \{x \in M; (x \in A) \vee (x \in B)\}, \\ \text{Differenz} \quad A \setminus B &:= \{x \in M; (x \in A) \wedge (x \notin B)\}, \\ \text{Komplement} \quad A^c &:= \{x \in M; x \notin A\}. \end{aligned}$$

Hinweis: Während es bei $A \cap B, A \cup B, A \setminus B$ auf die Wahl der „Grundmenge“ M nicht ankommt, ist A^c von M abhängig.

- Mengen heißen **disjunkt**, wenn sie kein gemeinsames Element haben.
- Die **Produktmenge** $A \times B$ zweier Mengen A, B ist definiert durch

$$A \times B := \{(a, b); (a \in A) \wedge (b \in B)\},$$

d.h. sie ist die Menge der geordneten Paare, in denen die erste Komponente aus A und die zweite aus B stammt.

- $A \times B$ ist eine neue Menge. Zwei Elemente $(a, b), (c, d)$ sind gleich, in Zeichen $(a, b) = (c, d)$, genau dann wenn $(a = c) \wedge (b = d)$.
- Analog definiert man für mehrere Mengen $A_i, i = 1, \dots, n$,

$$A_1 \times \cdots \times A_n = \{(a_1, \dots, a_n); a_i \in A_i, i = 1, \dots, n\}.$$

(a_1, \dots, a_n) heißt **geordnetes n-Tupel**. Falls $A_1 = \cdots = A_n = A$ schreibt man

$$A^n \quad \text{statt} \quad \underbrace{A \times \cdots \times A}_{n\text{-mal}}$$

Definition 1.1.4 Seien A, B Mengen. Eine **Funktion** oder **Abbildung** von A nach B ist eine Teilmenge f der Produktmenge $A \times B$ derart, dass zu jedem $x \in A$ genau ein $y \in B$ existiert mit $(x, y) \in f$.

Bezeichnungen:

- Statt $(x, y) \in f$ schreibt man $y = f(x)$ oder $f : A \rightarrow B : x \mapsto f(x)$.
 $y = f(x)$ nennt man den **Funktionswert von f an der Stelle x** .
- A heißt **Definitionsbereich von f** , in Zeichen $D(f) := A$. B heißt der **Zielbereich von f** .
- Sei $C \subseteq A$, dann ist das **Bild von C unter f** definiert durch

$$f(C) := \{f(x); x \in C\}.$$

Insbesondere heißt $f(A)$ **Wertebereich von f** .

- Eine Funktion $f : A \rightarrow B$ heißt
 - surjektiv** $\Leftrightarrow f(A) = B$ (Wertebereich ist die gesamte Menge B),
 - injektiv** $\Leftrightarrow (x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2))$
(verschiedene Argumente haben verschiedene Bilder),
 - bijektiv** \Leftrightarrow injektiv und surjektiv.
- Sei $f : A \rightarrow B$ bijektiv, dann ist die **Umkehrabbildung** f^{-1} definiert als

$$f^{-1} := \{(f(x), x); x \in A\}.$$

- Die **identische Abbildung** ist gegeben durch

$$\text{id} : A \rightarrow A : x \mapsto x$$

- **Gleichheit von zwei Abbildungen:** Seien $f : A \rightarrow B, g : C \rightarrow D$ zwei Abbildungen. Es gilt:

$$f = g \quad \Leftrightarrow \quad ((A = C) \wedge (B = D) \wedge (\forall x \in A : f(x) = g(x))).$$

- Die **Restriktion** einer Funktion $f : A \rightarrow B$ auf $A_0 \subset A$ ist definiert durch:

$$f|_{A_0} := \{(x, f(x)); x \in A_0\}.$$

1.2 Reelle Zahlen

Bezeichnungen:

$\mathbb{N} = \{1, 2, 3, \dots\}$	natürliche Zahlen
$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$	
$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$	ganze Zahlen
$\mathbb{Q} = \left\{ \frac{a}{n}; a \in \mathbb{Z}, n \in \mathbb{N} \right\}$	rationale Zahlen
\mathbb{R}	reelle Zahlen

In \mathbb{Q} gelten die bekannten Rechenregeln für Addition und Multiplikation (Körper-Axiome), die Regeln für die Ordnungsrelation \leq (s.u.), aber nicht das Vollständigkeitsaxiom 1.2.4. Durch eine wohlbegründete, aber komplizierte Konstruktion kann \mathbb{Q} zum Körper \mathbb{R} der reellen Zahlen erweitert werden. Dabei bleiben die Rechengesetze und die Anordnung erhalten. Hinzu kommt die Vollständigkeit.

Definition 1.2.1 Dezimaldarstellung der reellen Zahlen.

- Jedes $n \in \mathbb{N}$ besitzt eine eindeutige Darstellung

$$n = a_k \cdot 10^k + a_{k-1} \cdot 10^{k-1} + \dots + a_1 \cdot 10 + a_0$$

mit $k \in \mathbb{N}_0$; $a_0, \dots, a_k \in \{0, 1, \dots, 9\}$; $a_k \neq 0$.

Kurz: $n = a_k \dots a_0$.

Entsprechend bei anderen Basen, z.B. 2 statt 10, „Binärsystem“.

- Jedes $\alpha \in \mathbb{R}$ besitzt eine eindeutige Darstellung

$$\alpha = g + \frac{b_1}{10} + \frac{b_2}{10^2} + \dots$$

mit $b_1, b_2, \dots \in \{0, 1, \dots, 9\}$, $g \in \mathbb{Z}$.

Kurz: $\alpha = g + 0, b_1 b_2 \dots$

der Fall $b_k = b_{k+1} = \dots = 9$ ab einer Nummer k ist ausgeschlossen.

$$0, b_1 \dots b_{k-1} 99 \dots \quad (b_{k-1} \neq 9)$$

wird mit $0, b_1 \dots (b_{k-1} + 1) 00 \dots$ identifiziert (bzw. $0, 99 \dots = 1 + 0, 00 \dots$). Zur genauen Bedeutung der „unendlichen Summe“

$$\frac{b_1}{10} + \frac{b_2}{100} + \dots \text{ s. Kap. 5.1.}$$

- Genau die rationalen Zahlen haben eine Dezimaldarstellung, die ab einer Stelle periodisch wird. Bsp.

$$\frac{1}{12} = 0,0833 \dots = 0,08\bar{3},$$

$$\frac{3}{7} = 0,428571428571 \dots = 0,\overline{428571}.$$

Axiom 1.2.2 (Ordnungsaxiome).

- Man kann zwei beliebige reelle Zahlen der Größe nach vergleichen, d.h. es gilt immer eine der drei Möglichkeiten

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

Man schreibt

$$x \leq y \quad \text{für} \quad (x < y) \vee (x = y).$$

Die Ordnungsrelation ist verträglich mit den Operatoren $+$ und \cdot . Es gilt

$$x \leq y \wedge a \leq b \implies a + x \leq b + y \tag{1}$$

$$x \leq y \wedge a \geq 0 \implies ax \leq ay$$

$$x \leq y \iff -y \leq -x \tag{2}$$

$$0 < x \leq y \iff 0 < \frac{1}{y} \leq \frac{1}{x}.$$

Definition 1.2.3 Sei $S \subseteq \mathbb{R}$. S heißt **nach oben beschränkt**, wenn es eine Zahl $b \in \mathbb{R}$ gibt, so dass

$$\forall x \in S: \quad x \leq b.$$

Man nennt b eine **obere Schranke** von S .

- Analog definiert man die Begriffe **nach unten beschränkt** und **untere Schranke**.
- S heißt **beschränkt**, falls S eine obere und eine untere Schranke besitzt.

Axiom 1.2.4 (Vollständigkeitsaxiom).

Jede nach oben beschränkte, nichtleere Menge reeller Zahlen besitzt eine kleinste obere Schranke.

- Die durch das Axiom gesicherte kleinste obere Schranke s einer Menge S ist eindeutig bestimmt und wird **Supremum** von S ($s = \sup S$) genannt. $\sup S$ kann charakterisiert werden durch die Eigenschaft

$$s \text{ ist obere Schranke von } S \wedge \forall \varepsilon > 0 \exists x \in S : s - \varepsilon < x \leq s.$$

- Analog bei Beschränktheit nach unten das Infimum einer Menge.
Hinweis: Auf \mathbb{Q} gelten die Körper- und Ordnungsaxiome, aber nicht das Vollständigkeitsaxiom. \mathbb{R} ist die „kleinste“ vollständige Erweiterung von \mathbb{Q} .

Definition 1.2.5 Der **Betrag** $|a|$ einer reellen Zahl $a \in \mathbb{R}$ ist definiert durch

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

- Aus der Definition ergeben sich folgende Rechenregeln:

$$\begin{aligned} -|a| &\leq a \leq |a|, \\ |-a| &= |a|, \end{aligned} \tag{1}$$

$$\begin{aligned} |ab| &= |a||b|, \\ \left| \frac{a}{b} \right| &= \frac{|a|}{|b|} \quad \text{falls } b \neq 0, \end{aligned} \tag{2}$$

$$|a| \leq b \Leftrightarrow -b \leq a \leq b. \tag{3}$$

- Aus (1) und (3) folgt die **Dreiecksungleichung**: $\forall a, b \in \mathbb{R}$ gilt:

$$|a + b| \leq |a| + |b|. \tag{4}$$

- Seien $a, b \in \mathbb{R}$. Dann wird mit $|a - b|$ der **Abstand** der zu a und b gehörigen Punkte auf der Zahlengeraden bezeichnet. Also gilt für $a, x \in \mathbb{R}, \varepsilon > 0$:

$$|a - x| \leq \varepsilon \Leftrightarrow a - \varepsilon \leq x \leq a + \varepsilon$$

1.2.6 (Intervalle) Seien $a, b \in \mathbb{R}$ mit $a < b$. Dann definiert man:

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R}; a \leq x \leq b\} && \text{abgeschlossenes Intervall,} \\ (a, b) &:= \{x \in \mathbb{R}; a < x < b\} && \text{offenes Intervall.} \end{aligned}$$

- analog: halboffene Intervalle $(a, b], [a, b)$
- Es wird abkürzend geschrieben

$$\begin{aligned} (-\infty, a] &:= \{x \in \mathbb{R}; x \leq a\} \\ (b, \infty) &:= \{x \in \mathbb{R}, x > b\} \text{ usw.,} \\ \mathbb{R} &= (-\infty, \infty). \end{aligned}$$

Hinweis: Die Symbole $\pm \infty$ dürfen nicht als Zahlen verstanden werden.

- Das offene Intervall $(a - \varepsilon, a + \varepsilon)$ heißt ε -**Umgebung** von a .

1.2.7 (vollständige Induktion) Aussagen oder Eigenschaften können von Elementen einer Menge A abhängen, z.B.

$$\forall n \in A : E(n) \quad (\text{wenn } A \subseteq \mathbb{N}).$$

Falls $A = \{n_0, n_0 + 1, \dots\} \subseteq \mathbb{Z}$, kann man die Richtigkeit solcher Aussagen mit vollständiger Induktion beweisen, d.h.

- 1) **Induktionsbeginn:** Man zeigt, dass $E(n)$ für $n = n_0$ gilt.
- 2) **Induktionsschritt:** Für beliebiges $n \geq n_0$ setzt man die Gültigkeit von $E(n)$ voraus (**Induktionsvoraussetzung**) und leitet daraus die Gültigkeit von $E(n+1)$ her. Dann gilt $E(k)$ für alle $k \geq n_0$.

- **Bernoulli–Ungleichung** (Jacob B., 1654–1705):

$$\forall h \in [-1, \infty), \forall n \in \mathbb{N}: (1+h)^n \geq 1+nh \quad (5)$$

Auf der vollständigen Induktion beruht auch das Verfahren der **Definition durch Rekursion**.

- Sei $a \in \mathbb{R} \setminus \{0\}$. Die **Potenzen** a^n , für $n \in \mathbb{N}_0$ sind definiert durch:

1. $a^0 := 1$
2. $a^{n+1} := a^n \cdot a$

- Sei $n \in \mathbb{N}_0$. Dann ist **Fakultät** $n!$ definiert als:

1. $0! := 1$
2. $(n+1)! := (n+1)n!$

- **Summen- und Produktzeichen** Seien $a_i \in \mathbb{R}$, $m \leq n \in \mathbb{N}_0$, $i = m, \dots, n$. Dann definieren wir:

$$\sum_{i=m}^n a_i := a_m + a_{m+1} + \dots + a_n$$

$$\prod_{i=m}^n a_i := a_m \cdot a_{m+1} \cdot \dots \cdot a_n$$

- Es besteht die Ungleichung:

$$\left| \sum_{i=m}^n a_i \right| \leq \sum_{i=m}^n |a_i|. \quad (6)$$

- Für die endliche **geometrische Reihe** gilt:

$$\sum_{i=0}^n q^i = \frac{1-q^{n+1}}{1-q} \quad \text{falls } q \neq 1. \quad (7)$$

Definition 1.2.8 Für ganze Zahlen n, k mit $0 \leq k \leq n$ ist der **Binomialkoeffizient** $\binom{n}{k}$ definiert durch

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

- Nun dreht man die x -Achse gegen den Uhrzeigersinn um 90° und erhält die **y -Achse**.
- Für einen beliebigen Punkt $P_0 \in E$ fällt man das Lot auf die x - und die y -Achse und erhält die **x -Koordinate** x_0 und die **y -Koordinate** y_0 von P_0 . Man schreibt

$$P_0 = (x_0, y_0).$$

Der Punkt $0 = (0, 0)$ heißt **Ursprung** oder **Nullpunkt**.

Durch dieses Vorgehen haben wir eine bijektive Zuordnung von Punkten $P \in E$ und Zahlenpaaren $(x, y) \in \mathbb{R}^2$ erhalten. Man kann also Teilmengen im \mathbb{R}^2 als Punktmengen in E veranschaulichen und Gebiete in E mit Hilfe von Gleichungen beschreiben.

- Sei $I \subseteq \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine Funktion. Die Menge

$$G_f := \{(x, y); x \in I, y = f(x)\}$$

heißt **Graph der Funktion** f .

- Sei $C \subseteq E$ eine Kurve in E , versehen mit einem Kartesischen Koordinatensystem. Sei $X = (x, y)$ ein Punkt auf der Kurve C . Falls man genau für die Punkte auf C eine Abhängigkeit $F(x, y) = 0$ aufstellen kann, d.h.

$$C = \{(x, y); F(x, y) = 0\},$$

dann heißt $F(x, y) = 0$ die **Gleichung der Kurve** C und C heißt die **Lösungsmenge** der Gleichung $F(x, y) = 0$.

- Bsp.: $x^2 + y^2 - R^2 = 0$,
Gleichung für den Kreis vom Radius $R > 0$ mit dem Ursprung als Mittelpunkt.

1.3.1 (Winkel) Ein Winkel α entsteht durch Drehung eines Zeigers um einen Punkt der Ebene.

- Die Länge des zugehörigen Einheitskreisbogens sei ℓ . Wir nennen ℓ das **Bogenmaß von α** , wenn die Drehung in positiver Richtung (gegen Uhrzeigersinn) erfolgte. Falls die Drehung im Uhrzeigersinn erfolgte, ist $-\ell$ das Bogenmaß.
- Ein Winkel α habe das Gradmaß $(\alpha)^\circ$. Das Bogenmaß ℓ des Winkels α errechnet sich durch:

$$\ell = \frac{\pi}{180^\circ} (\alpha)^\circ$$

- Im Weiteren werden wir immer mit dem Bogenmaß arbeiten.

1.3.2 (Sinus, Cosinus) Auf dem Einheitskreis drehe man einen Zeiger der Länge 1 um den Winkel α . Dadurch zeigt der Zeiger auf den Punkt P . Die Koordinaten von P werden mit $\cos \alpha$ und $\sin \alpha$ bezeichnet. Da der Winkel $\alpha \in \mathbb{R}$ beliebig ist, erhalten wir die Funktionen:

$$\cos : \mathbb{R} \rightarrow [-1, 1] : \alpha \mapsto \cos \alpha$$

Cosinusfunktion,

$$\sin : \mathbb{R} \rightarrow [-1, 1] : \alpha \mapsto \sin \alpha$$

Sinusfunktion .

Der Zeiger habe nun die Länge $r > 0$. Nach einer Drehung um den Winkel α zeigt er auf den Punkt $Q = (a, b)$. Dieser hat die Koordinaten

$$\begin{aligned} a &= r \cos \alpha, \\ b &= r \sin \alpha. \end{aligned} \tag{1}$$

Diese Gleichungen kann man auch schreiben als

$$\begin{aligned} \cos \alpha &= \frac{a}{r}, \\ \sin \alpha &= \frac{b}{r}. \end{aligned} \tag{2}$$

Es liege ein Dreieck vor mit den Seiten a, b, c . α sei der von b und c eingeschlossene Winkel, ebenso β von c und a , γ von a und b .

1.3.3 (Cosinussatz) Es gilt:

$$a^2 = b^2 + c^2 - 2bc \cos \alpha.$$

1.3.4 (Sinussatz) Es gilt:

$$\frac{\sin \alpha}{a} = \frac{\sin \beta}{b} = \frac{\sin \gamma}{c}.$$

• **Drehung des Koordinatensystems:** Sei (x, y) ein Kartesisches Koordinatensystem. Das Koordinatensystem (x', y') entstehe aus (x, y) durch Drehung um den Ursprung um den Winkel α .

Satz 1.3.5 Hat ein Punkt $X \in E$ im ursprünglichen System die Koordinaten (x_0, y_0) und im gedrehten System die Koordinaten (x'_0, y'_0) so gelten die **Transformationsformeln**:

$$\begin{aligned} x_0 &= x'_0 \cos \alpha - y'_0 \sin \alpha, \\ y_0 &= x'_0 \sin \alpha + y'_0 \cos \alpha, \end{aligned} \tag{3}$$

$$\begin{aligned} x'_0 &= x_0 \cos \alpha + y_0 \sin \alpha, \\ y'_0 &= -x_0 \sin \alpha + y_0 \cos \alpha. \end{aligned} \tag{4}$$

Satz 1.3.6 Sei $d : E \rightarrow E : (x, y) \mapsto (x', y')$ eine Drehung der Ebene um den Winkel α um den Ursprung. Dann gilt:

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

1.4 Der Raum

Ähnlich wie im Abschnitt 1.3 kann man den \mathbb{R}^3 und den Anschauungsraum R miteinander identifizieren.

- Ein **Kartesisches Koordinatensystem** im Raum besteht aus drei sich in einem Punkt 0 rechtwinklig schneidenden Zahlengeraden, die ein Rechtssystem bilden. Man nennt sie die x -, y -, z -Achsen. („Rechtssystem“ bedeutet: In der xy -Ebene entsteht die y -Achse durch Drehung der x -Achse um 90° entgegen dem Uhrzeigersinn. Ebenso bei der (x, z) - und der (y, z) -Ebene).
- Die **Koordinatenebenen** sind die durch zwei Achsen aufgespannten Ebenen. Man nennt sie auch die (x, y) -, (x, z) - und (y, z) -Ebenen.

1.4.1 (Vektoren) Seien P, Q Punkte im Raum R . Es gibt genau eine Parallelverschiebung des Raumes, die P auf Q abbildet. Diese wird mit \overrightarrow{PQ} bezeichnet und heißt „**Vektor von P nach Q**“.

- Unter \overrightarrow{PQ} wird z.B. der Punkt S auf T abgebildet, d.h.

$$\overrightarrow{PQ} = \overrightarrow{ST}.$$

Also stellen zwei gleichlange, gleichgerichtete „Pfeile“ denselben Vektor dar.

- Man schreibt oft kürzer $\vec{a} = \overrightarrow{PQ}$. Die Parallelverschiebung, die alle Punkte fest läßt, d.h. durch \overrightarrow{PP} beschrieben ist, heißt $\vec{0} = \overrightarrow{PP} = \text{Nullvektor}$.

Definition 1.4.2 Seien $\vec{a} = \overrightarrow{PQ}$ und $\vec{b} = \overrightarrow{QR}$ Vektoren. Die zu \overrightarrow{PQ} umgekehrte Parallelverschiebung \overrightarrow{QP} bezeichnen wir mit $-\vec{a}$. Die Parallelverschiebung, die durch nacheinander Ausführen von \overrightarrow{PQ} und \overrightarrow{QR} entsteht, bezeichnen wir mit $\vec{a} + \vec{b}$ und nennen sie **Summe der Vektoren \vec{a} und \vec{b}** .

- Es gelten folgende Rechenregeln:

$$\begin{aligned} \vec{a} + \vec{0} &= \vec{a}, \\ \vec{a} + (-\vec{a}) &= \vec{0}, \\ \vec{a} + \vec{b} &= \vec{b} + \vec{a}, \\ \vec{a} + (\vec{b} + \vec{c}) &= (\vec{a} + \vec{b}) + \vec{c}. \end{aligned} \tag{1}$$

- Die Differenz zweier Vektoren wird erklärt durch:

$$\vec{a} - \vec{b} := \vec{a} + (-\vec{b}) \tag{2}$$

Definition 1.4.3 Die **Länge** bzw. die **Norm** eines Vektors $\vec{a} = \overrightarrow{PQ}$ ist die Länge der Strecke \overline{PQ} . Man schreibt dafür $|\vec{a}|$ bzw. $\|\vec{a}\|$.

- Für den **Nullvektor** $\vec{0} := \overrightarrow{PP}$ setzen wir $\|\vec{0}\| := 0$.

Definition 1.4.4 Sei $\alpha \in \mathbb{R}_0^+ := \{x \in \mathbb{R}, x \geq 0\}$ und \vec{a} ein Vektor. Dann bezeichnet man mit $\alpha \vec{a}$ denjenigen Vektor, der dieselbe Richtung wie \vec{a} hat, aber die α -fache Länge. Man nennt $\alpha \vec{a}$ das **α -fache skalare Vielfache** von \vec{a} . Sei $\alpha < 0$, dann setzen wir

$$\alpha \vec{a} := -(|\alpha| \vec{a})$$

- Für alle Vektoren \vec{a}, \vec{b} und alle $\alpha, \beta \in \mathbb{R}$ gelten folgende Rechenregeln:

$$\begin{aligned} \alpha(\beta \vec{a}) &= (\alpha\beta) \vec{a} \\ \alpha(\vec{a} + \vec{b}) &= \alpha \vec{a} + \alpha \vec{b} \\ (\alpha + \beta) \vec{a} &= \alpha \vec{a} + \beta \vec{a} \end{aligned} \quad (3)$$

$$\begin{aligned} \|\alpha \vec{a}\| &= |\alpha| \|\vec{a}\| \\ \|\vec{a} + \vec{b}\| &\leq \|\vec{a}\| + \|\vec{b}\| \end{aligned} \quad (4)$$

- Ein Vektor mit Norm 1 heißt **Einheitsvektor**. Sei $\vec{a} \neq \vec{0}$ ein Vektor, dann ist

$$\vec{a}_0 = \frac{\vec{a}}{\|\vec{a}\|}$$

ein Einheitsvektor.

Definition 1.4.5 Der Raum sei mit einem kartesischen Koordinatensystem versehen. Die drei Einheitsvektoren in positiver x -, y - und z -Richtung werden mit \vec{e}_1, \vec{e}_2 und \vec{e}_3 bezeichnet. Wir nennen $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine **kartesische Basis**.

- Sei $A = (a_1, a_2, a_3)$ ein Punkt im Raum. Der Vektor $\vec{a} = \overrightarrow{OA}$ heißt **Ortsvektor** des Punktes A .
- \vec{a} ist eindeutig zerlegbar als Summe:

$$\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3.$$

- Man nennt $a_i, i = 1, 2, 3$, die **Koordinaten** des Vektors \vec{a} und $a_i \vec{e}_i, i = 1, 2, 3$, die **Komponenten von \vec{a} in Richtung \vec{e}_i** .
- Für ein festes kartesisches Koordinatensystem schreibt man abkürzend:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (a_1, a_2, a_3)^\top \Leftrightarrow \vec{a} = \overrightarrow{OA} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3 \quad (5)$$

falls $A = (a_1, a_2, a_3)$.

- Im allgemeinen Fall $\vec{a} = \overrightarrow{PQ}$ mit $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$ haben wir die **Koordinatendarstellung**

$$\vec{a} = \overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}. \quad (6)$$

- Aus (5), (6) erhält man die Koordinatendarstellung für Summen und skalare Vielfache. Sei $\vec{a} = \sum_{i=1}^3 a_i \vec{e}_i$, $\vec{b} = \sum_{i=1}^3 b_i \vec{e}_i$ und $\alpha \in \mathbb{R}$, dann gilt:

$$\begin{aligned} \vec{a} + \vec{b} &= \sum_{i=1}^3 (a_i + b_i) \vec{e}_i, \\ \alpha \vec{a} &= \sum_{i=1}^3 (\alpha a_i) \vec{e}_i. \end{aligned} \quad (7)$$

- Aus (5) und dem Satz des Pythagoras ergibt sich für $\vec{a} = \sum_{i=1}^3 a_i \vec{e}_i$:

$$\|\vec{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2} \quad (8)$$

1.5 Produkte

1.5.1 (Winkel zwischen Vektoren) Seien \vec{a}, \vec{b} von $\vec{0}$ verschiedene Vektoren und sei P ein beliebiger Punkt in Raum, an dem man beide Vektoren abträgt. Dann ist der kleinere der positiv gemessenen Winkel, den die „Pfeile“ \vec{a} und \vec{b} im Scheitel P bilden, der **Winkel zwischen \vec{a} und \vec{b}** . Man schreibt $\angle(\vec{a}, \vec{b})$. Der so definierte Winkel hat die folgenden Eigenschaften:

$$\begin{aligned} 0 &\leq \angle(\vec{a}, \vec{b}) \leq \pi, \\ \angle(\vec{a}, \vec{b}) &= \angle(\vec{b}, \vec{a}), \\ \angle(\vec{a}, t\vec{a}) &= 0 && \text{falls } t > 0, \\ \angle(\vec{a}, t\vec{a}) &= \pi && \text{falls } t < 0, \\ \angle(-\vec{a}, \vec{b}) &= \pi - \angle(\vec{a}, \vec{b}). \end{aligned} \quad (1)$$

Definition 1.5.2 Man nennt \vec{a} **orthogonal** zu \vec{b} , wenn $\angle(\vec{a}, \vec{b}) = \frac{\pi}{2}$, und schreibt $\vec{a} \perp \vec{b}$. Es besteht die folgende Konvention: Für alle Vektoren \vec{a} gilt $\vec{a} \perp \vec{0}$. \vec{a} und \vec{b} heißen **parallel**, wenn $\vec{a}, \vec{b} \neq \vec{0}$ und $\vec{b} = t\vec{a}$ mit einem $t \in \mathbb{R}$.

Definition 1.5.3 Das **Skalarprodukt** $\vec{a} \cdot \vec{b}$ zweier Vektoren \vec{a} und \vec{b} ist definiert durch

$$\vec{a} \cdot \vec{b} := \begin{cases} \|\vec{a}\| \|\vec{b}\| \cos \angle(\vec{a}, \vec{b}) & \text{falls } \vec{a} \neq \vec{0} \wedge \vec{b} \neq \vec{0}, \\ 0 & \text{falls } \vec{a} = \vec{0} \vee \vec{b} = \vec{0}. \end{cases}$$

- Es bestehen die Gleichungen

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij} \quad i, j = 1, 2, 3, \quad (2)$$

wobei das **Kronecker Symbol** δ_{ij} definiert ist als

$$\delta_{ij} := \begin{cases} 0 & \text{falls } i \neq j, \\ 1 & \text{falls } i = j. \end{cases}$$

(Leopold Kronecker, 1821–1891)

- Sei $\vec{a} = \sum_{i=1}^3 a_i \vec{e}_i$, dann gilt:

$$a_i = \|\vec{a}\| \cos \angle(\vec{a}, \vec{e}_i) = \vec{a} \cdot \vec{e}_i,$$

und demzufolge

$$\vec{a} = \sum_{i=1}^3 (\vec{a} \cdot \vec{e}_i) \vec{e}_i. \quad (3)$$

- Es gelten folgende Rechenregeln:

$$\begin{aligned} a) & \quad \vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}, \\ b) & \quad (\alpha \vec{a}) \cdot \vec{b} = \alpha (\vec{a} \cdot \vec{b}), \\ c) & \quad (\vec{a} + \vec{b}) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}, \\ d) & \quad \vec{a} \cdot \vec{b} = 0 \Leftrightarrow \vec{a} \perp \vec{b}, \\ e) & \quad \|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}}. \end{aligned} \quad (4)$$

- Aus (4) b) & c) folgt:

$$\begin{aligned} & (\alpha_1 \vec{u}_1 + \dots + \alpha_n \vec{u}_n) \cdot (\beta_1 \vec{v}_1 + \dots + \beta_m \vec{v}_m) \\ &= \alpha_1 \beta_1 \vec{u}_1 \cdot \vec{v}_1 + \dots + \alpha_n \beta_m \vec{u}_n \cdot \vec{v}_m \\ &= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \vec{u}_i \cdot \vec{v}_j. \end{aligned} \quad (5)$$

- Aus (5), (2) und den Koordinatendarstellungen $\vec{a} = \sum_{i=1}^3 a_i \vec{e}_i$, $\vec{b} = \sum_{i=1}^3 b_i \vec{e}_i$ erhält man:

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \sum_{i=1}^3 a_i b_i, \\ \|\vec{a}\| &= \sqrt{a_1^2 + a_2^2 + a_3^2}. \end{aligned} \quad (6)$$

$$\cos \angle(\vec{a}, \vec{b}) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|} \quad \text{falls } \vec{a}, \vec{b} \neq \vec{0}, \quad (7)$$

$$\cos \angle(\vec{a}, \vec{e}_i) = \frac{a_i}{\sqrt{a_1^2 + a_2^2 + a_3^2}} \quad \text{falls } \vec{a} \neq \vec{0}. \quad (8)$$

Satz 1.5.4 Sei $\vec{b} \neq \vec{0}$ und \vec{a} ein beliebiger Vektor. Dann besitzt \vec{a} eine **orthogonale Zerlegung** längs \vec{b} , die gegeben ist durch

$$\vec{a} = \vec{a}_{\vec{n}} + \vec{a}_{\vec{b}}, \quad (9)$$

wobei

$$\begin{aligned} \vec{a}_{\vec{b}} &:= \frac{\vec{a} \cdot \vec{b}}{\|\vec{b}\|^2} \vec{b}, \\ \vec{a}_{\vec{n}} &:= \vec{a} - \vec{a}_{\vec{b}}. \end{aligned} \quad (10)$$

Hierbei zeigt $\vec{a}_{\vec{b}}$ in Richtung \vec{b} und $\vec{a}_{\vec{n}}$ steht senkrecht auf \vec{b} . $\vec{a}_{\vec{b}}$ heißt **Projektion von \vec{a} auf \vec{b}** .

1.5.5 (Vektorprodukt) Seien \vec{a}, \vec{b} zwei Vektoren. Das **Vektorprodukt** $\vec{a} \times \vec{b}$ ist der Vektor mit den Eigenschaften:

1) $\vec{a} \times \vec{b} := \vec{0}$ falls $\vec{a} = \vec{0}$ oder $\vec{b} = \vec{0}$ oder \vec{a} parallel zu \vec{b} ist,

2) sonst ist $\vec{a} \times \vec{b}$ derjenige Vektor,

a) der senkrecht auf \vec{a} und \vec{b} steht,

b) mit dem $(\vec{a}, \vec{b}, \vec{a} \times \vec{b})$ ein Rechtssystem bildet,

c) dessen Betrag gleich dem Flächeninhalt des von \vec{a}, \vec{b} aufgespannten Parallelogramms ist.

• Für die kartesischen Basisvektoren gilt:

$$\begin{aligned} \vec{e}_1 \times \vec{e}_1 &= \vec{e}_2 \times \vec{e}_2 = \vec{e}_3 \times \vec{e}_3 = \vec{0}, \\ \vec{e}_1 \times \vec{e}_2 &= \vec{e}_3 = -(\vec{e}_2 \times \vec{e}_1), \\ \vec{e}_2 \times \vec{e}_3 &= \vec{e}_1 = -(\vec{e}_3 \times \vec{e}_2), \\ \vec{e}_3 \times \vec{e}_1 &= \vec{e}_2 = -(\vec{e}_1 \times \vec{e}_3). \end{aligned} \quad (11)$$

• Es gelten folgende Rechenregeln:

$$\vec{a} \times \vec{a} = \vec{0}, \quad (12)$$

$$\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a}),$$

$$\alpha (\vec{a} \times \vec{b}) = (\alpha \vec{a}) \times \vec{b} = \vec{a} \times (\alpha \vec{b}) \quad (\alpha \in \mathbb{R}),$$

$$\vec{a} \times (\vec{b} + \vec{c}) = (\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c}), \quad (13)$$

$$(\vec{a} + \vec{b}) \times \vec{c} = (\vec{a} \times \vec{c}) + (\vec{b} \times \vec{c}),$$

$$\|\vec{a} \times \vec{b}\|^2 = \|\vec{a}\|^2 \|\vec{b}\|^2 - (\vec{a} \cdot \vec{b})^2. \quad (14)$$

- Aus (13) und (14) folgt:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \quad (15)$$

- (15) impliziert:

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}. \quad (16)$$

- Aus (16), (9) und (10) folgt, dass die zu \vec{b} orthogonale Komponente gegeben ist durch:

$$\vec{a}_{\vec{b}} = \frac{1}{\|\vec{b}\|^2} \vec{b} \times (\vec{a} \times \vec{b}). \quad (17)$$

Definition 1.5.6 Seien \vec{a} , \vec{b} und \vec{c} Vektoren. Das **Spatprodukt** dieser Vektoren ist definiert durch:

$$[\vec{a}, \vec{b}, \vec{c}] := \vec{a} \cdot (\vec{b} \times \vec{c}).$$

Satz 1.5.7 Der von den Vektoren \vec{a} , \vec{b} und \vec{c} aufgespannte Parallelepid (Spat) hat das Volumen

$$V = |[\vec{a}, \vec{b}, \vec{c}]|.$$

- Seien $\vec{a} = \sum_{i=1}^3 a_i \vec{e}_i$, $\vec{b} = \sum_{i=1}^3 b_i \vec{e}_i$ und $\vec{c} = \sum_{i=1}^3 c_i \vec{e}_i$. Dann folgt aus (15) und (6)

$$\begin{aligned} [\vec{a}, \vec{b}, \vec{c}] &= a_1(b_2 c_3 - b_3 c_2) + a_2(b_3 c_1 - b_1 c_3) + a_3(b_1 c_2 - b_2 c_1) \\ &= \det(\vec{a}, \vec{b}, \vec{c}). \quad (\text{siehe Kapitel 6}) \end{aligned} \quad (18)$$

1.6 Komplexe Zahlen

Bisher haben wir Punkte z der Ebene, die mit einem kartesischen Koordinatensystem versehen ist, als Zahlenpaare $z = (x, y) \in \mathbb{R}^2$ aufgefasst. Jetzt schreiben wir den Punkt $z = (x, y)$ als

$$z := x + i y \quad (1)$$

und nennen dies eine **komplexe Zahl** mit **Realteil** $\operatorname{Re} z := x$ und **Imaginärteil** $\operatorname{Im} z := y$. Die x -Achse heisst **reelle Achse** und die y -Achse **imaginäre Achse**. Die Menge der komplexen Zahlen wird mit

$$\mathbb{C} := \{x + i y; x, y \in \mathbb{R}\} \quad (2)$$

bezeichnet. Für zwei komplexe Zahlen $z = x + i y$, $w = u + i v$ mit $x, y, u, v \in \mathbb{R}$ gilt:

$$z = w \iff (x = u) \wedge (y = v). \quad (3)$$

1.6.1 (Grundrechenarten in \mathbb{C}) Seien $z = x + i y$, $w = u + i v$ mit $x, y, u, v \in \mathbb{R}$ zwei komplexe Zahlen.

- Die **Summe** und die **Differenz** von z, w ist definiert als:

$$\begin{aligned} z + w &:= (x + u) + i(y + v), \\ z - w &:= (x - u) + i(y - v). \end{aligned} \quad (4)$$

- Die Zahl ix geht aus x durch Drehung um $\pi/2$ hervor. Die **Multiplikation** verallgemeinert dies. Wir definieren:

$$zw := (xu - yv) + i(xv + yu). \quad (5)$$

Insbesondere gilt also:

$$i^2 = -1. \quad (6)$$

- Die **Potenzen** z^n sind rekursiv definiert durch:

$$z^0 := 1, \quad z^n := zz^{n-1}, \quad n \in \mathbb{N}. \quad (7)$$

- Sei $w \neq 0$. Dann ist die **Division** definiert als:

$$\frac{z}{w} = \frac{x + iy}{u + iv} := \frac{xu + yv}{u^2 + v^2} + i \frac{yu - xv}{u^2 + v^2}. \quad (8)$$

- \mathbb{C} , versehen mit der Addition und Multiplikation, erfüllt die Körperaxiome. Fasst man die reellen Zahlen x als spezielle komplexe Zahlen $x + i \cdot 0$ auf, dann erweist sich \mathbb{C} als Erweiterungskörper von \mathbb{R} .

Warnung: Die Anordnung $<$ kann nicht auf \mathbb{C} fortgesetzt werden. In \mathbb{C} gelten Ungleichungen nur in Bezug auf den Betrag (s.u.)

Definition 1.6.2 Sei $z = x + iy$. Die zu z **konjugierte komplexe Zahl** \bar{z} ist definiert durch:

$$\bar{z} := x - iy.$$

- Wir haben folgende Rechenregeln für $z, w \in \mathbb{C}$:

$$\begin{aligned} \overline{z + w} &= \bar{z} + \bar{w}, \\ \overline{z\bar{w}} &= \bar{z}w, \\ \overline{\left(\frac{z}{w}\right)} &= \frac{\bar{z}}{\bar{w}}, \quad \text{falls } w \neq 0, \end{aligned} \quad (9)$$

$$\begin{aligned} \overline{\bar{z}} &= z, \\ \operatorname{Re} z &= \frac{1}{2}(z + \bar{z}), \\ \operatorname{Im} z &= \frac{1}{2i}(z - \bar{z}). \end{aligned} \quad (10)$$

- Der **Betrag** $|z|$ von $z = x + iy$ ist definiert als:

$$|z| := \sqrt{x^2 + y^2}. \quad (11)$$

Es gelten die Rechenregeln:

$$\begin{aligned} |z| &= \sqrt{z\bar{z}}, \\ |zw| &= |z| |w|, \\ \left| \frac{z}{w} \right| &= \frac{|z|}{|w|}, \quad \text{falls } w \neq 0, \\ |z| &= |\bar{z}|. \end{aligned} \quad (12)$$

Die Dreiecksungleichung

$$|z + w| \leq |z| + |w| \quad (13)$$

kann durch Induktion auf n Summanden verallgemeinert werden

$$\left| \sum_{j=1}^n z_j \right| \leq \sum_{j=1}^n |z_j|. \quad (14)$$

- Zur Darstellung komplexer Zahlen mittels Polarkoordinaten s. 2.3.3.

1.6.3 (Quadratische Gleichungen) Die Gleichung $x^2 + 1 = 0$ hat in \mathbb{R} keine Lösung. Allerdings sind $x_{1,2} = \pm i$ Lösungen in \mathbb{C} . Allgemeiner sei

$$ax^2 + bx + c = 0,$$

mit $a, b, c \in \mathbb{R}$, $a \neq 0$ eine quadratische Gleichung. Dann sind die Lösungen gegeben durch:

$$x_{1,2} = -\frac{b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac}, \quad (15)$$

wobei $d := b^2 - 4ac$ die **Diskriminante** genannt wird. Im Falle $d < 0$ setzen wir $\sqrt{d} := i\sqrt{-d}$.

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

In Definition (1.1.4) im Kapitel 1 wurden Funktionen für allgemeine Mengen definiert. Nun betrachten wir den Spezialfall einer **reellen Funktion einer Veränderlichen**

$$f : D \subseteq \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x). \quad (1)$$

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 2.1.1 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.

Definition 2.1.2 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 oder streng monoton fallend** (bzw. **strikt oder streng monoton wachsend**), wenn für alle $x_1 < x_2 \in D$ die Ungleichung $f(x_1) > f(x_2)$ (bzw. $f(x_1) < f(x_2)$) besteht.

2.1.3 (Rechnen mit Funktionen)

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

$$\begin{aligned} (f \pm g)(x) &:= f(x) \pm g(x) \\ (f \cdot g)(x) &:= f(x) \cdot g(x) \\ \left(\frac{f}{g}\right)(x) &:= \frac{f(x)}{g(x)}, \quad \text{falls } g(x) \neq 0 \end{aligned} \tag{2}$$

d.h. die Operationen werden punktweise ausgeführt.

- 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:

$$(f \circ g)(x) := f(g(x)). \tag{3}$$

2.2 Polynome und rationale Funktionen

Definition 2.2.1 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

$$f(x) = a_0 + a_1x + \dots + a_nx^n = \sum_{j=0}^n a_jx^j. \tag{1}$$

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

Satz 2.2.2 Zwei Polynome sind genau dann gleich, wenn ihre Koeffizienten paarweise übereinstimmen, d.h.

$$\sum_{j=1}^n a_jx^j = \sum_{j=1}^n b_jx^j \quad \Leftrightarrow \quad a_j = b_j, \quad j = 1, \dots, n.$$

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

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

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

Satz 2.2.4 Sei $f(x) = \sum_{j=0}^n a_j x^j$, $a_n \neq 0$ 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

$$\begin{aligned} f(b) &= c_0 \\ f(x) &= (x - b) \sum_{j=1}^n c_j x^{j-1} + c_0, \end{aligned} \quad (3)$$

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

Definition 2.2.5 Als **Nullstelle** einer Funktion $f : D \rightarrow \mathbb{R}$ bezeichnet man jedes $x \in D$ mit $f(x) = 0$.

- Für Polynome erhält man aus (3):

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

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

$$f(x) = (x - b)^k g(x) \quad \text{und} \quad g(b) \neq 0. \quad (4)$$

Satz 2.2.7

a) Sei $f(x) = \sum_{j=0}^n a_j x^j$ 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

$$f(x) = \prod_{j=1}^r (x - b_j)^{\ell_j} q(x) \quad (5)$$

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

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

2.2.8 (Komplexe Polynome) Betrachten wir komplexe Polynome:

$$f(z) = \sum_{j=0}^n a_j z^j, \quad a_j \in \mathbb{C}. \quad (6)$$

- 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 $\pm i$.

Satz 2.2.9 (Fundamentalsatz der Algebra) ,

(Carl Friedrich Gauß, 1777–1855, 1801). Zu jedem Polynom der Form (6), mit $n \geq 1$ und $a_n \neq 0$, gibt es eine Zahl $w \in \mathbb{C}$ mit $f(w) = 0$.

Satz 2.2.10 Jedes komplexe Polynom der Form (6) mit $n \geq 1, a_n \neq 0$ besitzt eine **Faktorisierung über \mathbb{C}** der Form

$$f(z) = a_n (z - w_1)^{\ell_1} \cdots (z - w_k)^{\ell_k} \quad (7)$$

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

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

Lemma 2.2.11 Sei $w \in \mathbb{C}$ eine Nullstelle eines Polynoms f mit reellen Koeffizienten. Dann ist auch \bar{w} eine Nullstelle von f .

Satz 2.2.12 Jedes reelle Polynom $f(x) = \sum_{j=0}^n a_j x^j, n \geq 1, a_j \in \mathbb{R}, a_n \neq 0$, besitzt die **Faktorisierung über \mathbb{R}**

$$f(x) = a_n (x - b_1)^{\ell_1} \cdots (x - b_r)^{\ell_r} (x^2 + c_1 x + d_1)^{k_1} \cdots (x^2 + c_s x + d_s)^{k_s},$$

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

Satz 2.2.13 Die rationalen Nullstellen eines Polynoms

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

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.

Satz 2.2.14 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 mit Grad kleiner gleich n , so dass $p_n(x_i) = y_i$, $i = 0, \dots, n$.

• **Newton Interpolationsverfahren** (Isaac N., 1642–1727)

Man suche das Polynom p_n aus Satz 2.2.14 in der Form

$$p_n(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \dots + \alpha_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}) \quad (8)$$

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

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

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

$$\begin{array}{ccccccc} x_0 & y_0 & & & & & \\ & \diagdown & & & & & \\ x_1 & y_1 & y_{0,1} & & & & \\ & \diagup & \diagdown & & & & \\ x_2 & y_2 & y_{1,2} & y_{0,1,2} & & & \\ & & \vdots & \vdots & \cdots & & \\ & & \vdots & y_{n-2,n-1,n} & & & \\ & \diagdown & & & & & \\ \vdots & \vdots & y_{n-1,n} & & & & \\ x_n & y_n & & & & & \end{array}$$

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 der oberen Schrägzeile.

Definition 2.2.15 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**.

Satz 2.2.16 Jede rationale Funktion $\frac{p(x)}{q(x)}$ mit Zählergrad \geq Nennergrad lässt sich darstellen als

$$\frac{p(x)}{q(x)} = h(x) + \frac{r(x)}{q(x)} \quad (9)$$

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

Definition 2.2.17 Seien p, d Polynome. Man nennt d eine **Teiler** von p , wenn es ein Polynom p_0 gibt, so dass $p(x) = d(x)p_0(x)$.

Lemma 2.2.18 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 $\text{Grad } r < \text{Grad } q$. Es 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

$$D(f) = \{x \in \mathbb{R} : q(x) \neq 0\}. \quad (10)$$

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

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

Dann nennt man b einen **ℓ -fachen Pol** von f .

2.3 Trigonometrische Funktionen

In Paragraph 1.3 Punkt 1.3.2 haben wir die Sinus- und die Cosinusfunktion definiert. Aus der Definition erhält man sofort folgende Eigenschaften:

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

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

$$\begin{aligned} \cos x = 0 &\Leftrightarrow x = \pm \frac{1}{2}\pi, \pm \frac{3}{2}\pi, \pm \frac{5}{2}\pi, \dots, \\ \sin x = 0 &\Leftrightarrow x = 0, \pm\pi, \pm 2\pi, \dots \end{aligned} \quad (3)$$

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

$$\begin{aligned}\cos(x + y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x + y) &= \sin x \cos y + \cos x \sin y.\end{aligned}\tag{4}$$

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

$$\begin{aligned}\sin\left(x + \frac{\pi}{2}\right) &= \cos x, \\ \cos\left(x - \frac{\pi}{2}\right) &= \sin x,\end{aligned}\tag{5}$$

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

- In Formelsammlungen gibt es zahlreiche Identitäten, die sich am einfachsten aus der Euler-Formel (2.3.3) herleiten lassen. Hier einige Beispiele:

$$\begin{aligned}\cos(x - y) &= \cos x \cos y + \sin x \sin y, \\ \sin(x - y) &= \sin x \cos y - \cos x \sin y,\end{aligned}\tag{6}$$

$$\begin{aligned}\sin x + \sin y &= 2 \sin \frac{x+y}{2} \cos \frac{x-y}{2}, \\ \cos x + \cos y &= 2 \cos \frac{x+y}{2} \cos \frac{x-y}{2},\end{aligned}\tag{7}$$

$$\begin{aligned}\cos(2x) &= \cos^2 x - \sin^2 x = 2 \cos^2 x - 1, \\ \sin 2x &= 2 \sin x \cos x,\end{aligned}\tag{8}$$

$$\begin{aligned}1 + \cos x &= 2 \cos^2 \frac{x}{2}, \\ 1 - \cos x &= 2 \sin^2 \frac{x}{2},\end{aligned}\tag{9}$$

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

Definition 2.3.2 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:

$$\begin{aligned}\tan(-x) &= -\tan x, && \text{(ungerade Funktion),} \\ \cos(-x) &= \cos x, && \text{(gerade Funktion),}\end{aligned}\tag{10}$$

$$\begin{aligned}\tan(x + \pi) &= \tan x, && \text{(\pi-periodisch),} \\ \cos(x + \pi) &= -\cos x, && \text{(\pi-periodisch),}\end{aligned}\tag{11}$$

$$\tan(x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}, \quad \text{für „erlaubte“ } x, y.\tag{12}$$

2.3.3 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 $\varphi = \arg z$. Aber wir wissen, dass $x = r \cos \varphi$, $y = r \sin \varphi$ und somit erhalten wir

$$z = r(\cos \varphi + i \sin \varphi). \quad (13)$$

Diese Form der Darstellung von z heißt **Polardarstellung**.

- Die Exponentialfunktion e^z kann für $z \in \mathbb{C}$ durch die „Potenzreihe“ 5.3.7 definiert werden. Es zeigt sich

$$\begin{aligned} e^{i\varphi} &= \cos \varphi + i \sin \varphi, \quad \varphi \in \mathbb{R}. \\ e^{x+iy} &= e^x(\cos y + i \sin y), \quad x, y \in \mathbb{R} \end{aligned} \quad (14)$$

(Eulersche Formel, Leonhard E., 1707–1783). Die Exponentialfunktion genügt auf \mathbb{C} der Multiplikationsformel

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}.$$

2.3.4 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

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, \\ \varphi &= \begin{cases} \arccos \frac{x}{r} & \text{falls } y \geq 0, \\ -\arccos \frac{x}{r} & \text{falls } y < 0. \end{cases} \end{aligned} \quad (\text{s.3.3.4})$$

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 $z = x + iy$.

Satz 2.3.5 (Abraham de Moivre, 1667–1754). Für alle $\varphi, \psi \in \mathbb{R}$ gilt,

a) $e^{i\varphi}e^{i\psi} = e^{i(\psi+\varphi)},$

b) $(e^{i\varphi})^n = e^{in\varphi}, \quad n \in \mathbb{N},$

c) $\overline{e^{i\varphi}} = e^{-i\varphi} = \frac{1}{e^{i\varphi}}.$

• 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:

$$\begin{aligned} z \cdot w &= |z||w| e^{i(\varphi+\psi)}, \\ \frac{z}{w} &= \frac{|z|}{|w|} e^{i(\varphi-\psi)}, \quad \text{falls } w \neq 0. \end{aligned} \tag{15}$$

Definition 2.3.6 Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt **periodisch** mit einer **Periode** 2ℓ , wenn

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

Definition 2.3.7 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.

- 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

$$\begin{aligned} z(t) &:= A \cos(\omega t + \alpha) + iA \sin(\omega t + \alpha) \\ &= A e^{i(\omega t + \alpha)} \\ &= a e^{i\omega t}. \end{aligned} \tag{16}$$

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

Satz 2.3.8 *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

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

- Es gilt folgende Formel:

$$\sin \delta + \sin 2\delta + \cdots + \sin n\delta = \frac{\sin \frac{n+1}{2}\delta \sin \frac{n}{2}\delta}{\sin \frac{\delta}{2}} \quad (17)$$

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

$$a_1 e^{i\omega_1 t} + a_2 e^{i\omega_2 t} = \left(a_1 e^{i\frac{\omega_1 - \omega_2}{2} t} + a_2 e^{i\frac{\omega_2 - \omega_1}{2} t} \right) e^{i\frac{\omega_1 + \omega_2}{2} t} \quad (18)$$

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

- Im Allgemeinen ist eine modulierte Schwingung nicht periodisch. Ist der Quotient $\frac{\omega_1}{\omega_2} = \frac{n_1}{n_2}$ jedoch eine rationale Zahl, so ist die modulierte Schwingung periodisch. In diesem Fall besitzen die Schwingungen $z_1(t) = a_1 e^{i\omega_1 t}$ und $z_2(t) = a_2 e^{i\omega_2 t}$ die gemeinsame Periode $T := \frac{2\pi n_1}{\omega_1} = \frac{2\pi n_2}{\omega_2}$, d.h. $z_1(t) + z_2(t)$ ist periodisch, aber nicht notwendig harmonisch.

2.4 Zahlenfolgen und Grenzwerte

Einer der wichtigsten Begriffe der Analysis ist der des Grenzwerts. Wir fangen unsere Untersuchungen mit Folgen an und übertragen die Ergebnisse dann auf Funktionen.

Definition 2.4.1 *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) | $a_n = c$ | konstante Folge | c, c, \dots |
| 2) | $a_n = n$ | Folge der natürlichen Zahlen | $1, 2, 3, \dots$ |
| 3) | $a_n = a_0 + nd$ | arithmetische Folge | $a_0, a_0 + d, a_0 + 2d, \dots$ |
| 4) | $a_n = a_0 q^n$ | geometrische Folge | $a_0, a_0 q, a_0 q^2, \dots$ |
| 5) | $a_0 = 1, a_{n+1} = (n+1)a_n$ | rekursiv definierte Folge | |

Definition 2.4.2 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.

Definition 2.4.3 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 beliebig 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**.
- Die Aussage $a_n \rightarrow a$ ist gleichbedeutend damit, dass die Folge $b_n = a_n - a$ Nullfolge ist.

Satz 2.4.4 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 $a = b$, d.h. im Fall der Konvergenz ist der Grenzwert oder Limes eindeutig bestimmt.
- 2) Die Folge $(a_n)_{n \geq 0}$ ist beschränkt.

Definition 2.4.5 Man sagt, eine Folge $(a_n)_{n \geq 0}$ **divergiert** 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 $-\infty$.

- Für eine Folge (a_n) mit $a_n > 0 \forall n$ sind äquivalent:
 - 1) $a_n \rightarrow 0$,

$$2) \frac{1}{a_n} \rightarrow \infty.$$

Lemma 2.4.6 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 $x \leq -1$.

2.4.7 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:

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

Es können folgende Fälle auftreten:

- a) $|x| < 1 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \left(\frac{1}{1-x} + \frac{x^{n+1}}{1-x} \right) = \frac{1}{1-x}$,
- b) $x \geq 1 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} s_n = \infty$,
- c) $x \leq -1 \quad \Rightarrow \quad \text{divergiert.}$

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

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

2.4.8 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:

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

Definition 2.4.9 Ist $(a_n)_{n \geq 0}$ eine Folge und $n_0 < n_1 < \cdots$ eine aufsteigende Indexfolge, dann heißt die Folge a_{n_0}, a_{n_1}, \dots **Teilfolge** der Folge $(a_n)_{n \geq 0}$.

2.5 Rechenregeln für Grenzwerte und Konvergenzkriterien

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

Satz 2.5.1 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:

$$a) \lim_{n \rightarrow \infty} (a_n \pm b_n) = a \pm b,$$

$$b) \lim_{n \rightarrow \infty} (a_n b_n) = a \cdot b,$$

c) 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: $\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{1}{a}, \lim_{n \rightarrow \infty} \frac{b_n}{a_n} = \frac{b}{a}$,

$$d) \lim_{n \rightarrow \infty} |a_n| = |a|,$$

$$e) \lim_{n \rightarrow \infty} \sqrt{a_n} = \sqrt{a}, \text{ falls alle } a_n \geq 0.$$

Satz 2.5.2 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 $\lim_{n \rightarrow \infty} a_n = c$.

Satz 2.5.3 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: $\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n$.

• Da Folgen spezielle Funktionen sind, wissen wir schon, was eine **monoton wachsende** (bzw. **monoton fallende**) Folge ist, nämlich wenn für alle $n \geq 0$ gilt $a_{n+1} \geq a_n$ (bzw. $a_{n+1} \leq a_n$).

Satz 2.5.4 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

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

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.

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n. \quad (2)$$

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

Definition 2.5.5 Die **Exponentialfunktion** $e : x \mapsto e^x$ ist für alle $x \in \mathbb{R}$ definiert durch

$$e^x := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (3)$$

Beide Grenzwerte existieren auch für $x \in \mathbb{C}$ und sind gleich.

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 $f(x)$, wenn x sich a nähert, wobei $x \neq a$.

Definition 2.6.1 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 $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = c$.

Satz 2.6.2 Aus $\lim_{x \rightarrow a} f(x) = c$, $\lim_{x \rightarrow a} g(x) = d$, mit $c, d \in \mathbb{R}$ folgt:

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

Satz 2.6.3 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. $\lim_{x \rightarrow \infty} f(x) = c$).

- Es bestehen die Grenzwert-Aussagen:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1, \quad (1)$$

$$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0. \quad (2)$$

2.6.4 Asymptoten

- 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.

- b) 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.
- c) 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 $x \rightarrow -\infty$.

Definition 2.6.5 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.

Satz 2.6.6

- 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.

2.6.7 Korollar

- 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 $q(x) \neq 0$.

Satz 2.6.8 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.

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

2.6.9 Korollar *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. $f(\bar{x}) = 0$.*

• **Bisektion** Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion so, dass $f(b) \cdot f(a) < 0$. Dann existiert nach Korollar 2.6.9 eine Nullstelle in $[a, b]$. Diese kann man durch sukzessives Halbieren des Intervalls beliebig genau 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. Dabei stößt man entweder nach endlich vielen Schritten auf die Nullstelle, oder der Fehler wird in jedem Schritt mindestens halbiert.

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

Definition 3.1.1 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

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (1)$$

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

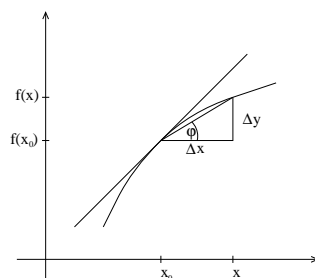
- Man nennt $\frac{\Delta f(x)}{\Delta x} := \frac{f(x_0+h)-f(x_0)}{h}$ den **Differenzenquotienten**. Man kann (1.2) auch durch $\lim_{x \rightarrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$ ersetzen.
- andere Bezeichnungen: $\frac{df(x_0)}{dx}$, $\frac{d}{dx} f$
- Falls f in jedem Punkt $x_0 \in I$ differenzierbar 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:

$f(x)$	$f'(x)$
c	0
$ax + b$	a
x^n	nx^{n-1} für $n \in \mathbb{N}$
\sqrt{x}	$\frac{1}{2\sqrt{x}}$ für $x > 0$

(2)

3.1.2 Geometrische Interpretation Sei $y = f(x)$ eine gegebene Funktion. Dann ist $\tan \varphi = \frac{\Delta y}{\Delta x}$ der Anstieg der Sekante durch die Punkte $(x_0, f(x_0))$, $(x, f(x))$.



Der Grenzwert $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ ist der Anstieg der Tangente an $y = f(x)$ im Punkt $(x_0, f(x_0))$, d.h. die Tangente des Graphen $y = f(x)$ im Punkt $(x_0, f(x_0))$ genügt der Formel

$$y = f'(x_0)(x - x_0) + f(x_0). \quad (3)$$

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

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

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 (*) erhalten wir $m = f'(x_0)$, d.h. die Tangente (3) ist die „beste“ lineare Approximation von $f(x)$ nahe $(x_0, f(x_0))$.

Satz 3.1.4 Jede in $x_0 \in I$ differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ ist in x_0 auch stetig.

- Stetigkeit von f ist **nicht** hinreichend für Differenzierbarkeit!

Satz 3.1.5 Sind Funktionen $f, g : I \rightarrow \mathbb{R}$ im Punkt $x \in I$ differenzierbar, so sind die in a), ..., d) links stehenden Funktionen differenzierbar und man berechnet die Ableitungen gemäß

a) $(f(x) + g(x))' = f'(x) + g'(x),$

b) $(f(x) \cdot g(x))' = f'(x)g(x) + f(x)g'(x),$

$$c) \left(\frac{f(x)}{g(x)} \right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}, \quad \text{falls } g(x) \neq 0.$$

$$d) \left(\frac{1}{g(x)} \right)' = -\frac{g'(x)}{g^2(x)}, \quad \text{falls } g(x) \neq 0.$$

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

$$p'(x) = \sum_{i=1}^n i a_i x^{i-1}. \quad (4)$$

- Wir haben folgenden Spezialfall von Satz 3.1.5 d)

$$\left(\frac{1}{x^n} \right)' = \frac{-n}{x^{n+1}} \quad n \in \mathbb{N}, x \neq 0. \quad (5)$$

Satz 3.1.6 Die Sinus- und die Cosinusfunktionen sind auf \mathbb{R} differenzierbar und es bestehen die Ableitungsregeln

$$a) (\sin x)' = \cos x,$$

$$b) (\cos x)' = -\sin x,$$

$$c) (\tan x)' = \frac{1}{\cos^2 x}, \quad x \neq (2k+1)\frac{\pi}{2}, k \in \mathbb{Z},$$

$$d) (\cot x)' = \frac{-1}{\sin^2 x}, \quad x \neq k\pi, k \in \mathbb{Z}.$$

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

$$(f(g(x)))' = f'(g(x)) \cdot g'(x). \quad (6)$$

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

$$f^{(0)}(x) := f(x)$$

$$f^{(1)}(x) := f'(x)$$

$$f^{(n)}(x) := \frac{d}{dx} f^{(n-1)}(x) = \frac{d^n}{dx^n} f(x).$$

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

- Mit Hilfe der Eulerschen Formel Kapitel 2 (14) und Satz 3.1.6 kann man zeigen, dass

$$\frac{d}{dt} e^{iwt} = iw e^{iwt}, \quad t \in \mathbb{R}. \quad (7)$$

3.2 Anwendungen der Differentiation

3.2.1 Maxima und Minima 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**.

Satz 3.2.2 Ist f in einem offenen Intervall I differenzierbar, so gilt:

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

• Aus Satz 2.2 und den Bemerkungen folgt, dass

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

Kandidaten für Extrema sind.

Satz 3.2.3 (Mittelwertsatz) Ist die Funktion f auf $[a, b]$ stetig und auf (a, b) differenzierbar, dann gibt es einen Punkt $x_0 \in (a, b)$ mit

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

Satz 3.2.4 Für eine auf dem Intervall $I \subseteq \mathbb{R}$ differenzierbare Funktion f gilt:

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

Satz 3.2.5 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).

Satz 3.2.6 Ist f auf (a, b) zweimal stetig differenzierbar und $x_0 \in (a, b)$ ein stationärer Punkt, dann gilt:

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

Definition 3.2.7 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:

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

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

Satz 3.2.8 Sei f auf einem Intervall I zweimal differenzierbar, dann gilt:

a) $f'' \geq 0$ auf $I \Rightarrow f$ ist konvex,

b) $f'' \leq 0$ auf $I \Rightarrow f$ ist konkav.

Definition 3.2.9 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.

Satz 3.2.10 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.

Satz 3.2.11 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

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

Satz 3.2.12 (Guillaume de l'Hospital, 1661–1704). 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 $f(x) \rightarrow \infty, g(x) \rightarrow \infty$ 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 $x \rightarrow \pm\infty$.

3.2.13 Kurvendiskussion Um eine Vorstellung von der Gestalt des Graphen $y = f(x)$ zu bekommen, führt man eine Kurvendiskussion durch, d.h.

- 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.

3.2.14 Nullstellen und Fixpunkte Oftmals benötigt man die Nullstellen einer Funktion. 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 zu f ist äquivalent zum Finden von Fixpunkten zu $g(x) = f(x) + x$.

Sei die Funktion f auf $[a, b]$ definiert. Dann heißt $x \in [a, b]$ **Fixpunkt** von f , wenn $f(x^*) = x^*$.

Satz 3.2.15 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

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}_0 \quad (2)$$

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

- 3) Es gilt die Abschätzung:

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

3.2.16 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

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 \in [a, b]. \quad (3)$$

Satz 3.2.17 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 in (3) definierte Iterationsfolge x_n für Startwerte x_0 aus diesem Intervall I gegen x^* konvergiert. Weiterhin gilt

$$|x_{n+1} - x^*| \leq M |x_n - x^*|^2,$$

wobei $M = \frac{\max\{|f''(x)|; x \in I\}}{\min\{|f'(x)|; x \in I\}}$.

3.3 Umkehrfunktionen

Definition 3.3.1 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.

$$x = g(y) \Leftrightarrow y = f(x). \quad (1)$$

• Aus (3.2) folgt

$$\begin{aligned} f(x) &= f(g(y)) = y, \\ g(y) &= g(f(x)) = x. \end{aligned} \quad (2)$$

Satz 3.3.2 a) 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.

b) Ist f über D umkehrbar mit der Umkehrfunktion $f^{-1} : f(D) \rightarrow \mathbb{R}$, dann liegen die Graphen $y = f(x)$ und $y = f^{-1}(x)$ symmetrisch zur Geraden $y = x$.

c) Die Umkehrfunktion $f^{-1} : f(D) \rightarrow \mathbb{R}$ einer über einem Intervall $I \subseteq \mathbb{R}$ umkehrbaren und differenzierbaren Funktion f ist in allen Punkten $y \in f(I)$ mit $f'(f^{-1}(y)) \neq 0$ differenzierbar und es gilt

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}. \quad (3)$$

3.3.3 n -te Wurzel 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 $x^n = (-x)^n$. 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:

$$y = \sqrt[n]{x} \Leftrightarrow y^n = x \begin{cases} \text{für } x \geq 0, & n \in \mathbb{N} \text{ gerade,} \\ \text{für } x \in \mathbb{R}, & n \in \mathbb{N} \text{ ungerade.} \end{cases} \quad (4)$$

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

$$\begin{aligned} x^{\frac{1}{n}} &:= \sqrt[n]{x}, \\ x^{\frac{m}{n}} &:= \left(x^{\frac{1}{n}}\right)^m. \end{aligned} \quad (5)$$

Aus der Definition ergeben sich sofort die Rechenregeln

$$\begin{aligned} x^\alpha x^\beta &= x^{\alpha+\beta}, \\ (x^\alpha)^\beta &= x^{\alpha\beta}, \\ x^\alpha y^\alpha &= (xy)^\alpha \end{aligned} \quad (6)$$

und die Formel für die Ableitung

$$(x^\alpha)' = \alpha x^{\alpha-1}, \quad x > 0. \quad (7)$$

3.3.4 Arcusfunktionen Wir betrachten die Umkehrfunktionen der Kreisfunktionen $\sin x$, $\cos x$, $\tan x$, $\cot x$. Dazu müssen wir geeignete Intervalle finden, auf denen diese Funktionen strikt monoton sind.

- Die Sinusfunktion ist auf $[-\pi/2, \pi/2]$ strikt monotonwachsend, denn $\sin'(x) = \cos x > 0$ falls $x \in (-\pi/2, \pi/2)$. Die Umkehrfunktion heißt **Arcussinus**, in Zeichen \arcsin , und ist definiert durch

$$\begin{aligned} \arcsin &: [-1, 1] \rightarrow [-\pi/2, \pi/2] \\ y = \arcsin x &\Leftrightarrow (\sin y = x) \wedge (y \in [-\pi/2, \pi/2]). \end{aligned} \quad (8)$$

Für die Ableitung gilt:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \quad \text{für } -1 < x < 1. \quad (9)$$

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

$$\begin{aligned} \arccos &: [-1, 1] \rightarrow [0, \pi] \\ y = \arccos x &\Leftrightarrow (\cos y = x) \wedge (y \in [0, \pi]). \end{aligned} \quad (10)$$

Analog zu (3.13) beweist man

$$(\arccos x)' = \frac{-1}{\sqrt{1-x^2}}, \quad \text{für } -1 < x < 1 \quad (11)$$

- Die Tangens- bzw. Cotangensfunktion ist auf dem Intervall $(-\pi/2, \pi/2)$ bzw. $(0, \pi)$ strikt monoton. Die Umkehrfunktion heißt **Arcustangens** bzw. **Arcuscotangens** und ist definiert als

$$\begin{aligned} \arctan : \mathbb{R} &\rightarrow (-\pi/2, \pi/2) \\ \arctan x = y &\Leftrightarrow (\tan y = x) \quad \wedge \quad (y \in (-\pi/2, \pi/2)) \end{aligned} \quad (12)$$

$$\begin{aligned} \operatorname{arccot} : \mathbb{R} &\rightarrow (0, \pi) \\ \operatorname{arccot} x = y &\Leftrightarrow \cot y = x \quad \wedge \quad y \in (0, \pi) \end{aligned} \quad (13)$$

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

$$\begin{aligned} (\arctan x)' &= \frac{1}{1+x^2}, \\ (\operatorname{arccot} x)' &= \frac{-1}{1+x^2} \end{aligned} \quad (14)$$

3.4 Exponential- und Logarithmusfunktion

Satz 3.4.1 Die Exponentialfunktion hat folgende Eigenschaften:

a) $e^0 = 1$, $e^x > 0$ für alle $x \in \mathbb{R}$.

b) Für alle $x \in \mathbb{R}$ gilt:

$$\frac{d}{dx} e^x = e^x. \quad (1)$$

Somit ist e^x beliebig oft differenzierbar.

c) 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.

d) Es bestehen die Gleichungen

$$\begin{aligned} e^{x+y} &= e^x e^y, & \forall x, y \in \mathbb{R}, \\ e^{-x} &= \frac{1}{e^x}, & \forall x \in \mathbb{R}. \end{aligned} \quad (2)$$

- 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

$$\exp(rx) = (\exp(x))^r \quad \forall x \in \mathbb{R}, r \in \mathbb{Q}. \quad (3)$$

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

$$e^x := \exp(x) = \lim_{n \rightarrow \infty} \exp(r_n) = \lim_{n \rightarrow \infty} e^{r_n}, \quad (4)$$

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

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

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

$$b) \lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \infty, \quad \forall n \in \mathbb{N},$$

das heißt die Exponentialfunktion wächst rascher als jede Potenz von x .

3.4.3 Exponentiale 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 4.1c)

$$u(t) = ce^{at}.$$

3.4.4 Der natürliche Logarithmus Die e -Funktion wächst strikt monoton, also existiert nach Satz 3.4a) auf $(0, \infty)$ eine Umkehrfunktion. Diese heißt der **natürliche Logarithmus** und ist definiert durch:

$$\begin{aligned} \ln : (0, \infty) &\rightarrow \mathbb{R} \\ y = \ln x &\Leftrightarrow e^y = x. \end{aligned} \quad (5)$$

Insbesondere hat man die Rechenregeln:

$$\begin{aligned} \ln(e^x) &= x, & x \in \mathbb{R}, \\ e^{\ln x} &= x, & x > 0. \end{aligned} \quad (6)$$

$$\begin{aligned} \ln e &= 1, \\ \ln 1 &= 0, \end{aligned} \quad (7)$$

$$\begin{aligned} x \in (0, 1) &\Rightarrow \ln x < 0, \\ x \in (1, \infty) &\Rightarrow \ln x > 0, \end{aligned} \quad (8)$$

$$\begin{aligned} \lim_{x \rightarrow 0^+} \ln x &= -\infty, \\ \lim_{x \rightarrow \infty} \ln x &= \infty. \end{aligned} \quad (9)$$

Satz 3.4.5 *Der natürliche Logarithmus ist strikt konkav und differenzierbar. Es gilt*

$$\frac{d}{dx} \ln x = \frac{1}{x}, \quad (10)$$

$$\begin{aligned} \ln(xy) &= \ln x + \ln y, \\ \ln \frac{x}{y} &= \ln x - \ln y, \end{aligned} \quad (11)$$

wobei $x, y > 0$.

• *Der natürliche Logarithmus wächst langsamer als jede n -te Wurzel, d.h. für alle $n \in \mathbb{N}$ hat man*

$$\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[n]{x}} = 0. \quad (12)$$

3.4.6 (Allgemeine Exponentialfunktionen und Logarithmen)

Sei $a > 0$. Die Formel (4.4) besagt, dass für $r \in \mathbb{Q}$ und $x \in \mathbb{R}$ 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)

$$a^x := e^{x \ln a} \quad \text{für alle } x \in \mathbb{R}, a > 0 \quad (13)$$

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

Satz 3.4.7 *Für die Exponentialfunktion zur Basis a , mit $a > 0$, gilt für alle $x, y \in \mathbb{R}$ und $b > 0$*

$$\begin{aligned} a^x a^y &= a^{x+y}, \\ (ab)^x &= a^x b^x, \end{aligned} \quad (14)$$

$$\begin{aligned} (a^x)^y &= a^{xy}, \\ \ln(a^x) &= x \ln a, \end{aligned} \quad (15)$$

$$\frac{d}{dx} a^x = a^x \ln a. \quad (16)$$

- Nun können wir auch **Potenzen** von x für beliebige $\alpha \in \mathbb{R}$ definieren. Aus (4.19) folgt für $x > 0$ und $\alpha \in \mathbb{R}$

$$x^\alpha = e^{\alpha \ln x}. \quad (17)$$

Man berechnet sofort

$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \frac{\alpha}{x} = \alpha x^{\alpha-1}. \quad (18)$$

- 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:

$$\begin{aligned} \ln_a : (0, \infty) &\rightarrow \mathbb{R} \\ y = \ln_a(x) &\Leftrightarrow a^y = x. \end{aligned} \quad (19)$$

Es bestehen die Rechenregeln

$$\begin{aligned} \ln_a(xy) &= \ln_a(x) + \ln_a(y), \\ \frac{d}{dx} \ln_a(x) &= \frac{1}{x \ln a}. \end{aligned} \quad (20)$$

Definition 3.4.8 Die *Hyperbelfunktionen* **sinushyperbolicus**, **cosinushyperbolicus** und **tangenshyperbolicus** sind für alle $x \in \mathbb{R}$ definiert durch

$$\begin{aligned} \sinh(x) &:= \frac{e^x - e^{-x}}{2}, \\ \cosh(x) &:= \frac{e^x + e^{-x}}{2}, \\ \tanh(x) &:= \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \end{aligned}$$

- Aus der Definition 3.4.8 und den Eigenschaften der e -Funktion erhält man folgende Rechenregeln:

$$\begin{aligned} \sinh(-x) &= -\sinh(x) \\ \cosh(-x) &= \cosh(x) \end{aligned} \quad (21)$$

$$\begin{aligned} \sinh(x+y) &= \sinh(x) \cosh(y) + \cosh(x) \sinh(y), \\ \cosh(x+y) &= \cosh(x) \cosh(y) + \sinh(x) \sinh(y), \end{aligned} \quad (22)$$

$$(\cosh(x))^2 - (\sinh(x))^2 = 1. \quad (23)$$

Für die Ableitungen gilt:

$$\begin{aligned} \sinh(x)' &= \cosh(x), \\ \cosh(x)' &= \sinh(x), \\ \tanh(x)' &= \frac{1}{\cosh^2(x)}. \end{aligned} \quad (24)$$

- Die Funktion $\sinh(x)$ ist auf \mathbb{R} strikt monoton wachsend, also existiert eine Umkehrfunktion. Diese heißt **areasinushyperbolicus** und wird mit $\operatorname{arsinh}(x)$ bezeichnet. Die Funktion $\cosh(x)$ ist auf \mathbb{R}_0^+ strikt monoton wachsend. Die Umkehrfunktion heißt **areacosinushyperbolicus** und wird mit $\operatorname{arcosh}(x)$ bezeichnet. Es gilt:

$$\begin{aligned} \operatorname{arsinh}(x) &= \ln \left(x + \sqrt{x^2 + 1} \right), \quad x \in \mathbb{R}, \\ \operatorname{arcosh}(x) &= \ln \left(x + \sqrt{x^2 - 1} \right), \quad x \geq 1. \end{aligned} \quad (25)$$

- Die Kettenregel und (4.33) liefern:

$$\begin{aligned}\frac{d}{dx} \operatorname{arsinh}(x) &= \frac{1}{\sqrt{1+x^2}}, & x \in \mathbb{R}, \\ \frac{d}{dx} \operatorname{arcosh}(x) &= \frac{1}{\sqrt{x^2-1}}, & x > 1.\end{aligned}\tag{26}$$

Kapitel 4

Integration

Die Integration (oder Aufleitung) ist die Umkehroperation zur Differentiation (oder Ableitung). Es wird also das Problem behandelt, wie man aus der Kenntnis der Ableitung einer Funktion die Funktion selbst wiederherstellt. 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 (Kapitel 3 Satz 3.2.3) $\xi_i \in (x_{i-1}, x_i)$ mit

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

Demzufolge gilt:

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

Wenn die Zerlegung fein genug 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

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

die **Riemannsche Summe von f** (Bernhard Riemann, 1826–1866). Der Grenzwert der Folge $(Z_n)_{n \geq 1}$ existiert, 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.

$$\int_a^b f(x) dx := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1}). \quad (2)$$

- Um Fallunterscheidungen zu vermeiden, setzen wir

$$\begin{aligned} \int_a^a f(x) dx &:= 0, \\ \int_b^a f(x) dx &:= - \int_a^b f(x) dx, \quad b > a. \end{aligned} \quad (3)$$

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

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

Genau genommen wird erst durch das Integral der Flächeninhalt definiert.

Satz 4.1.3 Seien f, g stückweise stetige Funktion auf $[a, b]$ und $\alpha, \beta \in \mathbb{R}$ und $c \in [a, b]$. Es gelten:

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

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

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

Satz 4.1.4 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 besteht die Ungleichung

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

Satz 4.1.5 (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

$$\int_a^b f(x) dx = f(\xi) (b - a). \quad (4)$$

Definition 4.1.6 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.

Satz 4.1.7 (Hauptsatz der Differential und Integralrechnung)

Ist f eine auf dem Intervall $I = [a, b]$ stetige Funktion, dann gilt:

a) Die durch

$$F(x) := \int_a^x f(t) dt \quad x \in I$$

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

$$\frac{d}{dx} \left(\int_a^x f(t) dt \right) = f(x). \quad (5)$$

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

• Hinweis: Statt F kann auch

$$F_{x_0}(x) = \int_{x_0}^x f(t) dt$$

genommen werden. Dabei ist x_0 ein beliebiger Wert aus I .

b) Mit einer beliebigen Stammfunktion F von f gilt:

$$\int_a^b f(x) dx = F(x) \Big|_a^b := F(b) - F(a). \quad (6)$$

• Berechnung des bestimmten Integrals:

1) Berechne eine Stammfunktion F von f , d.h. $F' = f$.

2) $\int_a^b f(x) dx = F(b) - F(a)$

Definition 4.1.8 Die Menge aller Stammfunktionen von f wird mit $\int f dx$ bezeichnet und heißt **unbestimmtes Integral** von f .

• Nach Satz 4.1.7 a) gilt

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

wobei c alle reellen Konstanten durchläuft und F eine feste Stammfunktion von f ist. Somit haben wir

$$\int f(x) dx = F + c \Leftrightarrow F' = f. \quad (7)$$

• Aus den Differentiationsformeln folgen somit folgende Integrationsformeln:

$F(x)$	$f(x) = F'(x)$	Bedingungen
$\frac{1}{n+1} x^{n+1}$	x^n	$n \neq -1$
$\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 < 1$
$\arctan x$	$\frac{1}{1+x^2}$	
$\frac{1}{a} e^{ax}$	e^{ax}	$a \neq 0$
$\cosh x$	$\sinh x$	
$\sinh x$	$\cosh x$	
$\frac{1}{2} \ln \frac{1+x}{1-x}$	$\frac{1}{1-x^2}$	$ x \leq 1$
$\ln(x + \sqrt{1+x^2})$	$\frac{1}{\sqrt{1+x^2}}$	

4.2 Integrationsregeln

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

$$\int (af(x) + bg(x)) dx = a \int f(x) dx + b \int g(x) dx. \quad (1)$$

4.2.2 Partielle Integration Die Produktregel $(uv)' = u'v + uv'$ liefert, dass uv eine Stammfunktion von $u'v + uv'$ ist, d.h.

$$\int u'(x)v(x) dx = uv - \int u(x)v'(x) dx. \quad (2)$$

Für das bestimmte Integral erhalten wir

$$\int_a^b u'v dx = u(x)v(x) \Big|_a^b - \int_a^b uv' dx. \quad (3)$$

- Aus (2) folgt mit $u' = 1$

$$\int v(x) dx = xv - \int xv'(x) dx, \quad (4)$$

insbesondere also:

$$\int \ln x dx = x \ln x - \int x \frac{1}{x} dx = x(\ln x - 1) + c. \quad (5)$$

- Für Integrale der Form

$$\begin{aligned} S_n &:= \int_a^b (\sin x)^n dx, & C_n &:= \int_a^b (\cos x)^n dx, \\ A_n &:= \int_a^b x^n \sin x dx, & B_n &:= \int_a^b x^n \cos x dx, \\ E_n &:= \int_a^b x^n e^x dx, & L_n &:= \int_a^b (\ln x)^n dx \end{aligned} \quad (6)$$

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

4.2.3 Substitutionsmethode Aus der Kettenregel für die Ableitung $(F(g(x)))' = F'(g(x)) g'(x)$ folgt mit $f(x) = F'(x)$

$$\int f(g(x)) g'(x) dx = F(g(x)) + c \quad (7)$$

und für das bestimmte Integral

$$\int_a^b f(g(x)) g'(x) dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(t) dt. \quad (8)$$

Es gibt zwei Möglichkeiten (8) anzuwenden.

1) Berechnung von $\int f(g(x)) g'(x) dx$

- a) Substitution $g(x) = t$, $g'(x) dx = dt$
- b) Berechnung von $\int f(t) dt = F(t) + c$
- c) Rücksubstitution $t = g(x)$ und somit $\int f(g(x)) g'(x) dx = F(g(x))$

2) Berechnung von $\int f(x) dx$

- a) Substitution $x = g(t)$, $dx = g'(t) dt$ mit „geeigneter“ umkehrbarer Funktion g
- b) Berechnung von $\int f(g(t)) g'(t) dt = H(t) + c$
- c) Auflösen von $x = g(t)$ nach t , d.h. $t = h(x)$, und somit ergibt sich $\int f(x) dx = H(h(x)) + c$

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

$$\int_0^{2\pi} \sin(kx + \varphi) dx = -\frac{1}{k} \cos(kx + \varphi) \Big|_0^{2\pi} = 0. \quad (9)$$

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

$$\begin{aligned} \int_0^{2\pi} \sin(kx) dx &= 0, \\ \int_0^{2\pi} \cos(kx) dx &= 0, \end{aligned} \quad (10)$$

falls $k \in \mathbb{Z} \setminus \{0\}$. Mit Hilfe der Additionstheoreme erhalten wir:

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

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

Somit gelten für $m, n \in \mathbb{Z}$ die folgenden **Orthogonalitätsbeziehungen**:

$$\begin{aligned} \int_0^{2\pi} \sin(mx) \sin(nx) dx &= \begin{cases} 0 & \text{falls } m = n = 0, \\ 0 & \text{falls } m \neq n, \\ \pi & \text{falls } m = n \neq 0, \end{cases} \\ \int_0^{2\pi} \cos(mx) \cos(nx) dx &= \begin{cases} 2\pi & \text{falls } m = n = 0, \\ 0 & \text{falls } m \neq n, \\ \pi & \text{falls } m = n \neq 0, \end{cases} \\ \int_0^{2\pi} \cos mx \sin nx dx &= 0 \quad \text{für alle } m, n \in \mathbb{N}_0. \end{aligned} \quad (11)$$

4.2.4 Symmetrien: *Integrale lassen sich leichter berechnen, wenn man Symmetrien des Integranden und des Integrationsbereiches beachtet. Beispiele hierfür sind gerade und ungerade Integranden. 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 Sei $f(x) = \frac{p(x)}{q(x)}$ eine rationale Funktion mit Grad $p <$ Grad q und teilerfremden Polynomen p und q . Eine **Partialbruchzerlegung** besteht aus folgenden Schritten:

1) Finden einer Faktorisierung von $q(x)$, d.h.

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

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

2) Für jede Nullstelle $b \in \{b_1, \dots, b_r\}$ der Vielfachheit k und jedes quadratische Polynom $Q \in \{q_1, \dots, q_s\}$ der Vielfachheit l bilden wir Funktionen der Form

$$\begin{aligned} \frac{A_1}{x - b}, \quad \frac{A_2}{(x - b)^2}, \quad \dots, \quad \frac{A_k}{(x - b)^k}, \\ \frac{B_1x + C_1}{Q(x)}, \quad \frac{B_2x + C_2}{(Q(x))^2}, \quad \dots, \quad \frac{B_lx + C_l}{(Q(x))^l} \end{aligned}$$

wobei A_i, B_i, C_i reelle Koeffizienten sind.

3) Diese Funktionen heißen **Partialbrüche**. Man setzt nun die rationale Funktion $f(x) = \frac{p(x)}{q(x)}$ als Summe obiger Partialbrüche an, wobei die Koeffizienten A_i, B_i, C_i zu bestimmen sind.

• Die dabei auftretenden Integrale berechnen sich wie folgt:

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

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

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

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

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

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

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

4.3.2 Integration von $R(e^x)$ 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.

4.3.3 Integration von $R(x, \sqrt[k]{\frac{ax+b}{cx+d}})$ 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. Ein Integral vom Typ $\int R(x, \sqrt[k]{\frac{ax+b}{cx+d}}) dx$ mit $ad - bc \neq 0$ wird mit der Substitution

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

in ein Integral einer rationalen Funktion überführt.

4.3.4 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 .

4.3.5 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:

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

4.4 Uneigentliche Integrale

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

Definition 4.4.1 Sei $b \in \mathbb{R} \cup \{\infty\}$. Die Funktion f sei auf dem Intervall $[a, b)$ definiert und auf jedem abgeschlossenen Teilintervall $[a, c]$, $c < b$, 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 $(a, b]$

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

Beispiele:

1)

$$\int_1^\infty \frac{dx}{x} = \lim_{c \rightarrow \infty} \int_1^c \frac{dx}{x} = \lim_{c \rightarrow \infty} \ln c = \infty, \quad (\text{divergiert}). \quad (1)$$

$$\int_0^1 \frac{dx}{x} = \lim_{c \rightarrow 0^+} -\ln c = \infty, \quad (\text{divergiert}). \quad (2)$$

2) Sei $\alpha \in \mathbb{R} \setminus \{1\}$.

$$\int_1^\infty \frac{dx}{x^\alpha} = \lim_{c \rightarrow \infty} \frac{1}{\alpha-1} \left(1 - \frac{1}{c^{\alpha-1}}\right) = \begin{cases} \frac{1}{\alpha-1}, & \text{falls } \alpha > 1, \\ \infty, & \text{falls } \alpha < 1. \end{cases} \quad (3)$$

$$\int_0^1 \frac{dx}{x^\alpha} = \lim_{c \rightarrow 0^+} \frac{1}{\alpha-1} \left(\frac{1}{c^{\alpha-1}} - 1\right) = \begin{cases} \infty, & \text{falls } \alpha > 1 \\ \frac{1}{1-\alpha} & \text{falls } \alpha < 1. \end{cases} \quad (4)$$

3)

$$\int_{-\infty}^\infty e^{-x^2} dx = 2 \int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Die Konvergenz des Integrals ist leicht einzusehen. Die Berechnung des Wertes erfolgt mit der zweidimensionalen Integralrechnung oder mit komplexer Funktionentheorie.

4) Für $x > 0$ konvergiert das Integral

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt$$

(Gamma-Funktion). Man rechnet nach

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

Die Γ -Funktion stellt also eine stetige und differenzierbare Erweiterung der Fakultätsfunktion auf $\{x \in \mathbb{R}, x > 0\}$ dar.

Satz 4.4.2 Ist f auf $[a, \infty)$ und g auf $(0, b]$ stückweise stetig und sind $\alpha, K \in \mathbb{R}$, dann gilt:

$$a) |f(x)| \leq K \frac{1}{x^\alpha}, \quad a \leq x < \infty, \quad 1 < \alpha \quad \Rightarrow \quad \int_a^\infty f(x) dx \text{ konvergiert,}$$

$$b) |f(x)| \leq K \frac{1}{x^\alpha}, \quad 0 < x \leq b, \quad 0 < \alpha < 1 \quad \Rightarrow \quad \int_0^b f(x) dx \text{ konvergiert.}$$

Definition 4.4.3 Sei $f : (a, b) \rightarrow \mathbb{R}$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ eine auf jedem abgeschlossenen Teilintervall $[\alpha, \beta]$, $a < \alpha < \beta < b$ stückweise stetige Funktion und sei $c \in (a, b)$. Falls die beiden uneigentlichen Integrale

$$\int_a^c f(x) dx = \lim_{\alpha \rightarrow a^+} \int_\alpha^c f(x) dx,$$

$$\int_c^b f(x) dx = \lim_{\beta \rightarrow b^-} \int_c^\beta f(x) dx$$

konvergieren, heißt das **uneigentliche Integral** $\int_a^b f(x) dx$ konvergent und ist definiert durch

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx.$$

4.4.4 Ausnahmestellen im Innern Sei f auf $[a, b]$ definiert und sei $c \in (a, b)$ ein Punkt so, dass f auf $[a, c)$ und $(c, b]$ stückweise stetig ist. Falls die uneigentlichen Integrale $\int_a^c f(x) dx$ und $\int_c^b f(x) dx$ konvergieren, setzen wir

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Analog wird der Fall endlich vieler solcher „Ausnahmepunkte“ $x_i \in [a, b]$ behandelt.

4.4.5 Cauchyscher Hauptwert (Augustin Louis Cauchy, 1789–1857). Es kann passieren, dass die beiden uneigentlichen Integrale aus 4.4.4 divergieren, aber dass der Grenzwert

$$\lim_{\varepsilon \rightarrow 0^+} \left(\int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right) \quad (5)$$

existiert. In diesem Fall wird der Grenzwert in (5) **Cauchyscher Hauptwert** genannt und mit

$$CHW \int_a^b f(x) dx$$

bezeichnet.

4.5 Kurven-, Längen- und Flächenmessung

Definition 4.5.1 *Unter einer **Kurve** versteht man eine differenzierbare Abbildung eines Intervalls I in die Ebene. Sei der \mathbb{R}^2 mit einem festen kartesischen Koordinatensystem versehen. Dann nennt man die vektorwertige Funktion*

$$\vec{r}: [a, b] \rightarrow \mathbb{R}^2 : t \mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad (1)$$

wobei $x: [a, b] \rightarrow \mathbb{R}, y: [a, b] \rightarrow \mathbb{R}$ zwei differenzierbare Funktionen sind, eine **Parameterdarstellung** der Kurve, t den **Parameter** und $[a, b]$ das **Parameterintervall**.

- Die Darstellung (1) ist äquivalent zu den Gleichungen

$$x = x(t), y = y(t), \quad t \in [a, b]. \quad (2)$$

- Jede Kurve besitzt unendlich viele Parameterdarstellungen.
- **Beispiel:** Ein Kreis K mit Radius r rollt auf der x -Achse. Der Punkt P mit Abstand a vom Kreismittelpunkt beschreibt eine **Zykloide**, die die Parameterdarstellung

$$\begin{aligned} x &= rt - a \sin t, \\ y &= r - a \cos t, \end{aligned}$$

hat, wobei t der Rollwinkel ist.

Definition 4.5.2 *Zu jeder Parameterdarstellung $\vec{r}(t) = (x(t), y(t))^T$ einer Kurve K definiert man den **Tangentialvektor***

$$\dot{\vec{r}}(t) := \lim_{h \rightarrow 0} \frac{1}{h} (\vec{r}(t+h) - \vec{r}(t)) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix}, \quad (3)$$

wobei der Punkt die Ableitung nach t bedeutet.

- 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 .
- Sei in einem Kurvenpunkt $(x(t), y(t))^T$ der Tangentialvektor $\dot{\vec{r}}(t) \neq \vec{0}$. Der **Normalenvektor** $\vec{n}(t)$ entsteht aus dem Tangentialvektor durch Drehung um 90° in positive Richtung, d.h.

$$\vec{n}(t) := \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}. \quad (4)$$

- Zum Zeitpunkt t_0 haben die Tangente bzw. Normale an die Kurve K die Darstellung

$$\begin{aligned} \text{Tangente: } \quad x &= x(s) = x(t_0) + s\dot{x}(t_0), & y &= y(s) = y(t_0) + s\dot{y}(t_0), \\ \text{Normale: } \quad x &= x(s) = x(t_0) - s\dot{y}(t_0), & y &= y(s) = y(t_0) + s\dot{x}(t_0), \end{aligned}$$

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.

Definition 4.5.4 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.

Satz 4.5.5 Die Länge L einer Kurve mit regulärer Parameterdarstellung $x = x(t)$, $y = y(t)$, $\dot{x}^2(t) + \dot{y}^2(t) \neq 0$, $a \leq t \leq b$, beträgt

$$L = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt. \quad (5)$$

Folgerung 4.5.6 Der Graph $y = f(x)$ einer stetig differenzierbaren Funktion $f : [a, b] \rightarrow \mathbb{R}$ hat die Länge

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx. \quad (6)$$

4.5.7 Krümmung Es sei $\vec{r}(t) = (x(t), y(t))^T$ eine reguläre, zweimal differenzierbare Parameterdarstellung einer Kurve. Mit $\varphi(t)$ bezeichnen wir den positiv gemessenen Winkel zwischen der positiven x -Achse und dem Tangentialvektor \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 $P = (x(t), y(t))^T$ als

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

Satz 4.5.8 Die Krümmung einer Kurve mit regulärer, zweimal differenzierbarer Parameterdarstellung $x = x(t)$, $y = y(t)$, $t \in [a, b]$, beträgt im Kurvenpunkt $P(t) = (x(t), y(t))^T$

$$K(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}^2(t) + \dot{y}^2(t))^{\frac{3}{2}}}. \quad (7)$$

Folgerung 4.5.9 Die Krümmung des Graphen $y = f(x)$ einer zweimal differenzierbaren Funktion $f : [a, b] \rightarrow \mathbb{R}$ im Punkt $(x, f(x))$ beträgt

$$K(x) = \frac{f''(x)}{\left(\sqrt{1 + (f')^2(x)}\right)^3}. \quad (8)$$

• Einen durch den Kurvenpunkt $P = (x(t), y(t))^T$ 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:

$$r = \frac{1}{|K|}.$$

4.5.10 Volumen von Rotationskörpern Sei K ein Rotationskörper, der durch Rotation um die x -Achse entsteht. Sei $F(x)$, $x \in [a, b]$, der Flächeninhalt des Querschnittes. Das Volumenelement dV einer „dünnen Scheibe“ der Dicke dx ergibt sich also zu

$$dV = F(x) dx,$$

und somit erhalten wir

$$V = \int dV = \int_a^b F(x) dx.$$

Die Fläche $F(x)$ des Querschnittes errechnet sich durch

$$F(x) = \pi (f(x))^2,$$

wenn $f(x)$ die den Rotationskörper beschreibende Kurve ist. Somit gilt:

Satz 4.5.11 Ein durch Drehung der Kurve $y = f(x)$, $a \leq x \leq b$, um die x -Achse erzeugter Rotationskörper hat das Volumen

$$V = \pi \int_a^b f^2(x) dx.$$

4.5.12 Oberfläche von Rotationskörpern Sei K ein Rotationskörper der durch Drehung der Kurve $y = f(x)$, $x \in [a, b]$ entsteht. Das Oberflächenelement dM des Mantels wird approximiert durch die Mantelfläche eines Zylinders mit dem Radius $f(x)$ und der Höhe $ds = \sqrt{1 + (f')^2} dx$. Also gilt

$$M = \int dM = 2\pi \int f(x) \sqrt{1 + (f')^2} dx. \quad (9)$$

4.5.13 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

$$\int_a^b f(x) dx \approx h \left(\frac{y_0}{2} + y_1 + \dots + y_{n-1} + \frac{y_n}{2} \right). \quad (10)$$

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**. (Thomas Simpson, 1710–1761). Man ersetzt über dem Teilintervall $[x_i, x_{i+2}]$, $i = 0, 2, \dots, 2n - 2$, den Integranden $f(x)$ durch die Parabel die durch die Punkte (x_i, y_i) , (x_{i+1}, y_{i+1}) , (x_{i+2}, y_{i+2}) geht. Man erhält

$$S_0 := \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right),$$

$$S_n := \frac{h}{3} \left[f(a) + f(b) + 4f(a+h) + 2 \sum_{i=1}^{n-1} \left(f(a+2ih) + 2f(a+(2i+1)h) \right) \right],$$

wobei $h := \frac{b-a}{2^n}$.

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 (Jean-Baptiste F., 1768–1830). Bevor man sich mit Reihen von Funktionen beschäftigt, muss man sich genauer mit Reihen reeller Zahlen befassen.

5.1 Unendliche Reihen

Definition 5.1.1 Sei $(a_k)_{k \geq 0}$ eine Folge reeller Zahlen. Es soll die „unendliche Summe“ $\sum_{k=0}^{\infty} a_k = a_0 + a_1 + \dots$ sinnvoll definiert werden. Die Zahlen a_k heißen Glieder oder Summanden der Reihe $\sum_{k=0}^{\infty} a_k$, die endlichen Summen

$$S_n := \sum_{k=0}^n a_k \tag{1}$$

heißen Partialsummen. Wenn die Folge $(S_n)_{n \geq 0}$ konvergiert (bzw. divergiert) sagt man, dass 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** der Reihe und schreibt

$$S = \sum_{k=0}^{\infty} a_k. \tag{2}$$

Satz 5.1.2 (Cauchy-Kriterium) Die Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann konvergent, wenn es zu jeder noch so kleinen Zahl $\varepsilon > 0$ einen Index N_ε gibt, so dass für alle $n, m \geq N_\varepsilon$ mit $m < n$

$$|S_n - S_m| = |a_{m+1} + \dots + a_n| < \varepsilon$$

gilt.

Satz 5.1.3 Die Glieder einer konvergenten Reihe bilden eine Nullfolge.

Definition 5.1.4 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.

Satz 5.1.5 (Leibniz-Kriterium) (Gottfried Wilhelm L., 1646–1769) Für jede monoton fallende Nullfolge $(a_n)_{n \geq 0}$ konvergiert die alternierende Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$.

Satz 5.1.6 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

$$\sum_{k=0}^{\infty} (a_k \pm b_k) = a \pm b; \quad \sum_{k=0}^{\infty} (ca'_k) = ca'.$$

• Elementare Operationen wie Klammern Setzen oder Weglassen oder Umordnen der Glieder, die für endliche Reihen erlaubt sind, sind mit Vorsicht zu behandeln.

Satz 5.1.7 In einer konvergenten Reihe darf man beliebig Klammern setzen, d.h.

$$S = \sum_{k=1}^{\infty} a_k = a_1 + a_2 + \dots = (a_1 + \dots + a_{k_1}) + (a_{k_1+1} + \dots + a_{k_2}) + \dots$$

Definition 5.1.8 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**.

Satz 5.1.9 Jede absolut konvergente Reihe ist konvergent.

Satz 5.1.10 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.

• Es gilt:

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \begin{cases} \text{konvergiert,} & \text{falls } \alpha > 1, \\ \infty, & \text{falls } \alpha \leq 1. \end{cases} \quad (3)$$

Definition 5.1.11 Es sei $\sigma : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ eine bijektive Abbildung und $b_k = a_{\sigma(k)}$. Dann heißt die Reihe $\sum_{k=0}^{\infty} b_k$ eine **Umordnung** der Reihe $\sum_{k=0}^{\infty} a_k$

- Die Reihe $\sum_{k=0}^{\infty} a_k$ ist also auch eine Umordnung der Reihe $\sum_{k=0}^{\infty} b_k$, denn $\sigma^{-1} : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ ist wieder bijektiv und $b_{\sigma^{-1}(k)} = a_k$.
- Durch Umordnen der Reihe $\sum a_k$ entstehen im allgemeinen ganz neue Partialsummen, das heißt Umordnen der Reihe ist i.a. etwas anderes als Umordnen der Folge der Partialsummen.

Satz 5.1.12 Ist die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent mit der Summe S , dann konvergiert jede Umordnung dieser Reihe ebenfalls gegen S .

Satz 5.1.13 (Vergleichskriterium) *Besteht für die Reihenglieder die Abschätzung*

$$0 \leq |a_k| \leq b_k \quad \text{für } k \geq k_0$$

dann gilt:

$$a) \sum_{k=0}^{\infty} b_k \text{ konvergent} \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ absolut konvergent,}$$

$$b) \sum_{k=0}^{\infty} |a_k| = \infty \quad \Rightarrow \quad \sum_{k=0}^{\infty} b_k = \infty.$$

Satz 5.1.14 (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:

$$i) a < 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent,}$$

$$ii) a > 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist divergent.}$$

- Die Bedingung kann nicht durch $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \leq 1$ ersetzt werden!

Satz 5.1.15 (Wurzelkriterium) *Es sei $\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = a$. Dann gilt*

$$a) a < 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent,}$$

$$b) a > 1 \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k \text{ ist divergent.}$$

Satz 5.1.16 *Für absolut konvergente Reihen $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ besteht die Cauchysche Produktformel*

$$\left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right).$$

5.2 Reihen von Funktionen

Definition 5.2.1 *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 $(f_n)_{n \geq 0}$ **punktwise gegen f konvergiert**.*

- Die Konvergenzgeschwindigkeit hängt vom Punkt $x \in I$ ab, und damit gehen im Allgemeinen die Eigenschaften der Folge für die Grenzfunktion verloren.

Definition 5.2.2 Eine Funktionenfolge $(f_n)_{n \geq 0}$ **konvergiert gleichmäßig** auf I gegen die Funktion $f : I \rightarrow \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ einen für alle $x \in I$ gemeinsamen Index N_ε gibt, so dass gilt:

$$n \geq N_\varepsilon \quad \Rightarrow \quad \forall x \in I : |f(x) - f_n(x)| < \varepsilon.$$

• Mit diesem Konvergenzbegriff übertragen sich Eigenschaften der Folge auf den Grenzwert.

Satz 5.2.3 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.

Gegenbeispiel: Die Folge $f_n(x) = x^n$, $x \in [0, 1]$ konvergiert punktweise 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.

Satz 5.2.4 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 $a, b \in I$

$$\int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Gegenbeispiel: 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 gegen 0. Aber es gilt

$$1 = \lim_{n \rightarrow \infty} \int_{n=0}^1 h_n(x) dx \neq \int_{n=0}^1 \lim_{n \rightarrow \infty} h_n(x) dx = \int_{n=0}^1 0 dx = 0.$$

Satz 5.2.5 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

$$f'(x) = \left(\lim_{n \rightarrow \infty} f_n(x) \right)' = \lim_{n \rightarrow \infty} f'_n(x).$$

Gegenbeispiel: 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.

Definition 5.2.6 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.

Satz 5.2.7 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:

$$\int_a^b \left(\sum_{k=0}^{\infty} f_k(x) \right) dx = \int_a^b f(x) dx = \sum_{k=0}^{\infty} \left(\int_a^b f_k(x) dx \right).$$

Satz 5.2.8 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:

$$f'(x) = \left(\sum_{k=0}^{\infty} f_k(x) \right)' = \sum_{k=0}^{\infty} f'_k(x).$$

Satz 5.2.9 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 $\sum_{k=0}^{\infty} a_k x^k$.

Definition 5.3.1 Sei $\sum_{k=0}^{\infty} a_k x^k$, $x \in \mathbb{R}$, eine Potenzreihe und sei

$$M := \left\{ x \in \mathbb{R}; \sum_{k=0}^{\infty} a_k x^k \text{ konvergiert} \right\}.$$

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:

$$R = 0, \quad 0 < R < \infty, \quad R = \infty.$$

Satz 5.3.2 Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius R . Dann gilt:

a) $R = 0 \Leftrightarrow$ Reihe konvergiert nur für $x = 0$.

b) Ist $0 < R < \infty$ 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 $[-\varrho, \varrho]$.

c) Im Fall $R = \infty$ ist die Reihe für alle $x \in \mathbb{R}$ absolut konvergent. Gleichmäßige Konvergenz in jedem Intervall $[a, b]$.

• Der Satz sagt nichts über $|x| = R$ aus. Die Punkte $x = -R$ und $x = R$ müssen für jede Reihe neu untersucht werden.

5.3.3 Berechnung des Konvergenzradius Der Konvergenzradius kann durch folgende Formel berechnet werden:

$$R = \sup \left\{ r \geq 0; \text{ die Folge } (|a_n| r^n)_{n \geq 0} \text{ ist beschränkt} \right\},$$

wobei $R = \infty$, falls die Menge unbeschränkt ist. Oft ist es einfacher, folgende Formeln zu benutzen:

a) 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 $R = \frac{1}{a}$, falls $a \neq 0$, und $R = \infty$, falls $a = 0$.

b) Sei

$$a = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|},$$

dann ist $R = \frac{1}{a}$, falls $a \neq 0$, und $R = \infty$, falls $a = 0$.

Definition 5.3.4 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.

Satz 5.3.5 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:

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} k a_k x^{k-1}, \\ f''(x) &= \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}, \end{aligned} \tag{1}$$

etc. Die abgeleiteten Reihen haben alle den Konvergenzradius R .

Beispiel:

$$\begin{aligned} \frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k = f(x) && |x| < 1, \\ f'(x) &= -\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k x^{k-1} && |x| < 1, \\ f''(x) &= \frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k(k-1) x^{k-2} && |x| < 1. \end{aligned}$$

Satz 5.3.6 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 ($a = 0$, $b = x$)

$$F(x) := \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$$

eine Stammfunktion von f auf $(-R, R)$.

Satz 5.3.7 Wir haben folgende Potenzreihendarstellungen:

$$\begin{aligned} a) \quad e^x &= \sum_{k=0}^{\infty} \frac{1}{k!} x^k, && x \in \mathbb{R}, \\ b) \quad \sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}, && x \in \mathbb{R}, \\ c) \quad \cos x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, && x \in \mathbb{R}, \end{aligned}$$

$$d) \quad \ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}, \quad -1 < x \leq 1,$$

$$e) \quad \arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}, \quad |x| < 1.$$

Satz 5.3.8 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

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$$

• Die Formel aus Satz 5.3.8 enthält als Spezialfälle:

1) $\alpha = n \in \mathbb{N}$

$$\binom{\alpha}{k} = \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = 0 \quad \text{falls } n < k$$

$$\Rightarrow (1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k \quad (\text{klassische binomische Formel})$$

2) $\alpha = \frac{1}{2}$

$$\binom{\frac{1}{2}}{0} = 0, \quad \binom{\frac{1}{2}}{1} = \frac{1}{2}, \quad \binom{\frac{1}{2}}{2} = -\frac{1}{8},$$

$$\binom{\frac{1}{2}}{k} = (-1)^{k-1} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{2 \cdot 4 \cdots (2k)}, \quad k \geq 2$$

$$\Rightarrow \sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 \pm \dots, \quad (2)$$

falls $|x| < 1$.

3) $\alpha = -\frac{1}{2}$

$$\binom{-\frac{1}{2}}{0} = 1, \quad \binom{-\frac{1}{2}}{k} = (-1)^k \frac{1 \cdot 3 \cdots (2k-1)}{2 \cdot 4 \cdots (2k)}, \quad k \geq 1$$

$$\Rightarrow \frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 \pm \dots, \quad (3)$$

falls $|x| < 1$.

Definition 5.3.9 Eine unendliche Reihe der Form

$$\sum_{k=0}^{\infty} a_k (x - a)^k \quad (4)$$

heißt **Potenzreihe mit Zentrum a** .

- Für theoretische Überlegungen reicht es $a = 0$ zu betrachten, denn durch Substitution $z = x - a$ geht die Reihe $\sum_{k=0}^{\infty} a_k (x - a)^k$ in die Reihe $\sum_{k=0}^{\infty} a_k z^k$ mit Zentrum in 0 über.
- Der **Konvergenzradius** von (4) ist der Konvergenzradius der entsprechenden Reihe mit Zentrum 0.

Lemma 5.3.10 Sei $\sum_{k=0}^{\infty} a_k (x - a)^k$ eine Potenzreihe mit Konvergenzradius R . Dann gilt:

$$i) \ x \in (a - R, a + R) \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k (x - a)^k \text{ konvergiert,}$$

$$ii) \ x \notin [a - R, a + R] \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k (x - a)^k \text{ divergiert.}$$

5.3.11 Koeffizientenvergleich Sei f auf dem Intervall $(a - R, a + R)$ als Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k (x - a)^k$ dargestellt. Nach Satz 5.3.5 gilt für die n -te Ableitung

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1) \cdots (k-n+1) a_k (x - a)^{k-n}.$$

Setzt man $x = a$, so erhält man $f^{(n)}(a) = n! a_n$.

Satz 5.3.12 (Eindeutigkeit von Potenzreihen) Sei $R > 0$ und gelte für alle $x \in (a - R, a + R)$:

$$f(x) = \sum_{k=0}^{\infty} a_k (x - a)^k = \sum_{k=0}^{\infty} b_k (x - a)^k,$$

dann haben wir

$$a_k = b_k = \frac{f^{(k)}(a)}{k!}, \quad k = 0, 1, \dots$$

5.4 Taylor–Reihen (Brook T., 1685–1731)

Satz 5.4.1 Für jede auf dem offenen Intervall $I \subseteq \mathbb{R}$ $(n + 1)$ -mal stetig differenzierbare Funktion f und $a, x \in I$ gilt:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + R_{n+1}(x, a),$$

wobei

$$T_n(x, a) := \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k$$

das **Taylor-Polynom** ist und das **Restglied** $R_{n+1}(x, a)$ die Darstellungen

$$R_{n+1}(x, a) = \frac{1}{n!} \int_a^x (x - t)^n f^{(n+1)}(t) dt,$$

$$R_{n+1}(x, a) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - a)^{n+1}, \quad \xi \text{ ist ein bestimmter Wert zwischen } x \text{ und } a,$$

hat.

Satz 5.4.2 Ist die Funktion f auf dem Intervall I n -mal stetig differenzierbar und $a \in I$ mit

$$f'(a) = f''(a) = \dots = f^{(n-1)}(a) = 0, \quad f^{(n)}(a) \neq 0,$$

dann gilt:

i) a Extremstelle $\Leftrightarrow n$ gerade,

ii) n gerade, $f^{(n)}(a) < 0 \Rightarrow$ lokales Maximum,
 n gerade, $f^{(n)}(a) > 0 \Rightarrow$ lokales Minimum.

Definition 5.4.3 Sei f auf dem offenen Intervall I beliebig oft differenzierbar und sei $a \in I$. Die unendliche Reihe

$$T_f(x, a) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

heißt **Taylor-Reihe** mit Zentrum a . Gilt für alle $x \in (a - R, a + R)$ die Gleichung $f(x) = T_f(x, a)$, so sagt man, dass sich f **um** a **als Taylor-Reihe entwickeln lässt**.

Satz 5.4.4 Sei f auf dem Intervall I beliebig oft differenzierbar und sei $a \in I$. Dann konvergiert die Taylor-Reihe $T_f(x, a)$ genau für diejenigen $x \in I$ gegen $f(x)$, für die das Restglied $R_n(x, a)$ mit $n \rightarrow \infty$ gegen 0 strebt. Eine hinreichende Bedingung dafür ist, dass es Konstanten A, B gibt mit

$$|f^{(n)}(x)| \leq A B^n \quad \forall x \in I, \forall n \in \mathbb{N}.$$

5.4.5 Methoden der Reihenentwicklung

1) Die Taylor-Formel

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + R_{n+1}(x, a)$$

mit dem Nachweis, dass $R_n(x, a) \rightarrow 0$, liefert, dass die Taylor-Reihe gegen $f(x)$ konvergiert.

- 2) Bekannte **Reihen differenzieren oder integrieren**. Die Sätze 5.3.5 und 5.3.6 ermöglichen es, durch gliedweise Differentiation und Integration bekannter Reihen neue Reihenentwicklung zu erhalten (siehe Satz 5.3.7).
- 3) **Summe und/oder Produkt bekannter Reihen liefert neue Reihenentwicklungen**. Die Definition von $\sinh(x) = \frac{1}{2}(e^x + e^{-x})$ und $\cosh(x) = \frac{1}{2}(e^x - e^{-x})$ zusammen mit der Reihenentwicklung für e^x liefert:

$$\begin{aligned}\cosh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, & x \in \mathbb{R}, \\ \sinh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}, & x \in \mathbb{R}.\end{aligned}\tag{1}$$

- 4) **Potenzreihen in einander einsetzen**

Sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$, $g(x) = \sum_{k=0}^{\infty} b_k x^k$. Wenn man die Potenzen nach dem Cauchy-Produkt (Satz 5.1.16) berechnet, d.h.

$$g(x)^k = \left(\sum_{n=0}^{\infty} b_n x^n \right)^k =: \sum_{n=0}^{\infty} b_{kn} x^n,$$

dann erhält man

$$f(g(x)) = \sum_{k=0}^{\infty} a_k \left(\sum_{n=0}^{\infty} b_{kn} x^n \right).$$

Dies, nach steigenden x Potenzen geordnet, gibt

$$f(g(x)) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} a_k b_{kn} \right) x^n.$$

5.5 Fourier–Reihen (Jean–Baptiste F., 1768–1830)

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ (oder $\mathbb{R} \rightarrow \mathbb{C}$) P –periodisch, d.h. $\forall x \in \mathbb{R} : f(x + P) = f(x)$. Dabei ist P eine feste, positive reelle Zahl, die **Periode**. Da mit P auch kP ($k \in \mathbb{N}$) Periode zu f ist, wird, falls f nicht konstant ist, als P meist die kleinste Periode genommen. Für viele Anwendungen ist es nützlich, f durch eine Reihe über die einfachsten P –periodischen Funktionen

$\cos\left(\frac{2\pi}{P} nx\right)$ ($n \in \mathbb{N}_0$) und $\sin\left(\frac{2\pi}{P} nx\right)$ ($n \in \mathbb{N}$) darzustellen.

Definition 5.5.1 $P > 0$. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ (oder $\mathbb{R} \rightarrow \mathbb{C}$) stückweise stetig und P –periodisch.

- 1) Die Zahlen

$$a_n = \frac{2}{P} \int_0^P f(x) \cos\left(\frac{2\pi}{P} nx\right) dx \quad (n \in \mathbb{N}_0),$$

$$b_n = \frac{2}{P} \int_0^P f(x) \sin\left(\frac{2\pi}{P} nx\right) dx \quad (n \in \mathbb{N})$$

heißen die **Fourier-Koeffizienten** zu f . Ist insbesondere $P = 2\pi$, dann haben die Fourier-Koeffizienten die Gestalt

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx$$

2) Die Reihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi}{P} nx\right) + b_n \sin\left(\frac{2\pi}{P} nx\right) \right)$$

heißt die **Fourier-Reihe** zu f .

Satz 5.5.2 von Dirichlet (Gustav Lejeune D., 1805–1859) Sei f P -periodisch und stückweise glatt (d.h. das Intervall $[0, P]$ zerfalle in endlich viele Teile $I_\nu = (x_\nu, x_{\nu+1})$, so daß f in jedem I_ν stetig differenzierbar ist und die Grenzwerte $\lim_{x \rightarrow x_\nu^+} f(x)$ und $\lim_{x \rightarrow x_{\nu+1}^-} f(x)$ existieren. An den x_ν kann f unstetig sein.) Dann konvergiert die Fourier-Reihe zu f

- a) an den Stellen x_ν gegen $\frac{1}{2} \left(\lim_{x \rightarrow x_\nu^-} f(x) + \lim_{x \rightarrow x_\nu^+} f(x) \right)$
(arithmetisches Mittel aus links- und rechtsseitigem Grenzwert),
- b) an allen Stetigkeitsstellen x gegen $f(x)$.

Hinweis: Eine Fourier-Reihe wie im Satz kann gliedweise integriert, aber i.a. nicht gliedweise differenziert werden.

5.5.3. Beispiele.

$$1) \quad f(x) = \begin{cases} 1 & \text{für } 0 \leq x < \pi, \\ -1 & \text{für } -\pi \leq x < 0. \end{cases}$$

f ist ungerade. Dann verschwinden alle a_n .

$$b_n = \begin{cases} \frac{4}{\pi n}, & n \text{ ungerade,} \\ 0, & n \text{ gerade.} \end{cases}$$

$$2) \quad f(x) = \begin{cases} -\frac{\pi}{2} + x & \text{für } 0 \leq x \leq \pi \\ -\frac{\pi}{2} - x & \text{für } -\pi \leq x \leq 0 \end{cases}$$

(Sägezahn-Funktion).

$$f(x) = -\frac{4}{\pi} \left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \dots \right).$$

Kapitel 6

Lineare Algebra

6.1 Lineare Gleichungssysteme und Matrizen

Definition 6.1.1 Ein rechteckiges Zahlenschema der Form

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (1)$$

mit $a_{ij} \in \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$ heißt $m \times n$ **Matrix**. Die Zahlen a_{ij} heißen **Elemente** der Matrix \mathbf{A} . Man schreibt abkürzend

$$\mathbf{A} = (a_{ij}) \quad (2)$$

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 sie elementweise 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).
- Als Elemente können auch komplexe Zahlen zugelassen werden. Die so entstehenden Matrizen bilden die Menge $\mathbb{C}^{m \times n}$.

Definition 6.1.2 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** $\lambda \mathbf{A}$ elementweise definiert, d.h.

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= (c_{ij}), \\ \lambda \mathbf{A} &= (d_{ij}),\end{aligned}$$

wobei $c_{ij} := a_{ij} + b_{ij}$ und $d_{ij} := \lambda a_{ij}$ für $i = 1, \dots, m$, $j = 1, \dots, n$.

- Summen und skalare Vielfache für Spalten- und Zeilenvektoren definiert man komponentenweise:

$$\begin{aligned}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} &= \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \\ \lambda \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &= \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}.\end{aligned}$$

Zeilenvektoren werden analog behandelt.

- Wiederholung der 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

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ \vdots \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ a_n \end{pmatrix}$$

$$= a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

und somit lässt sich jeder Spaltenvektor $a \in \mathbb{R}^n$ eindeutig als Summe

$$\mathbf{a} = a_1 \mathbf{e}_1 + \cdots + a_n \mathbf{e}_n \quad (3)$$

mit den **Standard Basisspaltenvektoren**

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n := \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (4)$$

darstellen. 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 diese Operationen 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 $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m \times n}$, $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned}
 \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\
 (\mathbf{A} + \mathbf{B}) + \mathbf{C} &= \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\
 \mathbf{A} + \mathbf{0} &= \mathbf{A} \\
 \mathbf{A} + (-\mathbf{A}) &= \mathbf{0} \\
 (\lambda\mu)\mathbf{A} &= \lambda(\mu\mathbf{A}) \\
 1\mathbf{A} &= \mathbf{A} \\
 (\lambda + \mu)\mathbf{A} &= \lambda\mathbf{A} + \mu\mathbf{A} \\
 \lambda(\mathbf{A} + \mathbf{B}) &= \lambda\mathbf{A} + \lambda\mathbf{B}
 \end{aligned} \tag{5}$$

- Matrizen treten in natürlicher Weise bei Gleichungssystemen auf.

Definition 6.1.3 Ein *lineares Gleichungssystem* mit m Gleichungen in n Unbekannten x_1, \dots, x_n hat die Form

$$\begin{aligned}
 a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\
 &\vdots \\
 a_{m1}x_1 + \dots + a_{mn}x_n &= b_m,
 \end{aligned} \tag{6}$$

wobei a_{ij} die **Koeffizienten** und b_i die **Absolutglieder** sind. Falls $b_i = 0$, $i = 1, \dots, m$, heißt das System **homogen**, ansonsten **inhomogen**.

- (6) kann man auch als

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, m \tag{7}$$

schreiben.

- Wir suchen Werte c_i für die Unbekannten x_i so, dass das lineare Gleichungssystem (6) erfüllt ist. Solche Werte heißen **Lösung** des Systems (6) und werden entweder in der Form

$$x_1 = c_1, \dots, x_n = c_n$$

oder als Spaltenvektor

$$c = (c_1, \dots, c_n)^T$$

angegeben.

- Jedes homogene Gleichungssystem besitzt mindestens die **triviale Lösung**

$$x_1 = x_2 = \dots = x_n = 0.$$

- Nicht jedes System ist lösbar. Es können folgende Fälle auftreten:

a) keine Lösung:

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 3x_1 + 2x_2 &= 2 \end{aligned}$$

b) genau eine Lösung:

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 3x_1 + x_2 &= 5 \end{aligned}$$

Wir wollen das Verfahren am konkreten Beispiel vorstellen:

- Koeffizienten von x_1 in 1. Gleichung = 1, d.h. 1. Gleichung $\times \frac{1}{3}$

$$\begin{aligned} x_1 + \frac{1}{3}x_2 &= \frac{1}{3} \\ 3x_1 + x_2 &= 5 \end{aligned}$$

- Eliminieren von x_1 aus 2. Gleichung, d.h. 2. Gleichung - $3 \times$ 1. Gleichung; Koeffizient von x_2 in 2. Gleichung, d.h. 2. Gleichung $\times (-1)$

$$\begin{aligned} x_1 + \frac{2}{3} &= \frac{1}{3} \\ x_2 &= -4 \end{aligned}$$

- Eliminieren von x_2 aus 1. Gleichung, d.h. 1. Gleichung - $\frac{2}{3} \times$ 2. Gleichung

$$\begin{aligned} x_1 &= 3 \\ x_2 &= -4 \end{aligned}$$

Also ist $x_1 = 3, x_2 = -4$ die Lösung.

c) unendlich viele Lösungen

$$3x_1 + 2x_2 = 1$$

Das ist eine Geradengleichung, mit einem Parameter λ . Also ist

$$\begin{aligned} x_2 &= \lambda \\ x_1 &= \frac{1}{3}(1 - 2\lambda) \end{aligned}$$

eine Lösung für alle $\lambda \in \mathbb{R}$.

Definition 6.1.4 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

$$\mathbf{b} := (b_1, \dots, b_m)^T \in \mathbb{R}^m$$

heißt rechte Seite und die 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**. Man kann davon ausgehen, dass keine der Zeilen oder Spalten von \mathbf{A} der Nullvektor ist.

Satz 6.1.5 Folgende Umformungen eines Gleichungssystems verändern die Lösungsmenge nicht:

- a) Vertauschen zweier Gleichungen,
- b) Multiplikation einer Gleichung mit einer Zahl $\alpha \neq 0$,
- c) Addition eines Vielfachen einer Gleichung zu einer anderen.

Definition 6.1.6 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.

Satz 6.1.7 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.

6.1.8 (Das Gauß'sche Eliminationsverfahren) Das **Gauß'sche Eliminationsverfahren** besteht aus folgenden Schritten, die an der erweiterten Koeffizientenmatrix $(\mathbf{A}|\mathbf{b})$ eines linearen Gleichungssystems ausgeführt werden:

1. Schritt: Ausräumen der 1. Spalte. Wir bilden die Buchführungsmenge $B = \{1\}$ und räumen dann die 1. Spalte aus. Dazu vertauschen wir Spalten derart, daß in der 1. Zeile der 1. Spalte eine Zahl $a \neq 0$ steht. Durch Multiplikation der 1. Zeile mit a^{-1} machen wir daraus eine 1. Danach werden aus allen anderen Zahlen in der 1. Spalte durch Elementarumformungen vom Typ c) Nullen erzeugt.

2. Schritt: Wir suchen die nächste Spalte, die unterhalb der 1. Zeile nicht nur aus Nullen besteht. Wir fügen die Nummer dieser Spalte, es sei die j_2 -te Spalte, zur Buchführungsmenge dazu, d.h. $B = \{1, j_2\}$. Danach räumen wir die j_2 -te Spalte aus, d.h. durch eventuelles Vertauschen von Zeilen mit Index > 1 erhalten wir in der 2. Zeile der j_2 -ten Spalte

eine Zahl ungleich Null. Diese Zahl wird zu einer 1 umgeformt und danach **alle** anderen Glieder dieser Spalte zu Null umgeformt.

3. Schritt: Der zweite Schritt wird analog so lange wiederholt, bis alle Spalten der Koeffizientenmatrix \mathbf{A} ausgeräumt sind. Wir erhalten dadurch eine Buchführungsmenge $B = \{1, j_1, \dots, j_r\}$ mit $1 < j_2 < \dots < j_r \leq n$ und eine erweiterte Koeffizientenmatrix mit **Stufenform**. Etwa bei 6 Gleichungen mit 10 Unbekannten und $B = \{1, 5, 8, 9\}$.

$$\left(\begin{array}{cccccccccc|c} 1 & * & * & * & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & 0 & * & * \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & 1 & 0 & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 1 & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{array} \right)$$

- Die Buchführungsmenge $B = \{j_1, j_2, \dots, j_r\}$ ist dadurch gekennzeichnet, daß in der j_k -ten Spalte in der k -ten Zeile eine 1 steht und sonst nur Nullen.

Satz 6.1.9 (Lösbarkeit) Sei eine erweiterte Koeffizientenmatrix in Stufenform mit Buchführungsmenge $B = \{j_1, \dots, j_r\}$. Das zugehörige lineare Gleichungssystem besitzt genau dann Lösungen, wenn alle Absolutglieder b_i mit Index $i > r$ Null sind.

Satz 6.1.10 (Konstruktion aller Lösungen) Ist n die Anzahl der Unbekannten und ist das System lösbar, so werden die Lösungen durch $n - r$ Parameter beschrieben, wobei r die Anzahl der Elemente der Buchführungsmenge B ist.

Definition 6.1.11 Die im Beweis von Satz 6.1.10 in Abhängigkeit von den Parametern $c_j, j \notin B$, konstruierte Lösung heißt **allgemeine Lösung**. Jede spezielle Wahl der Parameter liefert eine **spezielle Lösung**.

Satz 6.1.12 (Eindeutigkeit) Das System ist genau dann eindeutig lösbar wenn es lösbar ist und $n = r$, d.h. $B = \{1, \dots, n\}$.

Folgerung 6.1.13 Das homogene System, d.h. $b_i = 0, i = 1, \dots, m$, besitzt genau dann nur die triviale Lösung, wenn $r = n$.

Satz 6.1.14 (Struktur der Lösungsmenge)

- 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.
- 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 und einer allgemeinen Lösung \mathbf{u} des zugehörigen homogenen Systems.

6.2 Matrizenmultiplikation

Definition 6.2.1 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 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b} \in \mathbb{R}^n$

$$\begin{aligned} (\mathbf{a}_1 + \mathbf{a}_2) \cdot \mathbf{b} &= \mathbf{a}_1 \cdot \mathbf{b} + \mathbf{a}_2 \cdot \mathbf{b}, \\ \mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) &= \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2, \\ \alpha (\mathbf{a} \cdot \mathbf{b}) &= (\alpha \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha \mathbf{b}). \end{aligned} \tag{1}$$

Definition 6.2.2 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

$$c_{ij} := \sum_{k=1}^n a_{ik} b_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, r, \tag{2}$$

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

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \tag{3}$$

schreiben, mit $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$, $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \dots, b_m)^T \in \mathbb{R}^m$.

- Für den Spezialfall $\mathbf{A} \in \mathbb{R}^{m \times n}$ und $\mathbf{s} \in \mathbb{R}^n$ gilt:

$$\mathbf{A} \cdot \mathbf{s} = \begin{pmatrix} \mathbf{z}_1 \cdot \mathbf{s} \\ \vdots \\ \mathbf{z}_m \cdot \mathbf{s} \end{pmatrix} \tag{4}$$

wobei $\mathbf{z}_i, i = 1, \dots, m$ die Zeilenvektoren von \mathbf{A} sind.

- Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$, seien $\mathbf{a}_i \in \mathbb{R}^m, i = 1, \dots, n$ die Spaltenvektoren von \mathbf{A} und sei $\mathbf{x} \in \mathbb{R}^n$. Dann gilt:

$$\mathbf{A} \cdot \mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n. \quad (5)$$

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} . auch wenn beide Produkte definiert sind.

- Die Matrix

$$\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \quad (6)$$

heißt $n \times n$ **Einheitsmatrix**, sie hat die Spaltendarstellung $\mathbf{E}_n = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ mit den Basisspaltenvektoren $\mathbf{e}_i \in \mathbb{R}^n$.

Satz 6.2.3 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:

- $(\mathbf{A}_1 + \mathbf{A}_2) \cdot \mathbf{B} = \mathbf{A}_1 \cdot \mathbf{B} + \mathbf{A}_2 \cdot \mathbf{B}$,
- $\mathbf{A} \cdot (\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A} \cdot \mathbf{B}_1 + \mathbf{A} \cdot \mathbf{B}_2$,
- $\alpha(\mathbf{A} \cdot \mathbf{B}) = (\alpha\mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\alpha\mathbf{B})$,
- $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$,
- $\mathbf{E}_m \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{E}_n = \mathbf{A}$.

- Selbst wenn beide Produkte $\mathbf{A} \cdot \mathbf{B}$ und $\mathbf{B} \cdot \mathbf{A}$ definiert sind, kann $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$ sein.

Beispiel:

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & -3 \\ 14 & 19 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 14 & 11 \end{pmatrix}$$

■

- 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 es gilt: $a \cdot b = 0 \Rightarrow (a = 0) \vee (b = 0)$.
- Aber es gilt:

$$\mathbf{0} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{0} = \mathbf{0}$$

- **Potenzen** von quadratischen Matrizen $\mathbf{A} \in \mathbb{R}^{m \times n}$ sind rekursiv definiert:

$$\mathbf{A}^0 = \mathbf{E}_n, \quad \mathbf{A}^k := \mathbf{A}^{k-1} \cdot \mathbf{A} = \underbrace{\mathbf{A} \cdot \dots \cdot \mathbf{A}}_{k\text{-mal}} \quad (7)$$

Definition 6.2.4 Zu jeder $m \times n$ Matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$ gibt es eine **transponierte Matrix** $\mathbf{A}^T = (a_{ij}) \in \mathbb{R}^{n \times m}$, deren i -te Zeile aus den Koeffizienten der i -ten Spalte von \mathbf{A} bestehen, d.h. für

$$\mathbf{A} = \begin{pmatrix} \boxed{a_{11}} & \cdots & \boxed{a_{1n}} \\ \vdots & & \vdots \\ \boxed{a_{m1}} & \cdots & \boxed{a_{mn}} \end{pmatrix} \in \mathbb{R}^{n \times m}$$

ist die transponierte Matrix gegeben durch

$$\mathbf{A}^T = \begin{pmatrix} \boxed{a_{11}} & \cdots & \boxed{a_{m1}} \\ \vdots & & \vdots \\ \boxed{a_{1n}} & \cdots & \boxed{a_{mn}} \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

- Die Bezeichnung ist konsistent mit unserer Schreibweise $(a_1, \dots, a_n)^T$ für Spaltenvektoren
- Es gelten folgende Rechenregeln für alle $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R}$:

$$\begin{aligned} (\mathbf{A} + \mathbf{B})^T &= \mathbf{A}^T + \mathbf{B}^T, \\ (\alpha \mathbf{A})^T &= \alpha \mathbf{A}^T, \\ (\mathbf{A}^T)^T &= \mathbf{A}, \\ (\mathbf{A} \cdot \mathbf{B})^T &= \mathbf{B}^T \cdot \mathbf{A}^T. \end{aligned} \quad (8)$$

Definition 6.2.5 Eine $n \times n$ Matrix \mathbf{A} heißt **symmetrisch**, wenn $\mathbf{A}^T = \mathbf{A}$ gilt; sie heißt **schiefsymmetrisch**, falls $\mathbf{A}^T = -\mathbf{A}$ gilt.

- Das Transponieren einer Matrix macht aus einer Zeile eine Spalte. Sei $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times n}$. Dann gilt:

$$\begin{aligned} i\text{-te Zeile von } \mathbf{A} & \quad (a_{i1}, \dots, a_{in}), \\ i\text{-te Zeile von } \mathbf{A}^T & \quad (a_{1i}, \dots, a_{ni}). \end{aligned}$$

Somit gilt für symmetrische Matrizen:

$$\mathbf{A} = \mathbf{A}^T \Leftrightarrow a_{ij} = a_{ji} \quad \forall i, j = 1, \dots, n, \quad (9)$$

und analog gilt für schiefsymmetrische Matrizen:

$$\mathbf{A} = -\mathbf{A}^T \Leftrightarrow a_{ij} = -a_{ji} \quad \forall i, j = 1, \dots, n. \quad (10)$$

Insbesondere gilt für die Diagonalelemente einer schiefsymmetrischen Matrix $a_{ii} = 0$.

Definition 6.2.6 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} .

Satz 6.2.7 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:

$$\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}.$$

- Für $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc \neq 0$ gilt:

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Satz 6.2.8

- 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:

$$(\mathbf{A}_1 \cdot \dots \cdot \mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1} \cdot \mathbf{A}_{n-1}^{-1} \cdot \dots \cdot \mathbf{A}_1^{-1}. \quad (11)$$

- 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.

6.2.9 (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.

Satz 6.2.10 Eine obere (bzw. untere) Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonalelemente von Null verschieden sind.

- 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

Definition 6.3.1 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 kommutativ, d.h. für alle $\mathbf{a}, \mathbf{b} \in V$ gilt:

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

(V.2) 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.3) Es gibt ein Element $\mathbf{0} \in V$, Nullvektor genannt, mit

$$\mathbf{a} + \mathbf{0} = \mathbf{a}$$

für alle $\mathbf{a} \in V$.

(V.4) 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.5) $1 \cdot \mathbf{a} = \mathbf{a}$ für alle $\mathbf{a} \in V$.

(V.6) $\lambda(\mu \mathbf{a}) = (\lambda\mu) \mathbf{a}$ für alle $\lambda, \mu \in \mathbb{R}, \mathbf{a} \in V$.

(V.7) $\lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$ für alle $\lambda \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in V$.

$$(V.8) \quad (\mu + \lambda) \mathbf{a} = \mu \mathbf{a} + \lambda \mathbf{a} \quad \text{für alle } \lambda, \mu \in \mathbb{R}, \mathbf{a} \in V.$$

Die Elemente des Vektorraums V nennt man **Vektoren**.

Beispiele:

1) $V = \mathbb{R}^{m \times n}$ ist ein Vektorraum mit den komponentenweise definierten Operationen Addition und skalarer Multiplikation (siehe (5)).

2) Insbesondere sind

$$\begin{aligned} \mathbb{R}_n &= \{(a_1, \dots, a_n); a_i \in \mathbb{R}\} \\ \mathbb{R}^n &= \{(a_1, \dots, a_n)^T; a_i \in \mathbb{R}\} \end{aligned}$$

Vektorräume

3) Die Menge

$$P_n := \{a_0 + a_1x + \dots + a_nx^n; a_i \in \mathbb{R}, i = 0, \dots, n\}$$

der Polynome vom Grad $\leq n$ mit punktweise definierter Addition und skalarer Multiplikation ist ein \mathbb{R} -Vektorraum.

4) Die Menge $C^0(I) := \{f : I \rightarrow \mathbb{R}; f \text{ stetig}\}$ 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.

- Im folgenden sei V ein \mathbb{R} -Vektorraum.

Definition 6.3.2 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.:

$$(U.1) \quad \mathbf{u}, \mathbf{v} \in U \quad \Rightarrow \quad \mathbf{u} + \mathbf{v} \in U,$$

$$(U.2) \quad \mathbf{u} \in U, \lambda \in \mathbb{R} \quad \Rightarrow \quad \lambda \mathbf{u} \in U,$$

$$(U.3) \quad \mathbf{0} \in U.$$

- Man sagt, dass ein Unterraum U **abgeschlossen** bezüglich Addition und skalarer Multiplikation ist, wenn (U.1) und (U.2) gelten. Aus den Axiomen (V.1) – (V.8) folgt somit sofort, dass auch U ein Vektorraum ist.

Beispiele:

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 .

Definition 6.3.3 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 alle i gilt. Die Menge aller Linearkombinationen der \mathbf{v}_i heißt **lineare Hülle** der \mathbf{v}_i und wird mit

$$\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) := \left\{ \sum_{i=1}^k \alpha_i \mathbf{v}_i; \alpha_i \in \mathbb{R}, i = 1, \dots, k \right\}$$

bezeichnet.

Lemma 6.3.4 Die lineare Hülle der Vektoren $\mathbf{v}_i \in V, i = 1, \dots, k$, ist ein Unterraum von V .

Definition 6.3.5 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

$$U = \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

Beispiele:

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 6.3.6

a) 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.

b) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ heißen **linear unabhängig**, wenn sie nicht linear abhängig sind, d.h.

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \dots = \alpha_k = 0.$$

- 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

$$x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k = \mathbf{0} \tag{1}$$

betrachten (vergleiche (5)) und überprüfen, ob es nur triviale Lösungen gibt. Es gilt:

- i) $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear abhängig \Leftrightarrow (1) besitzt eine nichttriviale Lösung.
- ii) $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear unabhängig \Leftrightarrow (1) besitzt nur triviale Lösung.

- Sonderfälle von (1)

1) $k = 1$: $\mathbf{v} \in V$ ist linear abhängig $\Leftrightarrow (\alpha \mathbf{v} = \mathbf{0} \Rightarrow \alpha = 0) \Leftrightarrow \mathbf{v} \neq \mathbf{0}$.

2) $k = 2$: (\mathbf{u}, \mathbf{v}) sind linear abhängig $\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}$, nicht beide gleich Null, mit

$$\alpha \mathbf{u} + \beta \mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad \begin{cases} \mathbf{u} &= -\frac{\beta}{\alpha} \mathbf{v}, & \text{falls } \alpha \neq 0, \\ \mathbf{v} &= -\frac{\alpha}{\beta} \mathbf{u}, & \text{falls } \beta \neq 0. \end{cases}$$

Falls $V = \mathbb{R}^2, \mathbb{R}^3$ heißt dies, dass die Vektoren \mathbf{v}, \mathbf{u} parallel sind.

Lemma 6.3.7 Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ sind genau dann linear abhängig, wenn sich einer von ihnen als Linearkombination der anderen darstellen läßt.

Lemma 6.3.8

- a) Jedes endliche System von Vektoren, das linear abhängige Vektoren enthält, ist linear abhängig.
- b) Jedes endliche System von Vektoren, das den Nullvektor enthält, ist linear abhängig.
- c) Jedes Teilsystem linear unabhängiger Vektoren ist linear unabhängig.

Satz 6.3.9 In einer Matrix in Stufenform sind die nichttrivialen Zeilenvektoren linear unabhängig.

Satz 6.3.10 Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Dann sind die folgenden Aussagen äquivalent:

- a) \mathbf{A} ist invertierbar.
- b) Die Spalten von \mathbf{A} sind linear unabhängig.
- c) Die Zeilen von \mathbf{A} sind linear unabhängig.

Satz 6.3.11 Für Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w} \in V$ gilt:

$$a) \operatorname{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}) = \operatorname{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) \Leftrightarrow \mathbf{w} \in \operatorname{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k)$$

b) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ sind linear unabhängig \Leftrightarrow Zur Erzeugung von $\operatorname{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ kann kein $\mathbf{v}_i, i = 1, \dots, k$, weggelassen werden. In diesem Fall nennt man $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ ein **minimales Erzeugendensystem**.

Definition 6.3.12 Ein System $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ von Vektoren aus V heißt eine **Basis** des \mathbb{R} -Vektorraumes V , wenn gilt:

(B.1) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ sind linear unabhängig,

(B.2) Die \mathbf{v}_i erzeugen V , d.h. $\operatorname{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_n) = V$.

Satz 6.3.13 Ist $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ eine Basis von V , dann gibt es zu jedem Vektor $\mathbf{a} \in V$ genau ein n -Tupel reeller Zahlen $(\alpha_1, \dots, \alpha_n)$ mit

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n.$$

Ferner sind je m Vektoren aus V linear abhängig, falls $m > n$.

Satz 6.3.14 Die Zeilen (bzw. Spalten) einer invertierbaren $n \times n$ Matrix bilden eine Basis des \mathbb{R}_n (bzw. \mathbb{R}^n).

Definition 6.3.15 Ein Vektorraum V heißt **endlichdimensional**, wenn es endlich viele Vektoren $\mathbf{w}_1, \dots, \mathbf{w}_r$ mit $V = \operatorname{Lin}(\mathbf{w}_1, \dots, \mathbf{w}_r)$ gibt.

- Die Räume $\mathbb{R}^n, \mathbb{R}_n, P_{\mathbb{R}}$ sind endlichdimensional.
- Der Raum $P(\mathbb{R}) = \{\text{Polynom mit beliebigem Grad}\}$ ist nicht endlichdimensional.

Satz 6.3.16 (Basisergänzungssatz) In einem endlichdimensionalen Vektorraum $V \neq \{\mathbf{0}\}$ bilden linear unabhängige Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ bereits eine Basis $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ von V oder man kann sie durch Hinzunahme weiterer Vektoren $\mathbf{u}_1, \dots, \mathbf{u}_l$ zu einer Basis $(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_l)$ von V ergänzen.

Satz 6.3.17 Jeder endlichdimensionale Vektorraum $V \neq \{\mathbf{0}\}$ besitzt eine endliche Basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$. Ist $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ ebenfalls eine Basis von V , so gilt: $m = n$.

Definition 6.3.18 Die gemeinsame Länge n aller Basen eines endlichdimensionalen Vektorraumes $V \neq \{\mathbf{0}\}$ heißt **Dimension** von V , abgekürzt $\dim V = n$. Man setzt $\dim\{\mathbf{0}\} = 0$.

Beispiele:

$$\dim \mathbb{R}^n = \dim \mathbb{R}_n = n,$$

$$\dim P_k(\mathbb{R}) = k + 1,$$

$$\dim\{x \in \mathbb{R}^3, 3x_1 + x_2 + x_3 = 0\} = 2. \quad (\text{zwei Parameter frei wählbar})$$

Lemma 6.3.19 *Ist r die Maximalzahl linear unabhängiger Vektoren aus $(\mathbf{v}_1, \dots, \mathbf{v}_k)$, dann gilt:*

$$r = \dim \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

Satz 6.3.20 *In einem \mathbb{R} -Vektorraum V der Dimension n gilt:*

- Je n linear unabhängige Vektoren aus V bilden eine Basis von V .*
- Jedes Erzeugendensystem von V mit n Elementen ist eine Basis von V .*
- Je $n + 1$ Vektoren aus V sind linear abhängig.*

Satz 6.3.21 *Jeder Unterraum U eines endlichdimensionalen Vektorraumes V ist endlichdimensional. Im Falle $U \neq V$ gilt: $\dim U < \dim V$.*

• Nach Satz 6.3.21 ist ein Unterraum U eines endlichdimensionalen Vektorraumes wieder ein endlichdimensionaler Vektorraum. Also gelten die Sätze 6.3.16, 6.3.17, 6.3.20 analog für U .

• Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A} \neq \mathbf{0}$. Dann ist $\text{Ker } \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n$ ein Unterraum des \mathbb{R}^n der Dimension $r < n$.

6.4 Elementarmatrizen

Definition 6.4.1 *Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$ eine Matrix mit den Zeilenvektoren $\mathbf{z}_1, \dots, \mathbf{z}_m \in \mathbb{R}_n$ und den Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$. Dann heißt*

$$\text{Lin}(\mathbf{z}_1, \dots, \mathbf{z}_m) \subseteq \mathbb{R}_n$$

der **Zeilenraum** von \mathbf{A} und

$$\text{Lin}(\mathbf{a}_1, \dots, \mathbf{a}_n) \subseteq \mathbb{R}^m$$

der **Spaltenraum** von \mathbf{A} .

- Nach Lemma 6.3.19 ist die Dimension des Spaltenraumes die Maximalzahl linear unabhängiger Spaltenvektoren von \mathbf{A} . Analog für den Zeilenraum.
- Beim Transponieren gehen Zeilen in Spalten über und umgekehrt. Also spiegeln sich die Eigenschaften des Spaltenraumes (Zeilenraumes) von \mathbf{A} wider im Zeilenraum (Spaltenraum) von \mathbf{A}^T .

- Aus $\mathbf{A}\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i$, $\mathbf{x} \in \mathbb{R}^n$ bzw. $\mathbf{y}\mathbf{A} = \sum_{j=1}^m y_j \mathbf{a}_j$, $\mathbf{y} \in \mathbb{R}_m$ folgt

$$\begin{aligned} \text{Spaltenraum von } \mathbf{A} &= \{\mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n\}, \\ \text{Zeilenraum von } \mathbf{A} &= \{\mathbf{y}\mathbf{A}, \mathbf{y} \in \mathbb{R}_m\} = \{\mathbf{y}^T \mathbf{A}, \mathbf{y} \in \mathbb{R}^m\}. \end{aligned} \tag{1}$$

Satz 6.4.2 *Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$. Für alle invertierbaren Matrizen $\mathbf{P} \in \mathbb{R}^{m \times m}$ und $\mathbf{Q} \in \mathbb{R}^{n \times n}$ gilt:*

\mathbf{A} und $\mathbf{A}\mathbf{Q}$ haben denselben Spaltenraum, \mathbf{A} und $\mathbf{P}\mathbf{A}$ denselben Zeilenraum.

- Elementare Zeilen- bzw. Spaltenumformungen werden in folgende Typen eingeteilt:

Typ 1: Vertauschen zweier Zeilen (Spalten),

Typ 2: Multiplizieren einer Zeile (Spalte) mit Faktor ungleich Null,

Typ 3: Addition eines Vielfachen einer Zeile (Spalte) zu einer anderen.

Definition 6.4.3 Eine $m \times m$ Matrix $\tilde{\mathbf{E}}$ heißt **Elementarmatrix vom Typ i** , $i=1,2,3$, wenn sie aus der $m \times m$ Einheitsmatrix \mathbf{E} durch **eine** elementare Zeilenumformung vom Typ i hervorgeht.

Beispiele: für $m = 3$:

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Vertauschen von } \mathbf{e}_1 \text{ und } \mathbf{e}_2 \Rightarrow \text{Typ 1}$$

$$\begin{pmatrix} \mathbf{e}_1 \\ \alpha \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Multiplikation von } \mathbf{e}_2 \text{ mit } \alpha \Rightarrow \text{Typ 2}$$

$$\begin{pmatrix} \mathbf{e}_1 + \alpha \mathbf{e}_3 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Addition des } \alpha \text{ fachen von } \mathbf{e}_3 \text{ zu } \mathbf{e}_1 \Rightarrow \text{Typ 3}$$

Allgemein für $m \geq 2$:

- 1) $E_{1,i,j}$ ($i \neq j$): Entsteht aus der $m \times m$ Einheitsmatrix E durch Vertauschen der i -ten und j -ten Zeile.
- 2) $E_{2,i,\alpha}$ ($\alpha \neq 0$): Entsteht aus E , indem man die i -te Zeile durch ihr α -faches ersetzt.
- 3) $E_{3,i,j,\alpha}$ ($i \neq j$, $\alpha \neq 0$): Entsteht aus E , indem man zur i -Zeile das α -fache der j -ten addiert.

Satz 6.4.4

- a) Entsteht $\tilde{A} \in \mathbb{R}^{m \times n}$ durch eine (einzige) Zeilenumformung, dann gilt $\tilde{A} = \tilde{E}A$ mit der zugehörigen Elementarmatrix \tilde{E} .
- b) Die Elementarmatrizen sind invertierbar, die Inversen sind ebenfalls Elementarmatrizen. Insbesondere hat man

$$\begin{aligned} (E_{1,i,j})^{-1} &= E_{1,i,j}, & (E_{2,i,\alpha})^{-1} &= E_{2,i,\alpha^{-1}}, \\ (E_{3,i,j,\alpha})^{-1} &= E_{3,i,j,-\alpha}. \end{aligned}$$

- c) Entsteht \tilde{A} aus $A \in \mathbb{R}^{m \times n}$ durch die nacheinander ausgeführten Zeilenumformungen U_1, \dots, U_k und sind $\tilde{E}_1, \dots, \tilde{E}_k$ die zugehörigen Elementarmatrizen, dann gilt

$$\tilde{A} = P A \quad \text{mit} \quad P = \tilde{E}_k \cdots \tilde{E}_1, \quad P^{-1} = \tilde{E}_1^{-1} \cdots \tilde{E}_k^{-1}.$$

- d) Sei $A \in \mathbb{R}^{m \times m}$ invertierbar. Dann lässt sich A durch endlich viele elementare Zeilenumformungen U_1, \dots, U_k in $E \in \mathbb{R}^{m \times n}$ umwandeln. Sind $\tilde{E}_1, \dots, \tilde{E}_k$ die zugehörigen Elementarmatrizen, dann gilt

$$\tilde{E}_k \dots \tilde{E}_1 A = E, \quad \text{also} \quad A^{-1} = \tilde{E}_k \dots \tilde{E}_1.$$

Die gleichen Aussagen gelten sinngemäß für Spaltenumformungen. dann werden die Elementarmatrizen von rechts an A multipliziert.

Folgerung 6.4.5 Der Zeilenraum einer Matrix ändert sich nicht bei elementaren Zeilenumformungen, der Spaltenraum nicht bei elementaren Spaltenumformungen.

Folgerung 6.4.6 In einer Matrix in Stufenform bilden die nichttrivialen Zeilenvektoren eine Basis des Zeilenraumes. Die Dimension des Zeilenraumes ist die Anzahl der Elemente der Buchführungsmenge.

- Man kann statt Zeilenumformungen auch Spaltenumformungen durchführen und erhält so die sogenannte „Spaltenstufenform“, z.B.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ * & * & * & 0 & 0 \end{pmatrix}$$

Folgerung 6.4.7 In einer Matrix in „Spaltenstufenform“ bilden die nichttrivialen Spaltenvektoren eine Basis des Spaltenraumes. Die Dimension ist die Anzahl der Elemente der Buchführungsmenge.

- Wir bezeichnen mit

$$\begin{pmatrix} \mathbf{E}_s & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$$

die Matrix die an den ersten s Diagonalstellen eine 1 und sonst nur Nullen hat.

Satz 6.4.8

- a) Die Dimension des Spaltenraumes von \mathbf{A} ist gleich der Dimension des Zeilenraumes von \mathbf{A} .
- b) Es gibt invertierbare Matrizen $\mathbf{P} \in \mathbb{R}^{m \times m}$, $\mathbf{Q} \in \mathbb{R}^{n \times n}$, so dass

$$\mathbf{PAQ} = \begin{pmatrix} \mathbf{E}_r & 0 \\ 0 & 0 \end{pmatrix}, \quad (2)$$

wobei $r = \dim$ (Zeilenraumes von \mathbf{A})

- c) Gilt für invertierbare Matrizen $\mathbf{P} \in \mathbb{R}^{m \times m}$ und $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_n) \in \mathbb{R}^{n \times n}$ die Beziehung $\mathbf{PAQ} = \begin{pmatrix} \mathbf{E}_s & 0 \\ 0 & 0 \end{pmatrix}$, dann ist $s = r$ und $(\mathbf{q}_{r+1}, \dots, \mathbf{q}_n)$ bilden eine Basis von Kern \mathbf{A} .

Definition 6.4.9 Der **Rang** einer Matrix ist die Dimension ihres Zeilenraumes und wird mit $\text{Rang } \mathbf{A}$ bezeichnet.

Folgerung 6.4.10 Für jede $m \times n$ Matrix \mathbf{A} gilt:

$$\text{Rang } \mathbf{A} + \dim(\text{Ker } \mathbf{A}) = n.$$

Folgerung 6.4.11 Für $\mathbf{A} \in \mathbb{R}^{m \times n}$ gilt:

$$\text{Rang}(\mathbf{A}) = \text{Rang}(\mathbf{A}^T), \quad (\text{i})$$

und für alle invertierbaren Matrizen $\mathbf{P} \in \mathbb{R}^{m \times m}$, $\mathbf{Q} \in \mathbb{R}^{n \times n}$

$$\text{Rang}(\mathbf{PAQ}) = \text{Rang}(\mathbf{A}). \quad (\text{ii})$$

6.5 Determinanten

6.5.1 (Zweireihige Determinanten) Das von zwei Vektoren $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ der (x, y) -Ebene erzeugte Parallelogramm hat den Flächeninhalt (siehe Kapitel 1, 1.5.5)

$$F = \left| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right| = |a_1 b_2 - a_2 b_1|.$$

Die Zahl

$$\det \mathbf{A} := a_1 b_2 - a_2 b_1 \quad (1)$$

heißt **Determinante** der Matrix $\mathbf{A} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$. Man berechnet sie nach dem Schema

$$\det \begin{pmatrix} \overset{+}{a_1} & & b_1 \\ & \times & \\ \underset{-}{a_2} & & b_2 \end{pmatrix} = a_1 b_2 - a_2 b_1,$$

und es gilt $\det \mathbf{A} = \pm F$. Also haben wir:

$$\det \mathbf{A} = \pm F = 0 \Leftrightarrow \mathbf{a}, \mathbf{b} \text{ parallel oder } (\mathbf{a} = \mathbf{0} \text{ oder } \mathbf{b} = \mathbf{0}),$$

anders gesagt

$$\begin{aligned} \det \mathbf{A} = 0 &\Leftrightarrow \mathbf{a}, \mathbf{b} \text{ sind linear abhängig} \\ &\Leftrightarrow \mathbf{A} \text{ ist nicht invertierbar.} \end{aligned}$$

6.5.2 (Dreireihige Determinanten) Der Spat dreier Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ hat das Volumen (Kapitel 1, Satz 1.5.7)

$$V = |[\mathbf{a}, \mathbf{b}, \mathbf{c}]|$$

$$= |a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)|$$

Die Zahl

$$\det \mathbf{A} = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$$

ist die **Determinante** der Matrix $\mathbf{A} = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$. Das Volumen des Spats ist Null, wenn er entartet ist, und somit sind die Vektoren linear abhängig sind, d.h.

$$\det \mathbf{A} = 0 \Leftrightarrow \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ linear abhängig}$$

oder anders gesagt

$$\begin{aligned} \det \mathbf{A} \neq 0 &\Leftrightarrow \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ linear unabhängig,} \\ &\Leftrightarrow \mathbf{A} \text{ invertierbar,} \\ &\Leftrightarrow \text{Rang } \mathbf{A} = 3. \end{aligned}$$

Auf Grund der Definition des Spatproduktes hat man für $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{3 \times 3}$

$$\det \mathbf{A} = a_{11} \det \begin{pmatrix} a_{12} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}, \quad (2)$$

d.h. die Berechnung einer Determinante einer 3×3 Matrix kann auf die Berechnung von Determinanten von 2×2 Matrizen zurückgeführt werden.

Definition 6.5.3 Sei $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times n}$. Wir definieren die **Determinante** $\det \mathbf{A}$ rekursiv durch.

$$n = 1 : \det \mathbf{A} := a_{ij}$$

$$\begin{aligned} n \geq 2 : \det \mathbf{A} &:= a_{11} \det \mathbf{A}_{11} - a_{21} \det \mathbf{A}_{21} + \dots + (-1)^{n+1} a_{n1} \det \mathbf{A}_{n1}, \\ &\text{wobei } \mathbf{A}_{i1} \text{ die } (n-1) \times (n-1) \text{ Matrix ist, die aus } \mathbf{A} \text{ durch Entfernen der 1.} \\ &\text{Spalte und der } i\text{-ten Zeile entsteht.} \end{aligned}$$

Beispiel:

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 6 & 3 & 2 \\ 2 & 4 & 3 & 1 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} -2 & 1 & 0 \\ 6 & 3 & 2 \\ 4 & 3 & 1 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ 6 & 3 & 2 \\ 4 & 3 & 1 \end{pmatrix} \\ &\quad + 0 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 0 \\ 6 & 3 & 2 \end{pmatrix} \\ &= -2(3 \cdot 1 - 2 \cdot 3) - 6(1 \cdot 1 - 0 \cdot 3) + 4(1 \cdot 2 - 0 \cdot 3) \\ &\quad - 3[2(3 \cdot 1 - 2 \cdot 3) - 6(0 \cdot 1 - 1 \cdot 3) + 4(0 \cdot 2 - 1 \cdot 3)] \\ &\quad - 2[2(1 \cdot 2 - 0 \cdot 3) + 2(0 \cdot 2 - 1 \cdot 3) + 6(0 \cdot 0 - 1 \cdot 1)] \\ &= -46 \end{aligned}$$

- Wenn man die rekursive Definition 6.5.3 vollständig entwickelt erhält man:

$$\det \mathbf{A} = \sum_{i=i_1, \dots, i_n} (-1)^{\varepsilon(i)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}, \quad (3)$$

wobei man über alle Permutationen i der Menge $\{1, 2, \dots, n\}$ summiert und $\varepsilon(i)$ die Anzahl der Vertauschungen ist, die man benötigt um i in $\{1, 2, \dots, n\}$ zu bringen. Man benötigt also $n!$ Summanden!

Satz 6.5.4 Die Determinante einer oberen Dreiecksmatrix $\mathbf{A} = (a_{ij})$ ist das Produkt seiner Diagonalelemente a_{ii} , d.h.

$$\det \mathbf{A} = \prod_{i=1}^n a_{ii}$$

Insbesondere gilt: $\det \mathbf{E} = 1$, $\det(\alpha \mathbf{E}) = \alpha^n$.

Satz 6.5.5 a) \det ist **linear in jeder Zeile**, d.h.

$$\det \begin{pmatrix} z_1 \\ \vdots \\ \alpha z_i \\ \vdots \\ z_n \end{pmatrix} = \alpha \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix}, \quad \det \begin{pmatrix} z_1 \\ \vdots \\ a+b \\ \vdots \\ z_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ a \\ \vdots \\ z_n \end{pmatrix} + \det \begin{pmatrix} z_1 \\ \vdots \\ b \\ \vdots \\ z_n \end{pmatrix}.$$

b) \det ist **alternierend**, d.h.

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ z_n \end{pmatrix} = - \det \begin{pmatrix} z_1 \\ \vdots \\ z_j \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix},$$

insbesondere ist $\det \mathbf{A} = 0$ falls zwei Zeilen von \mathbf{A} identisch sind.

Satz 6.5.6 Seien $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$.

a) Entsteht $\tilde{\mathbf{A}}$ aus \mathbf{A} durch eine elementare Zeilen- bzw. Spaltenumformung vom Typ i , dann gilt:

$$\begin{aligned} \text{Typ 1:} & \quad \det \tilde{\mathbf{A}} = - \det \mathbf{A}, \\ \text{Typ 2:} & \quad \det \tilde{\mathbf{A}} = \alpha \det \mathbf{A} \quad \alpha \neq 0, \\ \text{Typ 3:} & \quad \det \tilde{\mathbf{A}} = \det \mathbf{A}. \end{aligned} \quad (4)$$

b) Es bestehen die folgenden Aussagen:

$$\det \mathbf{A} = \det \mathbf{A}^T,$$

insbesondere bleibt Satz 6.5.5 richtig mit „Spalten“ statt „Zeilen“.

c)

$$\det(\mathbf{A} \cdot \mathbf{B}) = \det \mathbf{A} \cdot \det \mathbf{B}. \quad (5)$$

d)

$$\begin{aligned} \mathbf{A} \text{ invertierbar} &\Leftrightarrow \det \mathbf{A} \neq 0, \\ \text{Rang } \mathbf{A} < n &\Leftrightarrow \det \mathbf{A} = 0. \end{aligned}$$

Folgerung 6.5.7 Für $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ haben wir

a) $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{B} \cdot \mathbf{A}),$

b) $\det(\mathbf{A}^k) = (\det \mathbf{A})^k, \quad k \in \mathbb{N},$

c) $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1},$ falls \mathbf{A} invertierbar,

d) $\det(\mathbf{B}^{-1} \mathbf{A} \mathbf{B}) = \det \mathbf{A},$ falls \mathbf{B} invertierbar,

e) falls $\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$ oder $\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ mit Unter- oder Blockmatrizen $\mathbf{B} \in \mathbb{R}^{k \times k}, \mathbf{D} \in \mathbb{R}^{(n-k) \times (n-k)},$ so gilt:

$$\det \mathbf{A} = \det \mathbf{B} \cdot \det \mathbf{D}.$$

6.5.8 Die Matrix $\tilde{\mathbf{A}} = (\mathbf{a}_j, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n)$ geht aus $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ durch $(j-1)$ sukzessive Vertauschungen von Spalten hervor. Also gilt

$$\det \tilde{\mathbf{A}} = (-1)^{j-1} \det \mathbf{A}.$$

Entwickelt man $\det \tilde{\mathbf{A}}$ nach Definition 6.5.3 erhält man die **Entwicklung von $\det \mathbf{A}$ nach der j -ten Spalte**

$$\det \mathbf{A} = \sum_{i=1}^n (-1)^{j+i} a_{ij} \det \mathbf{A}_{ij},$$

wobei \mathbf{A}_{ij} die $(n-1) \times (n-1)$ Matrix ist, die aus \mathbf{A} durch Wegstreichen der i -ten Zeile und der j -ten Spalte entsteht. Durch Transposition erhält man die **Entwicklung von $\det \mathbf{A}$ nach der i -ten Zeile**

$$\det \mathbf{A} = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij}.$$

Bei der Berechnung von $\det \mathbf{A}, \mathbf{A} \in \mathbb{R}^{n \times n}, n \geq 3$ sollte man die Struktur der Matrix beachten (evtl. Blockmatrix) und durch Zeilen- und Spaltenumformungen möglichst viele Nullen erzeugen.

Beispiel:

$$\begin{aligned} \det \begin{pmatrix} 1 & 3 & -5 & 1 & 4 \\ 4 & 2 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & -1 & 3 & 3 \end{pmatrix} &= \det \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ -1 & 3 & 3 \end{pmatrix} \\ &= -10 \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 5 & 6 \end{pmatrix} \\ &= -10 \cdot (-1) = 10 \end{aligned}$$

- **Cramersche Regel** (Gabriel C., 1704–1752):

Sei \mathbf{A} invertierbar. Dann hat das System $\mathbf{Ax} = \mathbf{b}$ die Lösung $\mathbf{x} = (x_1, \dots, x_n)^T$ mit

$$x_i = \frac{1}{\det \mathbf{A}} \det (\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n).$$

6.6 Lineare Abbildungen und Eigenwerte

Seien V, W Vektorräume über \mathbb{R} . Eine **Abbildung** $f : V \rightarrow W$ ordnet jedem Vektor $\mathbf{v} \in V$ einen eindeutig bestimmten Vektor $\mathbf{w} = f(\mathbf{v}) \in W$ zu.

Definition 6.6.1 Eine Abbildung $f : V \rightarrow W$ heißt **linear**, wenn gilt:

(L1) f ist homogen, d.h. $f(\alpha \mathbf{v}) = \alpha f(\mathbf{v}) \quad \forall \alpha \in \mathbb{R}, \mathbf{v} \in V$,

(L2) f ist additiv, d.h. $f(\mathbf{v} + \mathbf{u}) = f(\mathbf{u}) + f(\mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v} \in V$.

- (L1), (L2) sind äquivalent zu

$$\begin{aligned} f(\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n) &= \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_n f(\mathbf{v}_n) \\ &\forall \alpha_i \in \mathbb{R}, \mathbf{v}_i \in V. \end{aligned} \tag{1}$$

Beispiele:

- Die Multiplikation mit einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ liefert die lineare Abbildung $\ell_{\mathbf{A}} : \mathbb{R}^n \rightarrow \mathbb{R}^n : \mathbf{x} \rightarrow \mathbf{Ax}$ (siehe Satz 6.2.3)
- Der Differentialoperator $\frac{d}{dx} : C^1(\mathbb{R}) \rightarrow C(\mathbb{R}) : f \mapsto f'$ ist linear (siehe Kapitel 3, Satz 3.1.5).
- Das Integral $I : R([a, b]) \rightarrow \mathbb{R} : f \mapsto \int_a^b f(x) dx$, wobei $R([a, b])$ die Riemann integrierbaren Funktionen über $[a, b]$ sind, ist linear. (siehe Kapitel 4 Satz 4.1.3)

6.6.2 (Operationen mit linearen Abbildungen) Seien $f, g : V \rightarrow W$ lineare Abbildungen und $\alpha \in \mathbb{R}$. Wir definieren punktweise die Abbildungen

$$f + g : V \rightarrow W, \quad \alpha f : V \rightarrow W,$$

d.h.

$$\begin{aligned} (f + g)(v) &:= f(v) + g(v), \\ (\alpha f)(v) &:= \alpha f(v). \end{aligned}$$

Mit diesen Operationen ist die Menge der **Homomorphismen**

$$\text{Hom}(V, W) := \{f : V \rightarrow W \text{ linear}\}$$

wieder ein Vektorraum. Weiterhin ist auch die **Komposition** zweier linearer Abbildungen wieder linear, d.h. für $f : V \rightarrow W$ und $g : U \rightarrow V$ ist auch

$$f \circ g : U \rightarrow W \quad (f \circ g)(v) := f(g(v))$$

linear. Sei V ein endlichdimensionaler Vektorraum und $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ eine Basis von V . Nach Satz 6.3.13 existiert für jeden Vektor $\mathbf{v} \in V$ genau ein Vektor $\mathbf{a} = (a_1, \dots, a_n)^T \in \mathbb{R}^n$ mit $\mathbf{v} = \sum_{i=1}^n a_i \mathbf{v}_i$. Die Abbildung

$$k_{\mathbf{B}} : V \rightarrow \mathbb{R}^n : \mathbf{v} \mapsto \mathbf{a}$$

heißt **Koordinatenabbildung bezüglich der Basis \mathbf{B}** . Sie ist linear und mit ihr überträgt man Rechnungen im V in Rechnungen im \mathbb{R}^n und umgekehrt.

Satz 6.6.3 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Dann gilt für die Matrix $\mathbf{M}(f) := (f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)) \in \mathbb{R}^{m \times n}$, deren Spalten die Bilder der Standardbasisvektoren sind, folgende Beziehung:

$$\mathbf{M}(f)\mathbf{x} = f(\mathbf{x}).$$

Man nennt $\mathbf{M}(f)$ die **Abbildungsmatrix** von f bezüglich der Standardbasis.

Lemma 6.6.4 Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ zwei lineare Abbildungen mit Abbildungsmatrizen $\mathbf{M}(f), \mathbf{M}(g)$. Dann gilt:

- (i) $\mathbf{M}(\alpha f) = \alpha \mathbf{M}(f), \quad \alpha \in \mathbb{R}$
- (ii) $\mathbf{M}(f + g) = \mathbf{M}(f) + \mathbf{M}(g)$
- (iii) $\mathbf{M}(f \circ g) = \mathbf{M}(f) \cdot \mathbf{M}(g)$

Ist f **invertierbar**, d.h. zu jedem $\mathbf{y} \in \mathbb{R}^n$ gibt es genau ein $\mathbf{x} \in \mathbb{R}^n$ mit $f(\mathbf{x}) = \mathbf{y}$, so wird die inverse Abbildung mit f^{-1} bezeichnet und es gilt:

$$(iv) \mathbf{M}(f^{-1}) = (\mathbf{M}(f))^{-1}$$

6.6.5 (Änderung des Volumens) Seien $\mathbf{b}_1, \dots, \mathbf{b}_n$ n Vektoren aus \mathbb{R}^n . Ein n -**Spat** ist die Menge aller Linearkombinationen

$$\sum_{i=1}^n \lambda_i \mathbf{b}_i \quad 0 \leq \lambda_i \leq 1, i = 1, \dots, n.$$

In Analogie zum \mathbb{R}^3 bezeichnen wir mit

$$V = |\det(\mathbf{b}_1, \dots, \mathbf{b}_n)|$$

das Volumen des von $\mathbf{b}_1, \dots, \mathbf{b}_n$ aufgespannten n -Spats. Wegen

$$\begin{aligned} \det(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)) &= \det(\mathbf{M}(f) \mathbf{b}_1, \dots, \mathbf{M}(f) \mathbf{b}_n) \\ &= \det(\mathbf{M}(f) \text{ lr } \mathbf{b}_1, \dots, \mathbf{b}_n) \\ &= \det \mathbf{M}(f) \det(\mathbf{b}_1, \dots, \mathbf{b}_n) \end{aligned}$$

gibt $|\det(\mathbf{M}(f))|$ die Volumenverzerrung der linearen Abbildung an.

Definition 6.6.6 Seien $\mathbf{x} = (x_1, \dots, x_n)^T, \mathbf{y} = (y_1, \dots, y_n)^T$ Vektoren des \mathbb{R}^n . Die Zahl

$$\mathbf{x} \cdot \mathbf{y} := \sum_{i=1}^n x_i y_i$$

heißt **Skalarprodukt** von \mathbf{x} und \mathbf{y} , und

$$\|\mathbf{x}\| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

heißt **Länge** des Vektors \mathbf{x} . Ein Vektor \mathbf{x} heißt **Einheitsvektor**, wenn $\|\mathbf{x}\| = 1$.

Lemma 6.6.7 Für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \alpha \in \mathbb{R}$ gilt:

- i) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
- ii) $\alpha(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x}(\alpha \cdot \mathbf{y}) = \mathbf{y}(\alpha \cdot \mathbf{x})$
- iii) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$
- iv) $\mathbf{x} \cdot \mathbf{x} > 0$ für alle $\mathbf{x} > 0$
- v) $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
- vi) $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
- vii) $\|\mathbf{x} \cdot \mathbf{y}\| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ (Cauchy-Schwarzsche Ungleichung)
- viii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (Dreiecksungleichung)

Definition 6.6.8 Für $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ nennt man $\angle(\mathbf{x}, \mathbf{y}) := \varphi = \arccos \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$, $0 \leq \varphi \leq \pi$, den **Winkel** zwischen \mathbf{x} und \mathbf{y} . Man nennt $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ **orthogonal** falls $\mathbf{x} \cdot \mathbf{y} = 0$, d.h. $\varphi = \frac{\pi}{2}$. Der Nullvektor $\mathbf{0}$ steht orthogonal zu allen Vektoren des \mathbb{R}^n .

Definition 6.6.9

a) Eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heisst **orthogonal**, wenn sie das Skalarprodukt invariant lässt, d.h. es gilt

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

b) Eine Matrix heisst **orthogonal**, wenn gilt:

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{E}, \text{ d.h. } \mathbf{A}^{-1} = \mathbf{A}^T.$$

c) Eine Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ heisst **orthogonal**, wenn die Vektoren \mathbf{b}_i paarweise senkrecht aufeinander stehen, d.h.

$$\mathbf{b}_i \cdot \mathbf{b}_j = 0 \quad i \neq j = 1, \dots, n.$$

Die Basis heisst **orthonormal**, wenn sie orthogonal ist und aus Einheitsvektoren besteht, d.h.

$$\mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij} \quad \forall i, j = 1, \dots, n.$$

Satz 6.6.10 Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$. Dann sind äquivalent:

- (i) \mathbf{A} ist orthogonal,
 - (ii) $(\mathbf{A}\mathbf{x}) \cdot (\mathbf{A}\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
 - (iii) Die Spalten von \mathbf{A} sind eine Orthonormalbasis des \mathbb{R}^n .
- Insbesondere ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ orthogonal genau dann, wenn $\mathbf{M}(f) \in \mathbb{R}^{n \times n}$ orthogonal ist.

- $\mathbf{A}^T \mathbf{A} = \mathbf{E} \Leftrightarrow \mathbf{A} \mathbf{A}^T = \mathbf{E} \Leftrightarrow$ Zeilen sind Orthonormalbasis.

- Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ orthogonal. Dann ist f

- i) **Volumentreu**, d.h. $|\det \mathbf{M}(f)| = 1$,

- ii) **längentreu**, d.h. $\|f(\mathbf{x})\| = \|\mathbf{x}\|$,

- iii) **winkeltreu**, d.h. $\angle(f(\mathbf{x}), f(\mathbf{y})) = \angle(\mathbf{x}, \mathbf{y})$.

6.6.11 (Spiegelungen im \mathbb{R}^3) Eine Punktspiegelung ist durch die Vorschrift

$$s : \mathbf{x} \mapsto -\mathbf{x}$$

gegeben. Es gilt $\mathbf{M}(s) = -\mathbf{E}$.

Sei $\mathbf{a} = (a_1, a_2, a_3)^T$ ein Einheitsvektor. Durch die Gleichung $\mathbf{a} \cdot \mathbf{x} = 0$ wird eine zu \mathbf{a} orthogonale Ebene aufgespannt. Die Spiegelung an dieser Ebene ist gegeben durch

$$s : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : \mathbf{x} \mapsto \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{a}) \mathbf{a}.$$

Man rechnet aus, dass

$$s(\mathbf{x}) \cdot s(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$$

$$s(s(\mathbf{x})) = \mathbf{x}$$

und daher gilt (siehe Satz 6.6.10 und Lemma 6.6.4)

$$\mathbf{M}(s) = \mathbf{M}(s)^{-1} = \mathbf{M}(s)^T,$$

d.h. die Abbildungsmatrix $\mathbf{M}(s)$ ist eine symmetrische, orthogonale Matrix mit $\det(\mathbf{M}(s)) = -1$, für die gilt:

$$\mathbf{M}(s) = \mathbf{E} - 2\mathbf{a}\mathbf{a}^T.$$

6.6.12 (Drehungen im Raum) Eine Drehung um eine Koordinatenachse lässt stets eine Komponente unverändert und transformiert die beiden anderen Komponenten entsprechend der Formel für die Drehung einer Ebene (Kapitel 1, Satz 1.3.6). Somit erhält man z.B. für die Drehung um die z -Achse die Abbildungsmatrix

$$\mathbf{M}(\mathbf{D}_3) = \begin{pmatrix} \cos \alpha_3 & -\sin \alpha_3 & 0 \\ \sin \alpha_3 & \cos \alpha_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Im Allgemeinen sei $\mathbf{a} \neq \mathbf{0}$ ein Richtungsvektor und $\varphi \geq 0$ ein Drehwinkel, dann ist die entsprechende Drehung gegeben durch:

$$d(\mathbf{x}) = (\cos \varphi) \mathbf{x} + (1 - \cos \varphi) \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} + \frac{\sin \varphi}{\|\mathbf{a}\|} \mathbf{a} \times \mathbf{x}$$

Man kann zeigen, dass $\mathbf{D} \in \mathbb{R}^{3 \times 3}$ die Abbildungsmatrix einer Drehung ist genau dann, wenn

$$\mathbf{D}^T \cdot \mathbf{D} = \mathbf{E} \text{ und } \det \mathbf{D} = 1.$$

6.6.13 (Das Gram-Schmidtsche Orthonomierungsverfahren) (Jorgen Pedersen G., 1850–1916; Erhard S., 1876–1959).

Seien $\mathbf{b}_1, \dots, \mathbf{b}_k \in \mathbb{R}^n$, ($k \leq n$), linear unabhängige Vektoren. Man berechnet eine Orthonormalbasis $\mathbf{c}_1, \dots, \mathbf{c}_k$ der linearen Hülle $\text{Lin}(\mathbf{b}_1, \dots, \mathbf{b}_k)$ wie folgt:

- Setze $\mathbf{c}_1 = \frac{1}{\|\mathbf{b}_1\|} \mathbf{b}_1$.
- Berechne die zu \mathbf{c}_1 orthogonale Komponente \mathbf{c}'_2 von \mathbf{b}_2 , d.h.

$$\mathbf{c}'_2 = \mathbf{b}_2 - (\mathbf{b}_2 \cdot \mathbf{c}_1) \mathbf{c}_1.$$

- Normiere \mathbf{c}'_2 und setze $\mathbf{c}_2 := \frac{1}{\|\mathbf{c}'_2\|} \mathbf{c}'_2$.
- Berechne die zu \mathbf{c}_1 und \mathbf{c}_2 orthogonale Komponente \mathbf{c}'_3 von \mathbf{b}_3 , d.h.

$$\mathbf{c}'_3 = \mathbf{b}_3 - (\mathbf{b}_3 \cdot \mathbf{c}_1) \mathbf{c}_1 - (\mathbf{b}_3 \cdot \mathbf{c}_2) \mathbf{c}_2,$$

und normiere \mathbf{c}'_3 , d.h. setze

$$\mathbf{c}_3 = \frac{1}{\|\mathbf{c}'_3\|} \mathbf{c}'_3,$$

usw.

Es folgt

$$\text{Lin}(\mathbf{c}_1, \dots, \mathbf{c}_r) = \text{Lin}(\mathbf{b}_1, \dots, \mathbf{b}_r) \quad r = 1, \dots, k.$$

Definition 6.6.14 Seien $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ und $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_n)$ Basen des \mathbb{R}^n . Die **Basiswechselmatrix** $\mathbf{M}_S^B(id)$ von der Basis \mathbf{B} zur Basis \mathbf{S} ist diejenige Matrix, deren j -te Spalte aus den \mathbf{S} -Koordinaten des j -ten Basisvektors von \mathbf{B} besteht.

- $id : (\mathbb{R}^n, \mathbf{B}) \rightarrow (\mathbb{R}^n, \mathbf{S})$ „Bilder der alten Basisvektoren in die Spalten“

$$\mathbf{M}_S^B(id) = (k_S(\mathbf{b}_1), \dots, k_S(\mathbf{b}_n)). \quad (2)$$

Satz 6.6.15 Sei $\mathbf{M}_S^B(id)$ die Basiswechselmatrix für den Übergang von der Basis \mathbf{B} zur Basis \mathbf{S} . Dann gilt für alle $\mathbf{v} \in \mathbb{R}^n$:

$$k_S(\mathbf{v}) = \mathbf{M}_S^B(id) k_B(\mathbf{v}), \quad (3)$$

d.h. man erhält die \mathbf{S} -Koordinaten von \mathbf{v} indem man $\mathbf{M}_S^B(id)$ mit den \mathbf{B} -Koordinaten von \mathbf{v} multipliziert.

- Im Spezialfall dass $\mathbf{S} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ die Standardbasis ist, liefert (2)

$$\mathbf{M}_S^B(id) = (\mathbf{b}_1, \dots, \mathbf{b}_n) =: \mathbf{B},$$

und (3)

$$k_S(\mathbf{v}) = \mathbf{B} k_B(\mathbf{v}). \quad (4)$$

\mathbf{B} ist invertierbar und es gilt

$$k_B(\mathbf{v}) = \mathbf{B}^{-1} k_S(\mathbf{v}), \quad (5)$$

d.h. \mathbf{B}^{-1} ist die Basiswechselmatrix von der Standardbasis \mathbf{S} zur Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$.

Beispiel: Seien $\mathbf{S} = (\mathbf{e}_1, \mathbf{e}_2)$, $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2)$ zwei Basen des \mathbb{R}^2 , mit

$$\mathbf{b}_1 = (3, 4)^T, \quad \mathbf{b}_2 = (-1, 2)^T.$$

a) Übergang $\mathbf{B} \rightarrow \mathbf{S}$. Die \mathbf{S} -Koordinaten von \mathbf{b}_i in die Spalten

$$\mathbf{M}_S^B(id) = \begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix}.$$

Sei $k_B(\mathbf{v}) = (1, 3)^T$ die \mathbf{B} -Koordinaten von \mathbf{v} . Welcher Vektor ist dies?

$$k_S(\mathbf{v}) = \mathbf{M}_S^B(id) \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 10 \end{pmatrix}.$$

b) Übergang $\mathbf{S} \rightarrow \mathbf{B}$. Aus (5) folgt:

$$\mathbf{M}_B^S(id) = (\mathbf{M}_B^S(id))^{-1} = \frac{1}{10} \begin{pmatrix} 2 & 1 \\ -4 & 3 \end{pmatrix}.$$

Also sind die \mathbf{B} -Koordinaten von $\mathbf{v} = (1, 1)^T$

$$k_B(\mathbf{v}) = \frac{1}{10} \begin{pmatrix} 2 & 1 \\ -4 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 3 \\ -3 \end{pmatrix}.$$

6.6.16 (Ortsvektoren) Die Zeilen- n -tupel (x_1, \dots, x_n) , $x_i \in \mathbb{R}$, nennt man **Punkte** im n -dimensionalen **Punktraum** \mathbb{R}^n . Zu jedem Punkt $\mathbf{X} = (x_1, \dots, x_n)$ gehört ein eindeutig bestimmter Vektor $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ und umgekehrt. Ist $\mathbf{P} \in \mathbb{R}^n$ ein fester Punkt, so ist

$$\overrightarrow{\mathbf{P}\mathbf{X}} = \mathbf{x} - \mathbf{p} = \begin{pmatrix} x_1 - p_1 \\ \vdots \\ x_n - p_n \end{pmatrix} \quad (6)$$

der **Ortsvektor** des Punktes \mathbf{X} bezüglich \mathbf{P} . Insbesondere ist \mathbf{x} der Ortsvektor von \mathbf{X} bezüglich des Nullpunktes $\mathbf{0} = (0, \dots, 0)$.

Definition 6.6.17 Ein affines Koordinatensystem $K = (\mathbf{P}, \mathbf{b}_1, \dots, \mathbf{b}_n)$ des Punktraumes \mathbb{R}^n besteht aus einem festen Bezugspunkt $\mathbf{P} = (p_1, \dots, p_n)$ und einer Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ des \mathbb{R}^n . Die Koordinaten x'_i des Ortsvektors $\overrightarrow{\mathbf{P}\mathbf{X}}$ bezüglich der Basis \mathbf{B} , d.h.

$$\overrightarrow{\mathbf{P}\mathbf{X}} = \sum_{i=1}^n x'_i \mathbf{b}_i,$$

heißen **Koordinaten des Punktes \mathbf{X} im Koordinatensystem K** , man schreibt $\mathbf{X}_K = (x'_1, \dots, x'_n) = \left(k_B \left(\overrightarrow{\mathbf{P}\mathbf{X}} \right) \right)^T$.

• Nach (5),(6) gilt

$$\begin{aligned} k_B \left(\overrightarrow{\mathbf{P}\mathbf{X}} \right) &= \mathbf{B}^{-1} k_S(\mathbf{x} - \mathbf{p}) \\ \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} &= \mathbf{B}^{-1} \begin{pmatrix} x_1 - p_1 \\ \vdots \\ x_n - p_n \end{pmatrix} \end{aligned} \quad (7)$$

und umgekehrt

$$\begin{aligned}
 k_S(\mathbf{x} - \mathbf{p}) &= \mathbf{B}k_B(\overrightarrow{\mathbf{P}\mathbf{X}}), \text{ d.h.} \\
 k_S(\mathbf{x}) &= \mathbf{B}k_B(\overrightarrow{\mathbf{P}\mathbf{X}}) + k_S(\mathbf{P}) \\
 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= \mathbf{B} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}.
 \end{aligned} \tag{8}$$

Beispiel: Seien $(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2)$ die kartesischen Koordinaten und $(\mathbf{P}, \mathbf{e}_r, \mathbf{e}_\varphi)$ Polarkoordinaten, wobei

$$\begin{aligned}
 k_S(\mathbf{e}_r) &= (\cos \varphi, \sin \varphi)^T, \\
 k_S(\mathbf{e}_\varphi) &= (-r \sin \varphi, r \cos \varphi)^T, \\
 \mathbf{M}_S^P(\varphi, r) &= \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.
 \end{aligned}$$

Wir wählen $r = 1$, $\varphi = \frac{\pi}{4}$ und $k_P(\mathbf{v}) = (1, 1)^T$ und erhalten

$$k_S(\mathbf{v}) = \mathbf{M}_S^P(\varphi, r) k_P(\mathbf{v}) + k_S(\mathbf{P}) = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Umgekehrt sei $r = 1$, $\varphi = \frac{\pi}{4}$ und $k_S(\mathbf{v}) = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{3}{2}\sqrt{2} \end{pmatrix}$ und $k_S(\mathbf{P}) = \frac{1}{2}\sqrt{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Wir haben

$$\begin{aligned}
 \mathbf{M}_S^P\left(\frac{\pi}{4}, 1\right) &= \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \\
 \mathbf{M}_P^S\left(\frac{\pi}{4}, 1\right) &= \left(\mathbf{M}_S^P\left(\frac{\pi}{4}, 1\right)\right)^{-1} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.
 \end{aligned}$$

Also gilt:

$$k_P(\mathbf{v}) = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} v_1 - p_1 \\ v_2 - p_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Definition 6.6.18 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis des \mathbb{R}^n . Die Matrix

$$\mathbf{M}_B^B(f) := (k_B(f(\mathbf{b}_1)), \dots, k_B(f(\mathbf{b}_n))),$$

in deren Spalten die \mathbf{B} -Koordinaten von $f(\mathbf{b}_j)$ stehen heißt **Abbildungsmatrix von f bezüglich \mathbf{B}** .

- Für die Standardbasis schreiben wir $\mathbf{M}_S^S(f)$ statt $\mathbf{M}(f)$ (vergleiche Satz 6.6.3).
- Es gilt: $k_B(f(\mathbf{x})) = \mathbf{M}_B^B(f) k_B(\mathbf{x})$.

Satz 6.6.19 Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung und sei $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von \mathbb{R}^n . Es gilt:

$$\begin{aligned} \mathbf{M}_B^B(f) &= \mathbf{M}_B^S(id) \mathbf{M}_S^S(f) \mathbf{M}_S^B(id) \\ &= \mathbf{B}^{-1} \mathbf{M}_S^S(f) \mathbf{B}. \end{aligned} \quad (9)$$

Beispiel: Sei s eine Spiegelung an der zu \mathbf{a} , $\|\mathbf{a}\| = 1$, orthogonalen Ebene. Dann ist

$$\begin{aligned} s(\mathbf{x}) &= \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{a})\mathbf{a}, \\ \mathbf{M}_S^S(s) &= \begin{pmatrix} 1 - 2a_1^2 & -2a_2a_1 & -2a_3a_1 \\ -2a_1a_2 & 1 - 2a_2^2 & -2a_3a_2 \\ -2a_1a_3 & -2a_2a_3 & 1 - 2a_3^2 \end{pmatrix}. \end{aligned}$$

Sei \mathbf{b} , $\|\mathbf{b}\| = 1$, $\mathbf{a} \cdot \mathbf{b} = 0$. Dann ist $\mathbf{B} = (\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b})$ eine Basis und es gilt:

$$s(\mathbf{a}) = -\mathbf{a}, \quad s(\mathbf{b}) = \mathbf{b}, \quad s(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times \mathbf{b},$$

denn $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$. Also gilt

$$\mathbf{M}_B^B(s) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

und $\mathbf{M}_B^B(s) = \mathbf{B}^{-1} \mathbf{S} \mathbf{B}$.

Definition 6.6.20 Zwei Matrizen $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n \times n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ gibt mit

$$\mathbf{C} = \mathbf{B}^{-1} \mathbf{A} \mathbf{B}.$$

• Durch geschickte Wahl einer Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ kann $\mathbf{M}_B^B(f)$ für eine gegebene lineare Abbildung f eine sehr einfache Form haben. Sei \mathbf{b}_1 derart, dass $f(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1$. Für die Basis \mathbf{B} deren erster Vektor \mathbf{b}_1 ist gilt:

$$\Rightarrow \mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & * & \cdots \\ 0 & \ddots & * \\ 0 & * & \ddots \end{pmatrix}.$$

Im Fall $f(\mathbf{b}_2) = \lambda_2 \mathbf{b}_2$ ist:

$$\mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \ddots \end{pmatrix}.$$

Wie kann man solche \mathbf{b}_i finden? Für welche Abbildung gibt es eine Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ so, dass

$$f(\mathbf{b}_i) = \lambda_i \mathbf{b}_i \quad i = 1, \dots, n.$$

Dies ist das sogenannte **Normalformen-** oder **Eigenwertproblem**. Es ist eng verbunden mit Nullstellen von Polynomen. Demzufolge lassen wir von nun an auch komplexe Skalare zu, d.h. \mathbb{R} wird durch \mathbb{C} ersetzt. Alle bisherigen Aussagen bleiben wahr.

Definition 6.6.21 Eine Zahl $\lambda \in \mathbb{C}$ heißt **Eigenwert** einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, wenn es wenigstens einen Spaltenvektor $\mathbf{b} \in \mathbb{C}^n, \mathbf{b} \neq \mathbf{0}$, gibt mit

$$\mathbf{A}\mathbf{b} = \lambda\mathbf{b}. \quad (10)$$

Jeder Vektor $\mathbf{b} \neq \mathbf{0}$, der (10) erfüllt, heißt **Eigenvektor** von \mathbf{A} zum Eigenwert λ .

Beispiel: Für die Spiegelung $s(\mathbf{x}) = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{a})\mathbf{a}, \|\mathbf{a}\| = 1$ gilt $s(\mathbf{a}) = -\mathbf{a}$ und also $\mathbf{M}_S^S(s)\mathbf{a} = -\mathbf{a}$, d.h. -1 ist Eigenwert und \mathbf{a} Eigenvektor von $\mathbf{M}_S^S(s)$.

Definition 6.6.22 Das charakteristische Polynom einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist definiert durch

$$\chi_{\mathbf{A}}(\lambda) := \det(\lambda\mathbf{E} - \mathbf{A}) \quad (11)$$

- Die Determinante in (11) ist explizit zu berechnen, wobei λ eine Variable ist. Man kann zeigen, dass gilt:

$$\det(\lambda\mathbf{E} - \mathbf{A}) = \lambda^n - \text{Sp}(\mathbf{A})\lambda^{n-1} + \dots + (-1)^n \det \mathbf{A},$$

wobei

$$\text{Sp}(\mathbf{A}) := \sum_{j=1}^n a_{jj}. \quad (12)$$

Satz 6.6.23 Es ist $\lambda \in \mathbb{C}$ genau dann Eigenwert von $\mathbf{A} \in \mathbb{C}^{n \times n}$, wenn λ eine Nullstelle des charakteristischen Polynoms $\chi_{\mathbf{A}}$ ist.

Definition 6.6.24 Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ eine Matrix und sei $\lambda \in \mathbb{C}$ ihr Eigenwert. Die **algebraische Vielfachheit** $k(\lambda)$ von λ ist die Ordnung von λ als Nullstelle des charakteristischen Polynoms. Der **Eigenraum** $E(\lambda)$ zum Eigenwert λ ist der Lösungsraum des homogenen Gleichungssystems $(\lambda\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, d.h.

$$E(\lambda) = \text{Ker}(\lambda\mathbf{E} - \mathbf{A}). \quad (13)$$

Die Dimension des Eigenraums heißt **geometrische Vielfachheit** $m(\lambda)$ des Eigenwerts λ .

Satz 6.6.25 Eigenvektoren $\mathbf{b}_1, \dots, \mathbf{b}_r \in \mathbb{C}^n$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$ der Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ sind linear unabhängig.

Satz 6.6.26 Besitzt eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ n linear unabhängige Eigenvektoren $\mathbf{b}_1, \dots, \mathbf{b}_r$ mit nicht notwendig verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$, dann gilt mit $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_r) \in \mathbb{C}^{n \times n}$

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}. \quad (14)$$

Satz 6.6.27 Eine Matrix besitzt genau dann n linear unabhängige Eigenvektoren, wenn die Summe der geometrischen Vielfachheiten der Eigenwerte gleich n ist, d.h.

$$\sum_{j=1}^k m(\lambda_j) = n. \quad (15)$$

Folgerung 6.6.28 Falls (15) gilt, ist \mathbf{A} ähnlich einer Diagonalmatrix, d.h. (14) gilt.

Folgerung 6.6.29 Falls für alle Eigenwerte von $\mathbf{A} \in \mathbb{C}^{n \times n}$ die geometrische und die algebraische Vielfachheit übereinstimmen, gibt es eine Basis \mathbf{B} des \mathbb{C}^n so, dass (14) gilt.

• Durch die Identifizierung von linearen Abbildungen und Matrizen erhalten wir folgenden Satz.

Satz 6.6.30 Sei $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ eine lineare Abbildung und sei $\mathbf{A} = \mathbf{M}_S^S(f) \in \mathbb{C}^{n \times n}$ ihre Abbildungsmatrix. Falls die geometrischen und die algebraischen Vielfachheiten aller Eigenwerte der Matrix \mathbf{A} übereinstimmen, gibt es eine Basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ des \mathbb{C}^n aus Eigenvektoren von \mathbf{A} so, dass gilt:

$$\mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

6.6.31 Berechnung von Eigenwerten und Eigenvektoren

a) Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$. Bestimme alle Nullstellen $\lambda_1, \dots, \lambda_r$ des charakteristischen Polynoms $\chi_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{E} - \mathbf{A})$, d.h.

$$\chi_{\mathbf{A}}(\lambda) = (\lambda_1 - \lambda)^{k_1} \cdots (\lambda_r - \lambda)^{k_r},$$

wobei λ_j die Eigenwerte von \mathbf{A} mit der algebraischen Vielfachheit $k(\lambda_j) = k_j$ sind.

b) Bestimme die Eigenräume $\mathbf{E}(\lambda_j) = \text{Ker}(\lambda_j \mathbf{E} - \mathbf{A})$, $j = 1, \dots, r$, und berechne für jeden eine Basis. Die Dimension von $\mathbf{E}(\lambda_j)$, $j = 1, \dots, r$ ist die geometrische Vielfachheit $m(\lambda_j)$ des Eigenwertes λ_j .

c) Überprüfe, ob für alle λ_j , $j = 1, \dots, r$ gilt $k(\lambda_j) = m(\lambda_j)$. Falls ja, so ist \mathbf{A} bezüglich der aus den Basen der Eigenräume $\mathbf{E}(\lambda_j)$ zusammengesetzten Basis ähnlich einer Diagonalmatrix.

• Falls $\mathbf{A} \in \mathbb{R}^{n \times n}$ und alle Eigenwerte reell sind, d.h. $\lambda_j \in \mathbb{R}$, dann gelten Satz 6.6.26, Folgerung 6.6.29 und Satz 6.6.30 analog mit Matrizen $\mathbf{B} \in \mathbb{R}^{n \times n}$ bzw. Basen $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ des \mathbb{R}^n . Diese können mit Hilfe des Verfahrens aus 6.6.31 berechnet werden.

• Es kann $m(\lambda_j) = k(\lambda_j)$.

Beispiele:

1) Sei $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, dann gilt:

$$\chi_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 1 \end{pmatrix} = (\lambda_1 + 1)(\lambda_1 - 3).$$

Also sind $\lambda_1 = -1$, und $\lambda_2 = 3$ die Eigenwerte von \mathbf{A} . Das System $(\lambda_1 \mathbf{E} - \mathbf{A}) \mathbf{v}_1 = \mathbf{0}$ hat die Lösung $\mathbf{v}_1 = (1, -1)^T$. Die Lösung von $(\lambda_2 \mathbf{E} - \mathbf{A}) \mathbf{v}_2 = \mathbf{0}$ ist $\mathbf{v}_2 = (1, -1)^T$. Also gilt: $\dim \mathbf{E}(\lambda_{ij}) = 1, j = 1, 2$, und \mathbf{A} ist ähnlich einer Diagonalmatrix, d.h. mit $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ gilt:

$$\mathbf{B}^{-1} \mathbf{A} \mathbf{B} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}.$$

2) Für $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ gilt $\chi_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{E} - \mathbf{A}) = (\lambda - 3)^2$. Also ist der Eigenwert von \mathbf{A} $\lambda_1 = \lambda_2 = 3$. Da $\lambda_1 \mathbf{E} - \mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ gilt $k(\lambda_1) = 2$ und $\mathbf{E}(\lambda_1) = \mathbb{R}^2$, d.h. $k(\lambda_1) = m(\lambda_1)$. Mit $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gilt

$$\mathbf{B}^{-1} \mathbf{A} \mathbf{B} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

3) Für $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ haben wir

$$\det(\lambda \mathbf{E} - \mathbf{A}) = \det \begin{pmatrix} \lambda - 2 & 1 \\ -1 & \lambda - 4 \end{pmatrix} = (\lambda - 3)^2.$$

Also ist der Eigenwert von \mathbf{A} $\lambda_1 = \lambda_2 = 3$ und $k(\lambda_1) = 2$. Der Eigenraum $\mathbf{E}(\lambda_1)$ hat die Dimension 1, d.h. $m(\lambda_1) = 1$ und $\mathbf{v}_1 = (1, -1)^T$ ist eine Basis von $\mathbf{E}(\lambda_1)$, denn

$$\lambda_1 \mathbf{E} - \mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Also ist $k(\lambda_1) \neq m(\lambda_1)$ und \mathbf{A} ist nicht ähnlich einer Diagonalmatrix.

Aber man kann \mathbf{A} mit folgendem Verfahren in eine einfachere Form bringen.

Sei \mathbf{v}_2 die Lösung von $(\lambda \mathbf{E} - \mathbf{A}) \mathbf{v}_2 = \mathbf{v}_1$, d.h. $\mathbf{v}_2 = (1, 0)^T$. Setze man nun $\mathbf{B} = (\mathbf{v}_2, \mathbf{v}_1)$, d.h. $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ dann gilt:

$$\mathbf{B}^{-1} \mathbf{A} \mathbf{B} = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}.$$

Man sagt, dass $\mathbf{B}^{-1}\mathbf{A}\mathbf{B}$ ein 3-er Block der Länge zwei ist. Allgemein ist ein λ -Block der Länge k gegeben durch

$$\begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \ddots & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

4) Für $\mathbf{A} = \begin{pmatrix} 1 & 5 \\ 1 & -1 \end{pmatrix}$ haben wir

$$\chi_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} \lambda - 1 & 5 \\ -1 & \lambda + 1 \end{pmatrix} = \lambda^2 + 4 = (\lambda - 2i)(\lambda + 2i).$$

Also sind $\lambda_1 = 2i$ und $\lambda_2 = -2i$ die Eigenwerte von \mathbf{A} . Eine Basis von $\mathbf{E}(\lambda_1)$ berechnet sich durch

$$\begin{pmatrix} 2i - 1 & 5 \\ -1 & 2i + 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2i - 1 \\ 2i - 1 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2i - 1 \\ 0 & 0 \end{pmatrix},$$

d.h. $\mathbf{v}_2 = (2i + 1, 1)^T$. Analog für $\mathbf{E}(\lambda_2)$:

$$\begin{pmatrix} -2i - 1 & 5 \\ -1 & -2i + 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2i - 1 \\ -2i - 1 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2i - 1 \\ 0 & 0 \end{pmatrix},$$

d.h. $\mathbf{v}_2 = (-2i + 1, 1)^T$. Somit gilt mit $\mathbf{B} = \begin{pmatrix} 2i + 1 & -2i + 1 \\ 1 & 1 \end{pmatrix}$

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{B} = \begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} \in \mathbb{C}^{2 \times 2}.$$

Satz 6.6.32 (Jordansche Normalform) Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ und seien alle Eigenwerte $\lambda_1, \dots, \lambda_r$ von \mathbf{A} reell. Dann ist \mathbf{A} ähnlich einer diagonalen Blockmatrix bestehend aus λ_j -Blöcken.

- Zu einem Eigenwert λ können verschiedene λ -Blöcke auftreten.

$$\left(\begin{array}{ccc|ccc} \lambda & 0 & 0 & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 1 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & \lambda \end{array} \right) \begin{array}{l} \text{Länge 3} \\ \text{Länge 2} \\ \text{Länge 1} \end{array}$$

- Bis auf die Reihenfolge ist die diagonale Blockmatrix eindeutig bestimmt.

6.7 Symmetrische Matrizen

Für eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ gilt unter geeigneten Voraussetzungen

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2.$$

Im nächsten Kapitel betrachten wir Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ für die analog gilt,

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \sum_{i=1}^n D_i f(\mathbf{x}_0)(x - x_0)_i + \frac{1}{2} \sum_{i,j=1}^n D_{ij}^2 f(\mathbf{x}_0)(x - x_0)_i(x - x_0)_j.$$

Wir setzen $f(\mathbf{x}_0) =: \alpha_0$, $\mathbf{a} := (D_1 f(\mathbf{x}_0), \dots, D_n f(\mathbf{x}_0))^T$ und $\mathbf{A} := (D_{ij}^2 f(\mathbf{x}_0))_{i,j=1,\dots,n}$ und erhalten

$$f(\mathbf{x}) \approx \alpha_0 + \mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0).$$

Aber es gilt $x_i x_j = x_j x_i$ und demzufolge kann man \mathbf{A} durch die Matrix mit den Komponenten $\frac{1}{2}(a_{ij} + a_{ji})$ ersetzen, d.h. durch ihren symmetrischen Anteil.

Definition 6.7.1 Eine Funktion $q : \mathbb{R}^n \rightarrow \mathbb{R}$ heisst **quadratische Form** falls

$$q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

mit einer symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, d.h. $\mathbf{A} = \mathbf{A}^T$. Ist \mathbf{A} eine Diagonalmatrix, so heisst q rein quadratisch.

6.7.2 Basiswechsel Sei $q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ eine Darstellung der quadratischen Form bezüglich der Standardbasis \mathbf{S} . Sei \mathbf{B} eine andere Basis des \mathbb{R}^n und $\mathbf{M}_S^B(id)$ die Basiswechselmatrix für den Übergang von \mathbf{B} nach \mathbf{S} . Nach (3) gilt

$$k_S(\mathbf{x}) = \mathbf{M}_S^B(id) k_B(\mathbf{x}) = \mathbf{B} k_B(\mathbf{x})$$

und somit ist

$$\begin{aligned} q(\mathbf{x}) &= k_S(\mathbf{x})^T \mathbf{A} k_S(\mathbf{x}) \\ &= (\mathbf{B} k_B(\mathbf{x}))^T \mathbf{A} \mathbf{B} k_B(\mathbf{x}) \\ &= k_B(\mathbf{x})^T \mathbf{B}^T \mathbf{A} \mathbf{B} k_B(\mathbf{x}) \end{aligned}$$

eine Darstellung von $q(\mathbf{x})$ bezüglich der Basis \mathbf{B} .

Definition 6.7.3 Als **Hauptachsensystem** der quadratischen Form $q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ (bzw. der symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$) bezeichnet man eine Orthonormalbasis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ des \mathbb{R}^n , wenn $q(\mathbf{x})$ bezüglich der Basis \mathbf{B} rein quadratisch ist.

- Nach 6.7.2 ist \mathbf{B} ein Hauptachsensystem einer symmetrischen Matrix \mathbf{A} genau dann, wenn $\mathbf{B}^T \mathbf{A} \mathbf{B}$ eine Diagonalmatrix ist. Also suchen wir eine invertierbare Matrix \mathbf{B} mit der $\mathbf{B}^T \mathbf{A} \mathbf{B}$ eine Diagonalmatrix ist. Dies ist nicht mit dem Normalformenproblem zu verwechseln, bei dem man für eine nicht notwendigerweise symmetrische Matrix \mathbf{A} eine invertierbare Matrix \mathbf{B} sucht, so dass $\mathbf{B}^{-1} \mathbf{A} \mathbf{B}$ eine Diagonalmatrix ist.

Satz 6.7.4 Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann gilt:

- i) Alle Eigenwerte von \mathbf{A} sind reell.
- ii) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.
- iii) Algebraische und geometrische Vielfachheit jedes Eigenwertes sind gleich.

Satz 6.7.5 (Hauptachsentransformation) Zu jeder quadratischen Form $q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ bzw. jeder symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, gibt es wenigstens ein Hauptachsensystem. Es wird wie folgt berechnet:

Zu jedem der verschiedenen Eigenwerte $\lambda_i, 1 \leq i \leq r$, von \mathbf{A} bestimmt man eine orthogonale Basis des Eigenraumes $\mathbf{E}(\lambda_i), 1 \leq i \leq r$. Diese Teilbasen $(\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_{k_i}^{(i)})$ zusammengesetzt ergeben das Hauptachsensystem

$$\mathbf{B} = \left(\mathbf{b}_1^{(1)}, \dots, \mathbf{b}_{k_1}^{(1)}, \dots, \mathbf{b}_1^{(r)}, \dots, \mathbf{b}_{k_r}^{(r)} \right)$$

für das

$$\mathbf{B}^T \mathbf{A} \mathbf{B} = \text{Diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{k_1\text{-mal}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{k_r\text{-mal}})$$

gilt. Demzufolge gilt

$$q(\mathbf{x}) = k_{\mathbf{B}}(\mathbf{x})^T \mathbf{B}^T \mathbf{A} \mathbf{B} k_{\mathbf{B}}(\mathbf{x}) = \lambda_1 (k_{\mathbf{B}}(\mathbf{x}))_1^2 + \dots + \lambda_r (k_{\mathbf{B}}(\mathbf{x}))_n^2 .$$

Definition 6.7.6 Eine quadratische Form $q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, bzw. die zugehörige symmetrische Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, heißt **positiv definit (negativ definit)**, wenn aus $\mathbf{x} \neq \mathbf{0}$ stets $q(\mathbf{x}) > 0$ ($q(\mathbf{x}) < 0$) folgt. Die quadratische Form heißt **indefinit**, wenn sie sowohl positive als auch negative Werte annimmt. Sie heißt **positiv (negativ) semidefinit** wenn stets $q(\mathbf{x}) \geq 0$ ($q(\mathbf{x}) \leq 0$) gilt.

Satz 6.7.7

- a) $\mathbf{D} = \text{Diag}(\alpha_1, \dots, \alpha_n)$ ist genau dann positiv definit, wenn alle α_i positiv sind.
- b) $\mathbf{A} = \mathbf{A}^T$ ist genau dann positiv definit, wenn $\mathbf{W}^T \mathbf{A} \mathbf{W}$ für irgendeine invertierbare Matrix $\mathbf{W} \in \mathbb{R}^{n \times n}$ positiv definit ist.
- c) $\mathbf{A} = \mathbf{A}^T$ ist genau dann positiv definit, wenn sämtliche Eigenwerte von \mathbf{A} positiv definit sind.
- Der Satz gilt analog für negativ definite symmetrische Matrizen, d.h. $\mathbf{A} = \mathbf{A}^T$ ist negativ definit genau dann wenn alle Eigenwerte negativ sind. $\mathbf{D} = \text{Diag}(\alpha_1, \dots, \alpha_n)$ ist genau dann negativ definit, wenn alle $\alpha_i, i = 1, \dots, n$ negativ sind.

6.7.8 Positivitätskriterium Es sei $\mathbf{A} = \mathbf{A}^T$. Es gilt $\mathbf{e}_i^T \mathbf{A} \mathbf{e}_i = a_{ii}$. Damit haben wir eine notwendige Bedingung:

$$\mathbf{A} = \mathbf{A}^T \text{ positiv definit} \Rightarrow a_{ii} > 0, \quad i = 1, \dots, n. \quad (1)$$

Sei $n = 2$ und $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, mit $a, d > 0$. Mit Zeilen- und Spaltenumformungen erhalten wir

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} \rightarrow \begin{pmatrix} a & b \\ 0 & d - \frac{b^2}{a} \end{pmatrix} \rightarrow \begin{pmatrix} a & 0 \\ 0 & d - \frac{b^2}{a} \end{pmatrix}$$

Also gilt:

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ positiv definit} \Leftrightarrow a > 0, \det \mathbf{A} = ad - b^2 > 0 \quad (2)$$

Satz 6.7.9 $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{n \times n}$, $\mathbf{A} = (a_{ij})$ ist genau dann positiv definit, wenn die Determinanten der **Hauptuntermatrizen** \mathbf{H}_i positiv sind:

$$\mathbf{H}_1 = a_{11} \quad \mathbf{H}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \dots, \mathbf{H}_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{1k} & \dots & a_{kk} \end{pmatrix}, \mathbf{H}_n = \mathbf{A}.$$

Satz 6.7.10 $\mathbf{A} = \mathbf{A}^T$ ist genau dann negativ definit, wenn für die Determinanten der Hauptuntermatrizen gilt:

$$(-1)^k \det \mathbf{H}_k > 0.$$

Kapitel 7

Differentiation von Funktionen mehrerer Variablen

In diesem Kapitel betrachten wir Funktionen der Form

$$\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m : \mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$$

wobei $\mathbf{x} = (x_1, \dots, x_n)^T$, $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T$. \mathbb{R}^n heißt Urbildraum, D **Definitionsbereich**, \mathbb{R}^m Bild- oder Zielraum und das **Bild von \mathbf{f}** ist

$$\mathbf{f}(D) := \{\mathbf{f}(\mathbf{x}); \mathbf{x} \in D\}.$$

7.1 Kurven im \mathbb{R}^n

7.1.1 (Parameterdarstellung) *Kurven stellt man sich am besten als die Bahn eines bewegten Punktes vor, wobei man jedem Zeitpunkt t einen Ortsvektor $\mathbf{f}(t)$ zuordnet. Sei also $I \subseteq \mathbb{R}$ ein Intervall und $\mathbf{f} : I \rightarrow \mathbb{R}^n$ eine gegebene vektorwertige Funktion. Die Funktion $f_i : I \rightarrow \mathbb{R}$, $1 \leq i \leq n$ heißt **i -te Komponentenfunktion** von \mathbf{f} .*

Definition 7.1.2 *Sei $\mathbf{f} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ eine Funktion. Die Funktion \mathbf{f} hat im Punkt $t_0 \in I$ den **Grenzwert \mathbf{c}** , in Zeichen $\lim_{t \rightarrow t_0} \mathbf{f}(t) = \mathbf{c}$, genau dann, wenn*

$$\lim_{t \rightarrow t_0} f_i(t) = c_i, \quad 1 \leq i \leq n.$$

- Der Grenzwert ist komponentenweise definiert und wird somit zurückgeführt auf Grenzwerte von Funktionen in einer Dimension.
- Analog heißt $\mathbf{f} : I \rightarrow \mathbb{R}^n$ **stetig** bzw. **differenzierbar**, wenn alle Komponentenfunktionen stetig bzw. differenzierbar sind. Die Ableitung wird komponentenweise berechnet, d.h.

$$\dot{\mathbf{f}}(t) = \frac{d}{dt} \mathbf{f}(t) = \lim_{h \rightarrow 0} \frac{1}{h} (\mathbf{f}(t+h) - \mathbf{f}(t)) = \begin{pmatrix} \dot{f}_1(t) \\ \vdots \\ \dot{f}_n(t) \end{pmatrix}. \quad (1)$$

Man nennt $\dot{\mathbf{f}}(t)$ den **Tangentialvektor** von \mathbf{f} an der Stelle t . (Die Schreibweise $\dot{\mathbf{f}}$ geht auf Newton zurück und wird vor allem in der Physik zur Beschreibung der Ableitung nach dem Zeit-Parameter t benutzt.)

Lemma 7.1.3 Für $\mathbf{f}, \mathbf{g} : I \rightarrow \mathbb{R}^n$ und $\alpha : I \rightarrow \mathbb{R}$ gelten folgende Rechenregeln:

- $\frac{d}{dt}(a\mathbf{f}(t) + b\mathbf{g}(t)) = a\dot{\mathbf{f}}(t) + b\dot{\mathbf{g}}(t)$, $a, b \in \mathbb{R}$,
- $\frac{d}{dt}(\mathbf{f}(t) \cdot \mathbf{g}(t)) = \dot{\mathbf{f}}(t) \cdot \mathbf{g}(t) + \mathbf{f}(t) \cdot \dot{\mathbf{g}}(t)$,
- $\frac{d}{dt}(\mathbf{f}(t) \times \mathbf{g}(t)) = \dot{\mathbf{f}}(t) \times \mathbf{g}(t) + \mathbf{f}(t) \times \dot{\mathbf{g}}(t)$, $n = 3$,
- $\frac{d}{dt}(\alpha(t)\mathbf{f}(t)) = \dot{\alpha}(t)\mathbf{f}(t) + \alpha(t)\dot{\mathbf{f}}(t)$.

- Aus

$$\|\mathbf{f}(t)\| = \text{const} \quad \text{folgt} \quad \mathbf{f}(t) \cdot \dot{\mathbf{f}}(t) = 0. \quad (2)$$

Definition 7.1.4 Sei $G \subseteq \mathbb{R}^n$ und $[a, b] \subseteq \mathbb{R}$. Wir nennen jede stetig differenzierbare Funktion $\mathbf{f} : [a, b] \rightarrow G$ eine **Parameterdarstellung** des Kurvenstücks $\{\mathbf{f}(t); t \in [a, b]\}$ mit Anfangspunkt $\mathbf{f}(a)$ und Endpunkt $\mathbf{f}(b)$. Ein Kurvenstück heißt **regulär**, wenn $\dot{\mathbf{f}}(t) \neq \mathbf{0}$ für alle $t \in (a, b)$ gilt.

- Mehrere Kurvenstücke

$$\begin{aligned} K_1 \quad \mathbf{f}_1 : I_1 = [a_1, b_1] &\rightarrow G \\ K_2 \quad \mathbf{f}_2 : I_2 = [a_2, b_2] &\rightarrow G, \quad a_2 = b_1 \\ &\vdots \\ K_r \quad \mathbf{f}_r : I_r = [a_r, b_r] &\rightarrow G, \quad a_r = b_{r-1} \end{aligned}$$

können zusammengesetzt werden zu einer **Kurve**

$$K := K_1 \cup \dots \cup K_r \quad \mathbf{f} : I := [a_1, b_r] \rightarrow G, \quad \mathbf{f} = \mathbf{f}_j \text{ auf } I_j.$$

Die Funktion $\mathbf{f} : I \rightarrow G$ muss in den Punkten a_2, \dots, a_r nicht differenzierbar sein.

Beispiel: Polygonzug = endlich viele aneinanderghängte Strecken.

- Eine Kurve K wird beschrieben durch das Parameter-Intervall I und die Funktion \mathbf{f} . Kurz: $K = (I, \mathbf{f})$.

Eine Kurve kann i.a. auf viele Arten parametrisiert werden. Z.B. der Halbkreis-Bogen in \mathbb{R}^2 mit dem Radius 1, von $(1, 0)^T$ nach $(-1, 0)^T$ durchlaufen.

- $I = [-1, 1], \quad \mathbf{f}(t) = (-t, (1 - t^2)^{1/2})^T,$
- $I = [0, \pi], \quad \mathbf{f}(t) = (\cos t, \sin t)^T.$

Auf die Frage wie man erkennt, ob zwei Parameterdarstellungen die gleiche Kurve beschreiben, soll hier nicht eingegangen werden.

- Eine Kurve ist i.a. nicht identisch mit der Menge ihrer Punkte.
Beispiel: Die Kurven K_1 : Einheitskreis, einmal im Uhrzeigersinn von $(1, 0)^T$ nach $(1, 0)^T$ durchlaufen und K_2 : Wie K_1 , aber zweimal durchlaufen, haben zwar gleiche Punktmengen, sind aber als verschieden anzusehen.
- Durch eine Parametrisierung ist auch angegeben, wie die Kurvenpunkte durchlaufen werden. Die Darstellung $([0, 2\pi], (\cos t, \sin t)^T)$ beschreibt den Einheitskreis, im positiven Sinn durchlaufen, während die Darstellung $([0, 2\pi], (\cos(2\pi - t), \sin(2\pi - t))^T)$ Durchlaufen im umgekehrten Sinn bedeutet. Auch hier wird man von verschiedenen Kurven sprechen, obwohl die Menge der Kurvenpunkte dieselbe ist.

Definition 7.1.5 Sei $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^n$ ein reguläres Kurvenstück. Dann heißt

$$s(t) := \int_a^t \|\dot{\mathbf{f}}(\tau)\| d\tau$$

die **Bogenlänge** des Kurvenstücks über $[a, t]$.

- Nach dem Hauptsatz der Differential- und Integralrechnung (Kapitel 4, Satz 4.1.7) gilt

$$\dot{s}(t) = \|\dot{\mathbf{f}}(t)\|. \quad (3)$$

Definition 7.1.6 Sei $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^3$ zweimal stetig differenzierbar mit $\dot{\mathbf{f}}(t) \neq \mathbf{0}$ und $\dot{\mathbf{f}}(t) \times \ddot{\mathbf{f}}(t) \neq \mathbf{0}$. Dann sind

$$\begin{aligned} \mathbf{T}(t) &:= \frac{\dot{\mathbf{f}}(t)}{\|\dot{\mathbf{f}}(t)\|} && \text{der \textbf{Tangenten(einheits)vektor},} \\ \mathbf{N}(t) &:= \frac{\dot{\mathbf{T}}(t)}{\|\dot{\mathbf{T}}(t)\|} && \text{der \textbf{Normalen(einheits)vektor},} \\ \mathbf{B}(t) &:= \mathbf{T}(t) \times \mathbf{N}(t) && \text{der \textbf{Binormalenvektor}.} \end{aligned} \quad (4)$$

Das Rechtssystem $(\mathbf{T}(t), \mathbf{N}(t), \mathbf{B}(t))$ wird begleitendes **Dreibein** genannt.

- Der Grenzwert der Änderungsrate $\frac{\Delta \mathbf{T}}{\Delta s} = \frac{\mathbf{T}(t_1) - \mathbf{T}(t_2)}{s(t_1) - s(t_2)}$ des Tangentialvektors heißt **Krümmungsvektor** und ist durch $\frac{\dot{\mathbf{T}}(t)}{\dot{s}(t)}$ gegeben. Seine Länge heißt Krümmung

$$\kappa(t) := \frac{\|\dot{\mathbf{T}}(t)\|}{\dot{s}(t)} = \frac{\|\dot{\mathbf{T}}(t)\|}{\|\dot{\mathbf{f}}(t)\|}. \quad (5)$$

Lemma 7.1.7 Wir haben die Darstellungen

$$\begin{aligned} \dot{\mathbf{f}}(t) &= \dot{s}(t) \mathbf{T}(t), \\ \ddot{\mathbf{f}}(t) &= \ddot{s}(t) \mathbf{T}(t) + \dot{s}(t)^2 \kappa(t) \mathbf{N}(t). \end{aligned}$$

Lemma 7.1.8 *Wir haben folgende Darstellungen:*

$$a) \kappa(t) = \frac{\|\dot{\mathbf{f}}(t) \times \ddot{\mathbf{f}}(t)\|}{\|\dot{\mathbf{f}}(t)\|^3},$$

$$b) \mathbf{B}(t) = \frac{\dot{\mathbf{f}}(t) \times \ddot{\mathbf{f}}(t)}{\|\dot{\mathbf{f}}(t) \times \ddot{\mathbf{f}}(t)\|}.$$

Beispiel: Eine Schraubenkurve hat die Darstellung

$$\mathbf{f}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ ht \end{pmatrix}.$$

Somit erhalten wir:

$$\dot{\mathbf{f}}(t) = \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, \quad \ddot{\mathbf{f}}(t) = \begin{pmatrix} -r \cos t \\ -r \sin t \\ 0 \end{pmatrix},$$

$$\|\dot{\mathbf{f}}(t)\| = \sqrt{r^2 + h^2} = \dot{s}(t) =: R,$$

$$\mathbf{T}(t) = \frac{1}{R} \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix},$$

$$\mathbf{N}(t) = \frac{1}{R} \begin{pmatrix} -r \cos t \\ -r \sin t \\ 0 \end{pmatrix} \cdot \frac{R}{r} = \begin{pmatrix} -\cos t \\ -\sin t \\ 0 \end{pmatrix},$$

$$\mathbf{B}(t) = \frac{1}{R} \begin{pmatrix} h \sin t \\ -h \cos t \\ r \end{pmatrix},$$

$$\kappa(t) = \frac{\|\dot{\mathbf{T}}\|}{\|\dot{\mathbf{f}}\|} = \frac{r}{R \cdot R} = \frac{r}{r^2 + h^2}.$$

7.2 Reellwertige Funktionen mehrerer Veränderlicher

7.2.1 (Grundlagen) Sei $D \subseteq \mathbb{R}^n$ und $f : D \rightarrow \mathbb{R}$ eine reellwertige Funktion. Dieser Funktion ist ein **Graph**

$$\Gamma_f := \left\{ \begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix}; \mathbf{x} \in D \right\} \subseteq \mathbb{R}^{n+1}$$

zugeordnet. Zur Veranschaulichung von f betrachtet man **Niveaumengen**, d.h. für $c \in \mathbb{R}$

$$N_c := \{\mathbf{x} \in D; f(\mathbf{x}) = c\} \subseteq \mathbb{R}^n,$$

oder auch Schnitte mit zu Koordinatenachsen parallelen Geraden, d.h. man betrachtet die „partiellen“ Funktionen von f , die gegeben sind durch

$$x_i \rightarrow f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n),$$

mit $\mathbf{a} = (a_1, \dots, a_n)^T \in D$.

Definition 7.2.2 Für zwei Punkte $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ wird ihr **Abstand** definiert durch

$$\|\mathbf{x} - \mathbf{y}\| := \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{\frac{1}{2}}.$$

Sei $\mathbf{a} \in \mathbb{R}^n$ ein beliebiger, aber fester Punkt und $r > 0$, dann heißt die Menge

$$B_r(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x} - \mathbf{a}\| < r\}$$

r -Umgebung von \mathbf{a} .

Definition 7.2.3 Sei D eine Teilmenge des \mathbb{R}^n .

- Ein Punkt $\mathbf{a} \in D$ heißt **innerer Punkt** von D , wenn es eine r -Umgebung $B_r(\mathbf{a})$ von \mathbf{a} gibt, die ganz in D liegt, d.h. $B_r(\mathbf{a}) \subseteq D$.
- D heißt **offen**, wenn jeder Punkt von D ein innerer Punkt ist.
- Ein Punkt $\mathbf{b} \in \mathbb{R}^n$ heißt **Randpunkt** von D , wenn jede r -Umgebung von \mathbf{b} sowohl mindestens einen Punkt aus D als auch einen nicht zu D gehörenden Punkt enthält. Die Menge aller Randpunkte heißt **Rand** von D und wird mit ∂D bezeichnet.
- Eine Menge ist abgeschlossen, wenn sie alle ihre Randpunkte enthält.

Beispiele:

- Die Kreisscheibe $B_r(\mathbf{a}) = \{(x, y)^T \in \mathbb{R}^2, (x - a_1)^2 + (y - a_2)^2 < r^2\}$ ist offen. Der Rand ist $\{(x, y)^T \in \mathbb{R}^2; (x - a_1)^2 + (y - a_2)^2 = r^2\}$.
- $C_r(\mathbf{a}) = \{(x, y)^T \in \mathbb{R}^2; (x - a_1)^2 + (y - a_2)^2 \leq r^2\}$ ist abgeschlossen.
- $B_r(\mathbf{a}) \cap \{(x, y)^T \in \mathbb{R}^2; x \geq 0\}$ ist weder offen noch abgeschlossen.

Definition 7.2.4 Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ und $\mathbf{a} \in D \cup \partial D$.

- f hat in \mathbf{a} den Grenzwert $c \in \mathbb{R}$, in Zeichen

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = c$$

wenn es zu jedem $\varepsilon > 0$ eine r -Umgebung $B_r(\mathbf{a})$ gibt, so daß

$$|f(\mathbf{x}) - c| \leq \varepsilon$$

für alle $\mathbf{x} \in D \cap B_r(\mathbf{a})$ gilt.

b) f heißt in $\mathbf{a} \in D$ **stetig**, wenn $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ gilt.

c) f heißt auf D **stetig**, wenn f in allen $\mathbf{a} \in D$ stetig ist.

- Analog zum Fall $n = 1$ gelten für Grenzwerte und stetige Funktionen die üblichen Rechenregeln, d.h. Summe, Produkt und Quotient stetiger Funktionen sind stetig.
- Die Projektion $p_i(\mathbf{x}) = x_i$ ist stetig und somit sind alle Polynome

$$p(\mathbf{x}) = \sum_{1 \leq k_i \leq m} a_{k_1, \dots, k_m} x_1^{k_1} \cdots x_n^{k_n}$$

stetig. Insbesondere sind lineare Abbildungen

$$\ell(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = \sum_{i=1}^n a_i x_i$$

stetig. Die rationale Funktion $r(\mathbf{x}) = \frac{p(\mathbf{x})}{q(\mathbf{x})}$ ist stetig, falls $q(\mathbf{x}) \neq 0$ für alle $\mathbf{x} \in D$.

- Achtung! Aus der Stetigkeit der partiellen Funktionen folgt nicht die Stetigkeit von f .

Definition 7.2.5

a) Eine Menge $D \subseteq \mathbb{R}^n$ heißt **beschränkt**, wenn es eine Konstante $K > 0$ mit

$$\|\mathbf{x}\| < K \quad \text{für alle } \mathbf{x} \in D \text{ gibt.}$$

b) Die abgeschlossenen und beschränkten Mengen des \mathbb{R}^n nennt man **kompakt**.

Satz 7.2.6 Jede auf einer kompakten Menge $D \subseteq \mathbb{R}^n$ stetige Funktion nimmt auf D ein Minimum und ein Maximum an, d.h. es gibt $\mathbf{a}, \mathbf{b} \in D$ mit

$$f(\mathbf{a}) \leq f(\mathbf{x}) \leq f(\mathbf{b}) \quad \text{für alle } \mathbf{x} \in D.$$

Satz 7.2.7 Jede auf einer kompakten Menge $D \subseteq \mathbb{R}^n$ stetige Funktion ist **gleichmäßig stetig**, d.h. zu $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $\mathbf{x}, \mathbf{y} \in D$ gilt:

$$\|\mathbf{x} - \mathbf{y}\| < \delta \quad \Rightarrow \quad |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon.$$

- Die Sätze 7.2.6 und 7.2.7 sind im Allgemeinen falsch, wenn D nicht kompakt ist.
- Sei $f : D \rightarrow \mathbb{R}$ stetig und D beschränkt, aber nicht kompakt. Falls für alle $\mathbf{x} \in \partial D$ die Funktion f in \mathbf{x} einen Grenzwert hat, kann man f stetig auf $D \cup \partial D$ fortsetzen durch

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \mathbf{x} \in D, \\ \lim_{\substack{\mathbf{y} \rightarrow \mathbf{x} \\ \mathbf{y} \in D}} f(\mathbf{y}) & \mathbf{x} \in \partial D. \end{cases}$$

Dann ist \tilde{f} stetig auf der kompakten Menge $D \cup \partial D$ und Satz 7.2.6 und Satz 7.2.7 gelten (für \tilde{f}).

- Sei $f : D \rightarrow \mathbb{R}$ stetig und sei $f(\mathbf{a}) > 0$ für $\mathbf{a} \in D$. Dann gibt es ein $r > 0$ so dass

$$f(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in D \cap B_r(\mathbf{a}).$$

Definition 7.2.8 Sei $D \subseteq \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ eine Funktion und $\mathbf{x} \in D$. Existiert der Grenzwert

$$\lim_{t \rightarrow 0} \frac{1}{t} (f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n))$$

so wird er als die **partielle Ableitung von f nach x_i an der Stelle \mathbf{x}** genannt und als $\frac{\partial f}{\partial x_i}(\mathbf{x})$ bezeichnet.

- Die partielle Ableitung ist die Ableitung der „partiellen Funktionen“

$$x_i \rightarrow f(x_1, \dots, x_i, \dots, x_n),$$

wobei alle Werte $x_j, j \neq i$ konstant gehalten werden.

- Eine andere Bezeichnung ist f_{x_i} .
- **Höhere partielle Ableitungen** werden analog definiert und z.B. mit $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}$ bezeichnet.
- f heißt **partiell differenzierbar (in \mathbf{x})** bzw. **stetig partiell differenzierbar (in \mathbf{x})** wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ an der Stelle \mathbf{x} existieren bzw. stetig sind.
- Analoge Definition für höhere Ableitungen.

Hinweis: Als Definitionsbereiche $D \subseteq \mathbb{R}^n$ für differenzierbare Funktionen (partielle Ableitung, totale Ableitung 7.2.12, Richtungsableitungen 7.2.15) sollen stets offene Mengen genommen werden.

Definition 7.2.9 Der **Gradient** einer Funktion $f : D \rightarrow \mathbb{R}^n$ ist der Vektor, der aus den partiellen Ableitungen von f besteht, d.h. für $\mathbf{x} \in D \setminus \partial D$ ist

$$\text{grad } f(\mathbf{x}) := \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)^T \in \mathbb{R}^n.$$

- Eine andere Bezeichnung ist $\nabla f(\mathbf{x})$.
- Sei $D \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N}_0$. Dann bezeichnen wir

$$C^k(D, \mathbb{R}) := \{f : D \rightarrow \mathbb{R}; f \text{ } k\text{-mal stetig partiell differenzierbar}\}$$

Satz 7.2.10 Für jede C^2 -Funktion $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ offen gilt:

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right), \quad 1 \leq i, j \leq n.$$

(Satz von Hermann Amandus Schwarz, 1843–1921).

- Der Satz gilt analog für k -te gemischte partielle Ableitungen, falls $f \in C^k(D, \mathbb{R})$.

Definition 7.2.11 Für zwei Funktionen $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, $k \in \mathbb{N}_0$ und $\mathbf{x}_0 \in D$ schreibt man

$$f(\mathbf{x}) = g(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{x}_0\|^k) \quad \text{für } \mathbf{x} \rightarrow \mathbf{x}_0$$

falls

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^k} = 0.$$

- Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar an der Stelle x_0 . Nach Kapitel 3 3.1.3 ist die beste lineare Approximation von f an der Stelle x_0 gegeben durch $g(x) = f(x_0) + f'(x_0)(x - x_0)$, d.h.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|).$$

Für $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ reicht partielle Differenzierbarkeit für das Analogon obiger Formel nicht.

Definition 7.2.12 Sei $D \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt in $\mathbf{x}_0 \in D$ **total differenzierbar**, wenn es einen Vektor $\mathbf{a} \in \mathbb{R}^n$ gibt, so dass für alle $\mathbf{x} \in D$ gilt

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|), \quad (1)$$

Satz 7.2.13 Ist f in $\mathbf{x}_0 \in D$ total differenzierbar mit $f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$, dann gilt:

a) f ist in \mathbf{x}_0 stetig,

$$b) \lim_{t \rightarrow 0} \frac{1}{t} (f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)) = \mathbf{a} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq 0,$$

c) f ist partiell differenzierbar in \mathbf{x}_0 und \mathbf{a} ist eindeutig bestimmt als $\mathbf{a} = \text{grad } f(\mathbf{x}_0)$.

- Wenn f total differenzierbar in \mathbf{x}_0 ist, so ist

$$f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \quad (2)$$

die beste lineare Approximation von f nahe \mathbf{x}_0 .

Satz 7.2.14 Jede C^1 -Funktion $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, mit D offen, ist auf D total differenzierbar, d.h. jede stetig partiell differenzierbare Funktion ist total differenzierbar.

Definition 7.2.15 Seien $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{x}_0 \in D$ und $\mathbf{v} \in \mathbb{R}^n$ ein Einheitsvektor, d.h. $\|\mathbf{v}\| = 1$. Falls der Grenzwert

$$\lim_{t \rightarrow 0} \frac{1}{t} (f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0))$$

existiert, wird er mit $\partial_{\mathbf{v}} f(\mathbf{x}_0)$ oder $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ bezeichnet und heißt **Richtungsableitung** von f an der Stelle \mathbf{x}_0 in Richtung \mathbf{v} .

- Dies ist eine Verallgemeinerung der partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ (dort: $\mathbf{v} = \mathbf{e}_i$).
- $\partial_{\mathbf{v}} f(\mathbf{x}_0)$ beschreibt das Verhalten von f längs der Geraden $\mathbf{x}_0 + t\mathbf{v}$.

Satz 7.2.16 Für jede auf der offenen Menge $D \subseteq \mathbb{R}^n$ total differenzierbare Funktion f und alle $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$ gilt:

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{i=1}^n f_{x_i}(\mathbf{x}_0) v_i$$

Satz 7.2.17 (Kettenregel) Für jede C^1 -Funktion $f : D \rightarrow \mathbb{R}$ ($D \subseteq \mathbb{R}^n$ offen) und für jedes Kurvenstück $\mathbf{x} : [a, b] \rightarrow D$ gilt:

$$\frac{d}{dt} f(\mathbf{x}(t)) = \text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t). \quad (3)$$

7.2.18 (Polarkoordinaten) Im \mathbb{R}^2 besteht folgende Beziehung zwischen den Polarkoordinaten (r, φ) und den kartesischen Koordinaten (x, y)

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

Eine Funktion f von (x, y) kann also durch

$$F(r, \varphi) := f(r \cos \varphi, r \sin \varphi)$$

als Funktion von (r, φ) betrachtet werden.

Die partiellen Ableitungen von F berechnen sich durch die Kettenregel als

$$\begin{aligned} F_r &= f_x \cos \varphi + f_y \sin \varphi, \\ F_\varphi &= -f_x r \sin \varphi + f_y r \cos \varphi. \end{aligned} \quad (4)$$

Dieses Gleichungssystem kann man nach f_x und f_y auflösen und erhält für $r \neq 0$

$$\begin{aligned} f_x &= F_r \cos \varphi - \frac{1}{r} F_\varphi \sin \varphi, \\ f_y &= F_r \sin \varphi + \frac{1}{r} F_\varphi \cos \varphi. \end{aligned} \quad (5)$$

Es gibt Formeln auch für höhere partielle Ableitungen. Insbesondere haben wir

$$\Delta f := f_{xx} + f_{yy} = F_{rr} + \frac{1}{r} F_r + \frac{1}{r^2} F_{\varphi\varphi}, \quad (6)$$

welches die Formeln für den **Laplace-Operator** Δ in kartesischen Koordinaten und in Polarkoordinaten sind. Analoge Formeln gelten für die Polar- oder Kugelkoordinaten im \mathbb{R}^3 . (Pierre Simon L., 1749–1827).

7.3 Anwendungen

7.3.1 (Richtung des stärksten Anstiegs) Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ eine C^1 -Funktion und $\mathbf{x}_0 \in D$. Es werde $\text{grad } f(\mathbf{x}_0) \neq \mathbf{0}$ vorausgesetzt. Für die Richtungsableitung in Richtung \mathbf{v} mit $\|\mathbf{v}\| = 1$ gilt

$$\begin{aligned}\partial_{\mathbf{v}} f(\mathbf{x}_0) &= \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} \\ &= \|\mathbf{v}\| \|\text{grad } f(\mathbf{x}_0)\| \cos \alpha,\end{aligned}$$

wobei α der Winkel zwischen \mathbf{v} und $\text{grad } f(\mathbf{x}_0)$ ist. Somit ist der Wert der Richtungsableitung maximal, wenn \mathbf{v} und $\text{grad } f(\mathbf{x}_0)$ in die gleiche Richtung zeigen. Oder: Richtung von $\text{grad } f(\mathbf{x}_0)$ = Richtung des stärksten Anstiegs in \mathbf{x}_0 .

Satz 7.3.2 (Taylor-Formel) $f \in C^{k+1}(D)$, $\mathbf{x} \in D$, $\|\mathbf{v}\| = 1$, $t > 0$. Die Verbindungsstrecke von \mathbf{x} und $\mathbf{x} + t\mathbf{v}$ liege in $D \setminus \partial D$. Dann gilt

$$f(\mathbf{x} + t\mathbf{v}) = f(\mathbf{x}) + \partial_{\mathbf{v}} f(\mathbf{x})t + \frac{1}{2!} \partial_{\mathbf{v}}^2 f(\mathbf{x})t^2 + \cdots + \frac{1}{k!} \partial_{\mathbf{v}}^k f(\mathbf{x})t^k + R_{f,k}(\mathbf{x}, t, \mathbf{v})$$

mit dem Restglied

$$R_{f,k}(\mathbf{x}, t, \mathbf{v}) = \frac{1}{(k+1)!} \partial_{\mathbf{v}}^{k+1} f(\mathbf{x} + \xi\mathbf{v})t^{k+1},$$

wobei ξ eine Zahl $\in (0, 1)$ ist.

- Sei $f \in C^{k+1}(D)$, dann ist für $\mathbf{x} - \mathbf{x}_0 = t\mathbf{v}$, $\|\mathbf{v}\| = 1$, $t > 0$ und $0 \leq j \leq k$

$$T_{f,j}(\mathbf{x}, \mathbf{x}_0) := \sum_{i=0}^j \partial_{\mathbf{v}}^i f(\mathbf{x}_0)t^i \quad (1)$$

das j -te Taylor-Polynom von f an der Stelle \mathbf{x}_0 . Es gilt

$$f(\mathbf{x}) = T_{f,j}(\mathbf{x}, \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^j). \quad (2)$$

Definition 7.3.3 Für eine C^2 -Funktion f heißt die symmetrische Matrix

$$\mathcal{H}_f(\mathbf{x}) := \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) \right)_{i,j=1,\dots,n}$$

die **Hesse-Matrix** von f an der Stelle \mathbf{x} . (Ludwig Otto H., 1811–1874).

- Bisher hatten wir es mit Funktionen der Form

$$y = g(x) \quad (*)$$

zu tun. Aber man kann auch eine Abhängigkeit zwischen x und y implizit durch eine Gleichung

$$f(x, y) = 0$$

vorgeben. Wie kann man daraus eine Funktion der Form (*) gewinnen?

Definition 7.3.4 Sei $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$. Man sagt, durch $f(x, y) = 0$ ist auf dem Intervall $I \subseteq \mathbb{R}$ eine **implizite Funktion** $g : I \rightarrow K$ mit Werten in $K \subseteq \mathbb{R}$ erklärt, wenn es zu jedem $x \in I$ genau ein $y \in K$ gibt mit $(x, y)^T \in D$ und $f(x, y) = 0$. Dieses y wird als $g(x)$ bezeichnet.

Beispiele:

- Sei $f(x, y) = 3x + 2y - 1 = 0$. Für alle x ist diese Gleichung eindeutig nach y auflösbar. Wir erhalten eine explizite Darstellung $y = g(x) = \frac{1 - 3x}{2}$.
- Sei $f(x, y) = e^y + y^3 + x^3 + x^2 - 1 = 0$. Für alle x hat $e^y + y^3 = 1 - x^2 - x^3$ genau eine Lösung $g(x)$, denn $h(y) = e^y + y^3$ ist strikt monoton steigend. Aber g ist nicht durch elementare Umformungen darstellbar. Daher ist nur die Existenz einer Funktion $g(x)$ mit

$$e^{g(x)} + g(x)^3 = 1 - x^2 - x^3$$

gesichert.

Satz 7.3.5 (über implizite Funktionen) Sei $D \in \mathbb{R}^2$ offen und $f : D \rightarrow \mathbb{R}$ eine C^1 -Funktion. Ist $(x_0, y_0)^T \in D$ ein Punkt der Niveaumenge $f(x, y) = 0$ mit $f_y(x_0, y_0) \neq 0$, dann gibt es Intervalle $I \subseteq \mathbb{R}$ und $K \subseteq \mathbb{R}$ mit den Mittelpunkten x_0 bzw. y_0 , so das gilt:

- $R = I \times K \subseteq D$, $f_y(x, y) \neq 0$ für alle $(x, y) \in R$.
- Es gibt eine eindeutig bestimmte, differenzierbare Funktion $g : I \rightarrow K$ mit:

$$f(x, y) = 0 \Leftrightarrow y = g(x).$$

Für die Ableitung von g gilt

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} \quad \forall x \in I.$$

- Nachdem wir wissen, dass g differenzierbar ist, kann man die Formel aus b) leichter einsehen durch die Kettenregel.

$$0 = f(x, g(x)) \quad \Rightarrow \quad 0 = f_x(x, g(x)) + f_y(x, g(x))g'(x)$$

Man kann auch höhere Ableitungen berechnen; z.B.

$$f_{xx} + f_{xy}g' + f_{yx}g' + f_{yy}(g')^2 + f_y g'' = 0.$$

Also gilt:

$$g'(x) = \frac{-f_x}{f_y}$$

$$g''(x) = -\frac{1}{f_y} (f_{xx} + f_{xy}g' + f_{yx}g' + f_{yy}(g')^2)$$

$$= -\frac{1}{(f_y)^3} (f_{xx} (f_y)^2 - 2f_{xy} f_x f_y + f_{yy} (f_x)^2) .$$

Somit kann man Extremstellen der impliziten Funktion $g(x)$ bestimmen:
Seien x und $y = g(x)$ derart, dass

$$f(x, y) = 0 \quad f_x(x, y) = 0 \quad f_y(x, y) \neq 0$$

dann ist x ein lokales Maximum der Funktion $g(x)$, falls

$$\frac{f_{xx}(x, y)}{f_y(x, y)} > 0$$

ist und x ist ein lokales Minimum, falls $\frac{f_{xx}(x, y)}{f_y(x, y)} < 0$ ist.

Beispiel: Sei $f(x, y) = e^y + y^3 + x^3 + x^2 - 1$. Dann gilt:

$$f_x = 3x^2 + 2x, \quad f_y = e^y + 3y^2, \quad f_{xx} = 6x + 2 .$$

Sei g die in Satz 7.3.5 b) definierte Funktion, so gilt:

$$g'(x) = -\frac{f_x}{f_y} = -\frac{3x^2 + 2x}{e^{g(x)} + 3(g(x))^2} .$$

Hat g an der Stelle x_0 eine Extremstelle, d.h. $g'(x_0) = 0$, so gilt:

$$0 = \frac{f_x(x_0, g(x_0))}{f_y(x_0, g(x_0))} = \frac{3x_0^2 + 2x_0}{e^{g(x_0)} + 3(g(x_0))^2} .$$

Dies ergibt $x_0 = 0$ oder $x_0 = -\frac{2}{3}$. Wir haben

$$g''(x_0) = -\frac{f_{xx}(x_0, g(x_0))}{f_y(x_0, g(x_0))} = -\frac{6x_0 + 2}{e^{g(x_0)} + 3(g(x_0))^2} .$$

Also ist $g''(0) = -2$ und damit ist 0 ein Maximum.

Weiter gilt $g''(-\frac{2}{3}) = 2$ und damit ist $-\frac{2}{3}$ ein Minimum.

Definition 7.3.6 Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Ein Punkt $\mathbf{a} \in D$ heißt **lokales Maximum** (bzw. **Minimum**) von f , wenn es eine r -Umgebung $B_r(\mathbf{a})$ gibt so, dass für alle $\mathbf{x} \in B_r(\mathbf{a}) \cap D$ gilt

$$f(\mathbf{x}) \leq f(\mathbf{a}) \quad (\text{ bzw. } f(\mathbf{a}) \leq f(\mathbf{x})) .$$

Falls diese Ungleichungen für alle $\mathbf{x} \in D$ gelten, heißt \mathbf{a} **globales Maximum** (bzw. **Minimum**). Ein **Extremum** ist ein Minimum oder ein Maximum.

Satz 7.3.7 (Lokale Extrema im Innern) Ist f auf $B_r(\mathbf{a})$, $\mathbf{a} \in \mathbb{R}^n$, eine C^1 -Funktion, so gilt:

$$\mathbf{a} \text{ ist lokale Extremstelle} \quad \Rightarrow \quad \text{grad } f(\mathbf{a}) = \mathbf{0} .$$

- Ein Punkt $\mathbf{a} \in \mathbb{R}^n$ für den $\text{grad } f(\mathbf{a}) = \mathbf{0}$ gilt, heißt **stationärer Punkt**. Ein stationärer Punkt, der kein Extrempunkt ist, heißt **Sattelpunkt**.
- Satz 7.3.7 besagt, dass man Extrempunkte im Innern zwischen den stationären Punkten suchen muss.

Satz 7.3.8 Ist $D \subseteq \mathbb{R}^n$ offen, $\mathbf{a} \in D$ ein stationärer Punkt einer Funktion $f \in C^2(D, \mathbb{R})$ und $\mathcal{H}_f(\mathbf{a}) = (f_{x_i x_j}(\mathbf{a}))$ die Hesse-Matrix von f in \mathbf{a} . Dann gilt:

- Ist $\mathcal{H}_f(\mathbf{a})$ positiv definit, so ist \mathbf{a} ein lokales Minimum.
- Ist $\mathcal{H}_f(\mathbf{a})$ negativ definit, so ist \mathbf{a} ein lokales Maximum.
- Ist $\mathcal{H}_f(\mathbf{a})$ indefinit, so ist \mathbf{a} ein Sattelpunkt.

Folgerung 7.3.9 Sei $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ ein C^2 -Funktion und \mathbf{a} ein stationärer Punkt. Dann gilt:

- $f_{xx}(\mathbf{a}) > 0$, und $\det \mathcal{H}_f(\mathbf{a}) > 0$, so ist \mathbf{a} lokales Minimum,
- $f_{xx}(\mathbf{a}) < 0$, und $\det \mathcal{H}_f(\mathbf{a}) > 0$, so ist \mathbf{a} lokales Maximum,
- $\det \mathcal{H}_f(\mathbf{a}) < 0$, so ist \mathbf{a} Sattelpunkt.

Beweis:

Sei $\mathbf{A} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ negativ definit. Da \mathbf{A} ähnlich zu $\mathbf{B} = \begin{pmatrix} a & 0 \\ 0 & d - \frac{b^2}{a} \end{pmatrix}$ ist, muß auch \mathbf{B} negativ definit sein. Dies ist äquivalent zu

$$a < 0, \quad ad - b^2 > 0.$$

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch:

$$f(x, y) = 6xy - 3y^2 - 2x^3$$

Wir versuchen, die Extremstellen zu ermitteln:

$$\begin{aligned} f_x(x, y) &= 6y - 6x^2 \\ f_y(x, y) &= 6x - 6y \\ \text{grad } f(x, y) = 0 &\Leftrightarrow x = y \text{ und } x^2 = y \end{aligned}$$

Mögliche Kandidaten sind also $(0, 0)$ und $(1, 1)$. Zur genauen Bestimmung berechnen wir die Hessematrix

$$\begin{aligned} \mathcal{H}_f(x, y) &= \begin{pmatrix} -12x & 6 \\ 6 & -6 \end{pmatrix} \\ \mathcal{H}_f(0, 0) &= \begin{pmatrix} 0 & 6 \\ 0 & -6 \end{pmatrix} && \text{indefinit} \\ \mathcal{H}_f(1, 1) &= \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix} && \text{negativ definit} \end{aligned}$$

Damit liegt bei $(0, 0)$ ein Sattelpunkt und bei $(1, 1)$ ein Maximum vor.

7.4 Vektorwertige Funktionen

Für Funktionen $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ werden die Begriffe aus Paragraph 7.2 über die Komponentenfunktionen $f_i(\mathbf{x}), i = 1, \dots, m$, auf die vektorwertigen Funktionen $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T$ übertragen.

Definition 7.4.1 Sei $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m, k \in \mathbb{N}$ und $\mathbf{x}, \mathbf{x}_0 \in D$.

- a) Der **Grenzwert**, die **partiellen Ableitungen** und die „**klein-o**“-Notation sind jeweils komponentenweise erklärt:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) := \begin{pmatrix} \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_1(\mathbf{x}) \\ \vdots \\ \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_m(\mathbf{x}) \end{pmatrix}; \quad \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(\mathbf{x}) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(\mathbf{x}) \end{pmatrix}$$

$$\mathbf{f}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^k) \Leftrightarrow f_j(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^k), \quad j = 1, \dots, m.$$

- b) \mathbf{f} ist genau dann **stetig, partiell differenzierbar** oder eine **C^k -Funktion**, wenn alle Komponentenfunktionen $f_i, i = 1, \dots, m$, stetig, partiell differenzierbar oder C^k -Funktionen sind.
- c) \mathbf{f} heißt im Punkt $\mathbf{x}_0 \in D$ **total differenzierbar**, wenn es eine Matrix $\mathcal{A} \in \mathbb{R}^{m \times n}$ und eine r -Umgebung $B_r(\mathbf{x}_0) \subseteq D$ gibt, so dass für alle $\mathbf{x} \in B_r(\mathbf{x}_0)$ gilt:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathcal{A} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|). \quad (1)$$

- Aus (1) folgt, dass die Komponentenfunktionen total differenzierbar sind. Also gilt für $k = 1, \dots, m$:

$$f_k(\mathbf{x}) = f_k(\mathbf{x}_0) + \text{grad } f_k(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|).$$

Somit ist \mathcal{A} in (1) eindeutig bestimmt als

$$\mathcal{A} = \mathcal{J}_{\mathbf{f}}(\mathbf{x}_0) := \begin{pmatrix} \text{grad } f_1(\mathbf{x}_0)^T \\ \vdots \\ \text{grad } f_m(\mathbf{x}_0)^T \end{pmatrix} = \left(\frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{x}_0) \right) = \left(\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right)_{i,j}.$$

$\mathcal{J}_{\mathbf{f}}(\mathbf{x}_0)$ heißt **Jacobi-Matrix** von \mathbf{f} an der Stelle \mathbf{x}_0 . (Carl Gustav J., 1804–1851).

Satz 7.4.2 Jede C^1 -Funktion $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (D offen) ist auf D total differenzierbar.

Beispiele:

- a) Sei $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear.

Dann existiert ein $\mathcal{A} \in \mathbb{R}^{m \times n}$ mit $\mathbf{f}(\mathbf{x}) = \mathcal{A} \cdot \mathbf{x}$. Also gilt $\mathcal{J}_{\mathbf{f}}(\mathbf{x}) = \mathcal{A}$.

b) Polarkoordinaten: Die Menge $[0, \infty) \times [0, 2\pi)$ wird mittels

$$\mathbf{x}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

auf die (x_1, x_2) -Ebene abgebildet, d.h. $\mathbf{x} : D \rightarrow \mathbb{R}^2$. Die Jacobi Matrix ist

$$\mathcal{J}_{\mathbf{x}}(r, \varphi) = \left(\frac{\partial \mathbf{x}}{\partial r}, \frac{\partial \mathbf{x}}{\partial \varphi} \right) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Satz 7.4.3 (Kettenregel) Sind $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ in $\mathbf{x}_0 \in D$ und $\mathbf{g} : G \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^q$ in $\mathbf{f}(\mathbf{x}_0) \in G$ total differenzierbar, dann ist die Komposition $\mathbf{g} \circ \mathbf{f}$ in \mathbf{x}_0 ebenfalls total differenzierbar und es gilt:

$$\mathcal{J}_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}_0) = \mathcal{J}_{\mathbf{g}}(\mathbf{f}(\mathbf{x}_0)) \cdot \mathcal{J}_{\mathbf{f}}(\mathbf{x}_0).$$

Definition 7.4.4 Jedem C^1 -Skalarenfeld $f : D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ wird das Vektorfeld „**Gradient**“ $\text{grad } f : D \rightarrow \mathbb{R}^3$

$$\text{grad } f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \frac{\partial f}{\partial x_3}(\mathbf{x}) \right)^T$$

zugeordnet. Jedem C^2 -Skalarenfeld $f : D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ ordnet der **Laplace Operator** Δ das Skalarenfeld $\Delta f : D \rightarrow \mathbb{R}$

$$\Delta f(\mathbf{x}) = \sum_{i=1}^3 \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$$

zu. Zu jedem C^1 -Vektorfeld $\mathbf{V} : D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definiert man ein Skalarenfeld **Divergenz** $\text{div } \mathbf{V} : D \rightarrow \mathbb{R}$ und ein Vektorfeld **Rotation** $\text{rot } \mathbf{V} : D \rightarrow \mathbb{R}^3$ durch die Vorschriften:

$$\text{div } \mathbf{V}(\mathbf{x}) = \sum_{i=1}^3 \frac{\partial f}{\partial x_i}(\mathbf{x}),$$

$$\text{rot } \mathbf{V}(\mathbf{x}) = \begin{pmatrix} \frac{\partial v_3}{\partial x_2}(\mathbf{x}) - \frac{\partial v_2}{\partial x_3}(\mathbf{x}) \\ \frac{\partial v_1}{\partial x_3}(\mathbf{x}) - \frac{\partial v_3}{\partial x_1}(\mathbf{x}) \\ \frac{\partial v_2}{\partial x_1}(\mathbf{x}) - \frac{\partial v_1}{\partial x_2}(\mathbf{x}) \end{pmatrix}.$$

Lemma 7.4.5 Für ein C^2 -Skalarenfeld f und ein C^2 -Vektorfeld \mathbf{V} gelten folgende Rechenregeln:

- a) $\text{rot}(\text{grad } f) = 0$,
- b) $\text{div}(\text{rot } \mathbf{V}) = 0$,

$$c) \operatorname{div}(\operatorname{grad} f) = \Delta f,$$

$$d) \operatorname{div}(f\mathbf{V}) = \operatorname{grad} f \cdot \mathbf{V} + f \operatorname{div} \mathbf{V},$$

$$e) \operatorname{rot}(f\mathbf{V}) = \operatorname{grad} f \times \mathbf{V} + f \operatorname{rot} \mathbf{V},$$

$$f) \operatorname{rot}(\operatorname{rot} \mathbf{V}) = \operatorname{grad}(\operatorname{div} \mathbf{V}) - \Delta \mathbf{V}.$$

7.4.6 (Nabla-Kalkül) Mit dem symbolischen „Vektor“ $\nabla := \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)^T$ („**Nabla Operator**“) kann man formal schreiben

$$\begin{aligned} \operatorname{grad} f &= \nabla f, \\ \operatorname{div} \mathbf{V} &= \nabla \cdot \mathbf{V}, \\ \Delta f &= \nabla \cdot \nabla f, \\ \operatorname{rot} \mathbf{V} &= \nabla \times \mathbf{V}. \end{aligned}$$

Kapitel 8

Integration von Funktionen mehrerer Variablen

8.1 Parameterintegrale

Viele Funktionen der Analysis besitzen Integraldarstellungen, die von einem Parameter abhängen, z.B.

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad x > 0 \quad \text{Gammafunktion}$$
$$F(x) = \int_0^{\infty} f(t) e^{-xt} dt \quad \text{Laplace-Transformierte}$$

Satz 8.1.1 Sei

$$D = [a, b] \times [c, d] = \{(x, y)^T \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\}$$

ein abgeschlossenes Rechteck in \mathbb{R}^2 und $f : D \rightarrow \mathbb{R}$ stetig.
Dann gilt für die Integralfunktion

$$F(x) = \int_c^d f(x, y) dy, \quad x \in [a, b],$$

1) F ist auf $[a, b]$ stetig.

2)

$$\int_a^b F(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy,$$

(Spezialfall des **Satzes von Fubini**, Guido F., 1879–1943).

3) Ist f auf $[a, b]$ stetig partiell nach x differenzierbar, dann ist F differenzierbar mit der Ableitung

$$F'(x) = \frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d \frac{\partial f}{\partial x}(x, y) dy.$$

Satz 8.1.2 Seien g, h C^1 -Funktionen und f stetig partiell nach x differenzierbar. Dann gilt:

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, y) dy = \int_{g(x)}^{h(x)} \frac{\partial f}{\partial x}(x, y) dy + f(x, h(x))h'(x) - f(x, g(x))g'(x).$$

Beispiel: Sei $g : \mathbb{R} \rightarrow \mathbb{R}$ stetig, $k \in \mathbb{R} \setminus \{0\}$,

$$F(x) = \frac{1}{k} \int_0^x g(y) \sin(k(x-y)) dy.$$

Dann gilt

$$F'(x) = \int_0^x g(y) \cos(k(x-y)) dy,$$

$$F''(x) = -k \int_0^x g(y) \sin(k(x-y)) dy + g(x),$$

d.h. F ist eine Lösung der Schwingungsgleichung

$$F''(x) + k^2 F(x) = -g(x).$$

- Falls das Integral $F(x)$ in 8.1.1 uneigentlich ist, d.h.

$$\int_c^d f(x, y) dy = \lim_{\varepsilon \rightarrow 0} \int_c^{d-\varepsilon} f(x, y) dy$$

oder

$$\int_c^\infty f(x, y) dy = \lim_{d \rightarrow \infty} \int_c^d f(x, y) dy$$

gilt Satz 8.1.1 nur unter zusätzlichen Voraussetzungen:

Satz 8.1.3 Sei $D = \{(x, y)^T \in \mathbb{R}^2 : a \leq x \leq b, c \leq y < d\}$ und sei f auf dem einseitig offenen Rechteck D stetig und stetig partiell nach x differenzierbar und es gebe Funktionen $g, h : [c, d) \rightarrow \mathbb{R}$, so dass

a) $|f(x, y)| \leq g(y)$, $|f_x(x, y)| \leq h(y)$ $\forall (x, y)^T \in D$ und

b) die uneigentlichen Integrale $\int_c^d g(y) dy$, $\int_c^d h(y) dy$ konvergieren.

Dann konvergiert für jedes $x \in [a, b]$ das uneigentliche Integral

$$F(x) := \int_c^d f(x, y) dy$$

und die so definierte Funktion F hat die Ableitung

$$F'(x) = \int_c^d f_x(x, y) dy.$$

Beispiel: Sei $f : [0, \infty)$ stückweise stetig und zu jedem $\varepsilon > 0$ gebe es ein $C(\varepsilon)$, so dass

$$|f(t)| \leq C(\varepsilon)e^{\varepsilon t}, \quad t \geq 0.$$

Dann existiert für jedes $x > 0$ die Laplace-Transformierte F von f $F(x) = \int_0^{\infty} f(t)e^{-xt} dt$.

(Anstelle $x > 0$ kann auch $z \in \mathbb{C}$ mit $\operatorname{Re} z > 0$ genommen werden).

F ist beliebig oft differenzierbar und es gilt

$$F'(x) = - \int_0^{\infty} t f(t)e^{-xt} dt, \quad F''(x) = \int_0^{\infty} t^2 f(t)e^{-xt} dt.$$

8.2 Integrale auf rechteckigen Bereichen

Definition 8.2.1 1) Seien $R = [a, b] \times [c, d]$ ein Rechteck in \mathbb{R}^2 und $f : R \rightarrow \mathbb{R}$ stetig. R werde in einander nicht überlappende achsenparallele Rechtecke R_1, \dots, R_k zerlegt. Δ_j sei der Flächeninhalt von R_j . $\xi_j \in R_j$ ($j = 1, \dots, k$). Dann heißt (wie im Eindimensionalen)

$$Z_k = Z_k(f, \xi) = \sum_{j=1}^k f(\xi_j) \Delta_j$$

eine **Riemannsche Summe** oder kurz **Riemann-Summe**.

2) Bei beliebiger Verfeinerung der Zerlegung, d.h. der Durchmesser aller R_j gehe gegen Null, konvergieren die Riemann-Summen gegen einen eindeutig bestimmten Wert. Dieser wird als $\int_R f(x, y) d(x, y)$, das **Riemann-Integral** (oder kurz: **Integral**) von f auf R definiert.

Satz 8.2.2 f stetig auf $R = [a, b] \times [c, d]$. Dann gilt

$$\int_R f(x, y) d(x, y) = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

(Satz von Fubini).

Definition und Satz 8.2.3 8.2.1 und 8.2.2 gelten sinngemäß für drei- und höherdimensionale Rechtecke.

8.3 Integrale auf Normalbereichen

Im Allgemeinen sind die Integrationsbereiche im Mehrdimensionalen keine Rechtecke. Es müssen daher zuerst einmal vernünftige Bereiche definiert werden.

Definition 8.3.1 1) $B \subseteq \mathbb{R}^2 = \{(x_1, x_2)^T; x_1, x_2 \in \mathbb{R}\}$ heißt **in x_1 -Richtung projizierbar** oder **Normalbereich in x_1 -Richtung**, wenn es $a_1, b_1 \in \mathbb{R}$ mit $a_1 \leq b_1$ und stetige

$$g_1, g_2 : I_1 = [a_1, b_1] \rightarrow \mathbb{R} \quad \text{mit} \quad g_1 \leq g_2 \quad \text{auf} \quad I_1 \quad \text{gibt,}$$

so dass

$$B = \{(x_1, x_2)^T : x_1 \in I_1, g_1(x_1) \leq x_2 \leq g_2(x_1)\}.$$

Analog in x_2 -Richtung projizierbar bzw. Normalbereich in x_2 -Richtung mit $I_2 = [a_2, b_2]$, stetigen $h_1, h_2 : I_2 \rightarrow \mathbb{R}, h_1 \leq h_2$ und

$$B = \{(x_1, x_2)^T : x_2 \in I_2, h_1(x_2) \leq x_1 \leq h_2(x_2)\}.$$

2) B heißt **zweidimensionaler Normalbereich**, wenn er x_1 - und x_2 -projizierbar ist.

Beispiel: Die abgeschlossene Kreisscheibe

$$B = \{(x_1, x_2)^T; x_1^2 + x_2^2 \leq R^2\}$$

ist Normalbereich. Z.B. $I_1 = [-R, R], g_1(x_1) = -(R^2 - x_1^2)^{1/2}, g_2(x_1) = (R^2 - x_1^2)^{1/2}$.

Hinweis: Oft ist ein Bereich nicht normal, kann aber in endlich viele Normalbereiche zerlegt werden, z.B. der Kreisring

$$B = \{(x_1, x_2)^T; R_1^2 \leq x_1^2 + x_2^2 \leq R_2^2\}, \quad 0 < R_1 < R_2.$$

Das Integral $\int_B f(x_1, x_2) d(x_1, x_2)$ kann ähnlich wie in 8.3.1 über Riemann-Summen eingeführt werden. Hierbei treten bei einer Zerlegung in achsenparallele Rechtecke solche auf, die nur zum Teil zu B gehören. Bei Projizierung kann das Integral als Doppelintegral eingeführt werden.

Satz 8.3.2 B zweidimensionaler Normalbereich (s. 8.3.1). $f : B \rightarrow \mathbb{R}$, stetig. Dann gilt

$$\int_B f(x_1, x_2) d(x_1, x_2) := \int_{a_1}^{b_1} \left(\int_{g_1(x_1)}^{g_2(x_1)} f(x_1, x_2) dx_2 \right) dx_1 = \int_{a_2}^{b_2} \left(\int_{h_1(x_2)}^{h_2(x_2)} f(x_1, x_2) dx_1 \right) dx_2.$$

Definition 8.3.3 1) $B \subseteq \mathbb{R}^3$ heißt **in x_3 -Richtung projizierbar** oder **Normalbereich in x_3 -Richtung**, wenn es einen Normalbereich $B_{12}^* \subseteq \mathbb{R}^2$ und stetige $h_1, h_2 : B_{12}^* \rightarrow \mathbb{R}$ mit $h_1 \leq h_2$ gibt, so dass

$$B = \{(x_1, x_2, x_3)^T; (x_1, x_2)^T \in B_{12}^*, h_1(x_1, x_2) \leq x_3 \leq h_2(x_1, x_2)\}.$$

2) B heißt **dreidimensionaler Normalbereich**, wenn er in alle Richtungen projizierbar ist.

Definition bzw. Satz 8.3.4 $B \subseteq \mathbb{R}^3$ Normalbereich. B wie in 8.3.3,

$B_{12}^* = \{(x_1, x_2)^T; a_1 \leq x_1 \leq b_1, g_1(x_1) \leq x_2 \leq g_2(x_1)\}$. $f : B \rightarrow \mathbb{R}$, stetig. Dann kann das über Riemann-Summen eingeführte Integral $\int_B f(x_1, x_2, x_3) d(x_1, x_2, x_3)$ berechnet werden durch

$$\int_{a_1}^{b_1} \left(\int_{g_1(x_1)}^{g_2(x_1)} \left(\int_{h_1(x_1, x_2)}^{h_2(x_1, x_2)} f(x_1, x_2, x_3) dx_3 \right) dx_2 \right) dx_1.$$

Ebenso – mit den entsprechenden Begrenzungsfunktionen – mit Hilfe der x_1 - oder x_2 -Projizierbarkeit.

Beispiel: $R > 0$.

$B = \{(x_1, x_2, x_3)^T; x_1^2 + x_2^2 + x_3^2 \leq R^2\}$ (abgeschlossene Kugel vom Radius R mit dem Mittelpunkt $(0, 0, 0)^T$). f stetig auf B . Dann gilt

$$\int_B f(x_1, x_2, x_3) d(x_1, x_2, x_3) = \int_{-R}^R \left(\int_{-(R^2-x_1^2)^{1/2}}^{(R^2-x_1^2)^{1/2}} \left(\int_{-(R^2-x_1^2-x_2^2)^{1/2}}^{(R^2-x_1^2-x_2^2)^{1/2}} f(x_1, x_2, x_3) dx_3 \right) dx_2 \right) dx_1.$$

Analog bei anderer Reihenfolge der x_i .

Definition 8.3.5 B Normalbereich in \mathbb{R}^2 bzw. \mathbb{R}^3 .

1) $\int_B d(x_1, x_2)$ ist der **Flächeninhalt** von B , falls $B \subseteq \mathbb{R}^2$

2) $\int_B d(x_1, x_2, x_3)$ ist das **Volumen** von B , falls $B \subseteq \mathbb{R}^3$

3) $\int_B x_i d(x_1, x_2, x_3)$ ($i = 1, 2, 3$) heißen die **Schwerpunkt-Koordinaten** von B .

Integrale über krummlinig begrenzte Bereiche B können schwer zu berechnen sein. Man versucht oft, Koordinaten-Transformationen $(y_1, y_2)^T \rightarrow (x_1, x_2)^T$ bzw. $(y_1, y_2, y_3)^T \rightarrow (x_1, x_2, x_3)^T$ so zu finden, dass B in den y -Koordinaten ein Rechteck ist.

8.3.6 (Substitutionsregel für Polarkoordinaten in \mathbb{R}^2 bzw. \mathbb{R}^3)

1) $R > 0$, $B = \{(x_1, x_2)^T; x_1^2 + x_2^2 \leq R^2\}$. Die Abbildung

$$\mathbf{x} : \left\{ \begin{array}{l} (0, R] \times [0, 2\pi) \rightarrow B \setminus \{\mathbf{0}\} \\ \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \mathbf{x}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \end{array} \right\}$$

ist bijektiv. Die Jacobi-Matrix

$$\mathcal{J}_{\mathbf{x}}(r, \varphi) = \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \varphi} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \varphi} \end{pmatrix} \text{ hat die Gestalt } \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Die **Funktionaldeterminante** $\text{Det } \mathcal{J}_{\mathbf{x}}(r, \varphi)$ hat den Wert r .
 f sei stetig auf B . Dann gilt

$$\begin{aligned} \int_B f(x_1, x_2) d(x_1, x_2) &= \int_0^R \left(\int_0^{2\pi} r f(r \cos \varphi, r \sin \varphi) d\varphi \right) dr \\ &= \int_0^{2\pi} \left(\int_0^R r f(r \cos \varphi, r \sin \varphi) dr \right) d\varphi \end{aligned}$$

2) $R > 0$, $B = \{(x_1, x_2, x_3)^T; x_1^2 + x_2^2 + x_3^2 \leq R^2\}$.

Die Koordinaten $(r, \varphi, \vartheta)^T$, (φ : geographische Länge, ϑ : geographische Breite)
 $0 \leq r \leq R$, $0 \leq \varphi < 2\pi$, $-\pi/2 \leq \vartheta \leq \pi/2$ mit $x_1 = r \cos \vartheta \cos \varphi$,
 $x_2 = r \cos \vartheta \sin \varphi$, $x_3 = r \sin \vartheta$ werden als **spärische** oder **Kugelkoordinaten**
bezeichnet. Die Zuordnung $\mathbf{x} : (r, \varphi, \vartheta) \rightarrow (x_1, x_2, x_3)$ ist im Wesentlichen bijektiv.
Die Funktionaldeterminante $\text{Det } \mathcal{J}_{\mathbf{x}}(r, \varphi, \vartheta)$ hat den Wert $r^2 \cos \vartheta$.
Sei $f : B \rightarrow \mathbb{R}$, stetig. Dann gilt

$$\begin{aligned} &\int_B f(x_1, x_2, x_3) d(x_1, x_2, x_3) \\ &= \int_0^R \left(\int_0^{2\pi} \left(\int_{-\pi/2}^{\pi/2} r^2 \cos \vartheta f(r \cos \vartheta \cos \varphi, r \cos \vartheta \sin \varphi, r \sin \vartheta) d\vartheta \right) d\varphi \right) dr. \end{aligned}$$

Hinweis auf die **allgemeine Substitutionsregel**. B Normalbereich in \mathbb{R}^2 (oder in endlich viele Normalbereiche in \mathbb{R}^2 zerlegbar). Es gebe eine C^1 - und bijektive Variablentransformation $\mathbf{x} : (y_1, y_2)^T \rightarrow (x_1, x_2)^T$ ($(y_1, y_2)^T$ aus einem Rechteck $R \subseteq \mathbb{R}^2$, $(x_1, x_2) \in B$). Die Funktionaldeterminante $\text{Det } \mathcal{J}_{\mathbf{x}}(y_1, y_2)$ sei auf B ungleich Null. Dann gilt für stetiges $f : B \rightarrow \mathbb{R}$

$$\int_B f(x_1, x_2) d(x_1, x_2) = \int_R |\text{Det } \mathcal{J}_{\mathbf{x}}(y_1, y_2)| f(x_1(y_1, y_2), x_2(y_1, y_2)) d(y_1, y_2).$$

Analog im Dreidimensionalen.

8.4 Kurven- und Oberflächenintegrale

Vor allem für physikalische und technische Anwendungen werden weitere Integralbegriffe benötigt: Integration längs einer Kurve in einem Vektorfeld (Bewegung eines Massenpunktes längs einer Kurve in einem Kraftfeld und die dabei aufgewendete Arbeit) oder Integration über ein nicht planes Flächenstück im \mathbb{R}^3 mit einer skalaren Funktion oder einer Vektorfunktion als Integrand.

8.4.1 (Kurvenintegrale im \mathbb{R}^n) $n \geq 1$, $B \subseteq \mathbb{R}^n$ offen. $a < b$, $I = [a, b]$, $\mathbf{f} = (f_1, \dots, f_n)^T : I \rightarrow B$ stetig und stückweise stetig differenzierbar. Dann definiert das Paar (I, \mathbf{f}) eine **Parameterdarstellung einer Kurve** K in B (d.h. alle Kurvenpunkte $\mathbf{f}(t)$, $t \in I$ liegen in B).

$\mathbf{V} = (V_1, \dots, V_n)^T : B \rightarrow \mathbb{R}^n$ sei ein stetiges Vektorfeld auf B . Dann heißt das Integral

$$\int_K \mathbf{V} \cdot d\mathbf{x} \left(= \int_K (V_1 dx_1 + \dots + V_n dx_n) \right) := \int_a^b (\mathbf{V}(\mathbf{f}(t)) \cdot \mathbf{f}'(t)) dt$$

das **Kurvenintegral von \mathbf{V} längs K** .

Erläuterung. $\mathbf{V}(\mathbf{f}(t)) \cdot \mathbf{f}'(t)$ verschwindet, wenn der Tangentialvektor \mathbf{f}' senkrecht zum Vektor \mathbf{V} an der Stelle $\mathbf{f}(t)$ steht. Physikalisch: Auf einer Bahn, die senkrecht zum Kraftfeld verläuft, kann ein Massenpunkt ohne Aufwand bewegt werden. Das andere Extrem liegt vor, wenn Tangentenvektor und Kraftvektor $\mathbf{V}(\mathbf{f}(t))$ in die gleiche Richtung zeigen.

Definition 8.4.2 $n \geq 1$, $B \subseteq \mathbb{R}^n$ offen. $\mathbf{V} : B \rightarrow \mathbb{R}^n$ stetiges Vektorfeld auf B . $\varphi : B \rightarrow \mathbb{R}$ heißt **Potential** von \mathbf{V} , wenn $\varphi \in C^1(B)$ und $\mathbf{V} = \text{grad } \varphi$ ist. Falls \mathbf{V} ein Potential besitzt, heißt es **Potentialfeld** (oder konservatives Feld oder Gradientenfeld).

- Vorsicht: In der Physik werden Potentiale durch „ $-\mathbf{V} = \text{grad } \varphi$ “ eingeführt!

Satz 8.4.3 $B \subseteq \mathbb{R}^n$ offen. $\mathbf{V} : B \rightarrow \mathbb{R}^n$ stetiges Vektorfeld. B sei zusammenhängend, d.h. je zwei Punkte aus B seien durch einen Polygonzug in B bzw. eine stückweise glatte Kurve in B verbindbar. Dann sind folgende Aussagen äquivalent.

1) \mathbf{V} ist Potentialfeld.

2) Für alle $\mathbf{u}, \mathbf{v} \in B$ und alle stückweise glatten Kurven K_1 und K_2 in B mit Anfangspunkt \mathbf{u} und Endpunkt \mathbf{v} gilt

$$\int_{K_1} \mathbf{V} \cdot d\mathbf{x} = \int_{K_2} \mathbf{V} \cdot d\mathbf{x},$$

(d.h. das Kurvenintegral ist unabhängig von der Wahl der Kurve von \mathbf{u} nach \mathbf{v}).

- Falls (1) oder (2) gilt, kann ein Potential φ von \mathbf{V} wie folgt berechnet werden: $\mathbf{y}_0 \in B$ werde beliebig, aber fest gewählt. Für $\mathbf{y} \in B$ wähle man irgendeine stückweise glatte Kurve K von \mathbf{y}_0 nach \mathbf{y} .

Dann ist $\varphi(\mathbf{y}) = \int_K \mathbf{V} \cdot d\mathbf{x}$ ein Potential von \mathbf{V} .

8.4.4 (Rotation.) $B \subseteq \mathbb{R}^3$ offen. $\mathbf{V} = (V_1, V_2, V_3)^T$ stetig differenzierbares Vektorfeld auf B . $\mathbf{y} = (y_1, y_2, y_3)^T \in B$. Sei $\delta > 0$ so klein, dass $B_\delta(\mathbf{y}) \subseteq B$. Sei $K_{3,\delta}$ die Randkurve des zur (x_1, x_2) -Ebene parallelen Quadrats

$$Q = \{(x_1, x_2, y_3)^T; |x_1 - y_1| \leq \delta/2, |x_2 - y_2| \leq \delta/2\},$$

entgegen dem Uhrzeigersinn durchlaufen. Dann gilt

$$\lim_{\delta \rightarrow 0} \delta^{-2} \int_{K_{3,\delta}} \mathbf{V} \cdot d\mathbf{x} = \left(\frac{\partial V_2}{\partial x_1} - \frac{\partial V_1}{\partial x_2} \right)(\mathbf{y})$$

Ähnlich ergeben sich bei Integration parallel zur (x_1, x_3) - bzw. (x_2, x_3) -Ebene die zweite bzw. die erste Komponente der Rotation von \mathbf{V} in \mathbf{y} .

Definition 8.4.5 (Flächen im \mathbb{R}^3). $B \subseteq \mathbb{R}^2$ offen und zusammenhängend. $\overline{B} := B \cup \partial B$ (Abschluss von B). $\mathbf{f} = (f_1, f_2, f_3)^T : B \rightarrow \mathbb{R}^3$ stetig partiell differenzierbar nach beiden Variablen u und v . \mathbf{f} sei stetig auf \overline{B} fortsetzbar.

1) Das Paar $F = (B, \mathbf{f})$ heißt **Fläche im \mathbb{R}^3** , wenn die Ableitungsvektoren

$$\mathbf{f}_u = \frac{\partial \mathbf{f}}{\partial u} = \left(\frac{\partial f_1}{\partial u}, \frac{\partial f_2}{\partial u}, \frac{\partial f_3}{\partial u} \right)^T \quad \text{und} \quad \mathbf{f}_v = \frac{\partial \mathbf{f}}{\partial v} = \left(\frac{\partial f_1}{\partial v}, \frac{\partial f_2}{\partial v}, \frac{\partial f_3}{\partial v} \right)^T$$

auf B linear unabhängig sind, d.h. eine Ebene aufspannen.

2) $F = (B, \mathbf{f})$ Fläche. Dann heißt die von den Vektoren $\mathbf{f}_u(u_0, v_0)$ und $\mathbf{f}_v(u_0, v_0)$ erzeugte Ebene, angesetzt an den Punkt $\mathbf{f}(u_0, v_0) \in \mathbb{R}^3$, die **Tangentialebene** in $\mathbf{f}(u_0, v_0)$.

3) $F = (B, \mathbf{f})$ Fläche, $(u_0, v_0)^T \in B$. Der Vektor

$$\mathbf{N}(u_0, v_0) = \frac{\mathbf{f}_u(u_0, v_0) \times \mathbf{f}_v(u_0, v_0)}{\|\mathbf{f}_u(u_0, v_0) \times \mathbf{f}_v(u_0, v_0)\|}$$

heißt **Normalenvektor** von \mathbf{f} an der Stelle $\mathbf{f}(u_0, v_0)$.

- $\mathbf{N}(u_0, v_0)$ ist ein normierter, zur Tangentialfläche orthogonaler Vektor.
- Hat \mathbf{f} die spezielle Gestalt $\mathbf{f}(u, v) = (u, v, f_3(u, v))$ (d.h. die Fläche ist auf B projizierbar), dann gilt

$$\mathbf{N} = \frac{\left(-\frac{\partial f_3}{\partial u}, -\frac{\partial f_3}{\partial v}, 1\right)^T}{\left\|\left(-\frac{\partial f_3}{\partial u}, -\frac{\partial f_3}{\partial v}, 1\right)^T\right\|}.$$

- Lässt sich ∂B als Kurve $K_B = (I, \mathbf{g})$ in \mathbb{R}^2 darstellen und ist die Fortsetzung von \mathbf{f} auf K_B stetig, dann beschreibt

$$K_F : I \rightarrow \mathbb{R}^3, t \rightarrow (f_1(g_1(t), g_2(t)), f_2(g_1(t), g_2(t)), f_3(g_1(t), g_2(t)))^T$$

die **Randkurve** von F .

Definition 8.4.6 $A \subseteq \mathbb{R}^3$ offen. $F = (B, \mathbf{f})$ Fläche in A (d.h. für alle $(u, v)^T \in B$ gilt $\mathbf{f}(u, v) \in A$). B Normalbereich in \mathbb{R}^2 (d.h. als Integrationsbereich geeignet).

- 1) $g : A \rightarrow \mathbb{R}$ stetig (g skalare Funktion auf A). Dann heißt das Integral

$$\int_F g d\sigma := \int_B g(\mathbf{f}(u, v)) \|\mathbf{f}_u(u, v) \times \mathbf{f}_v(u, v)\| d(u, v)$$

das **Oberflächenintegral (erster Art)** von g über F . Insbesondere ist

$$I(F) := \int_F d\sigma = \int_B \|\mathbf{f}_u(u, v) \times \mathbf{f}_v(u, v)\| d(u, v)$$

der **Oberflächeninhalt** von F .

- Einem „kleinen“ achsenparallelen Rechteck in (u, v) mit den Seitenlängen δu und δv entspricht auf F in erster Näherung ein Parallelogramm mit den Seitenvektoren $\mathbf{f}_u(u, v)\delta u$ und $\mathbf{f}_v(u, v)\delta v$. Sein Flächeninhalt ist

$$\|\mathbf{f}_u(u, v) \times \mathbf{f}_v(u, v)\| \delta u \delta v.$$

- 2) $\mathbf{V} = (V_1, V_2, V_3)^T$ stetiges Vektorfeld auf A . Dann heißt das Integral

$$\int_F (\mathbf{V} \cdot \mathbf{N}) d\sigma := \int_B (\mathbf{V}(f(u, v))) \cdot (\mathbf{f}_u(u, v) \times \mathbf{f}_v(u, v)) d(u, v)$$

das **Oberflächenintegral (zweiter Art)** von \mathbf{V} über F .

- Im Fall eines Strömungsfeldes \mathbf{V} beschreibt das Integral den Fluss durch die Fläche F . Er ist maximal an Stellen, an denen der Flussvektor \mathbf{V} in Richtung des Normalenvektors $\mathbf{f}_u \times \mathbf{f}_v$ zeigt. Er verschwindet, wenn \mathbf{V} parallel zur Tangentialebene liegt, d.h. orthogonal zu $\mathbf{f}_u \times \mathbf{f}_v$ ist. Daher das Skalarprodukt $\mathbf{V} \cdot (\mathbf{f}_u \times \mathbf{f}_v)$.

8.4.7 (Divergenz.) \mathbf{V} sei ein stetig partiell differenzierbares Vektorfeld auf A (A wie in 8.4.6). $\mathbf{x} \in A$. Für genügend kleines $\delta > 0$ sei W_δ der achsenparallele Würfel mit Seitenlänge δ und Mittelpunkt \mathbf{x} . Die sechs Deckflächen F_1, \dots, F_6 von W_δ seien so parametrisiert (d.h. als (B_j, \mathbf{f}_j) dargestellt), dass die Normalenvektoren nach außen zeigen. Dann gilt

$$\lim_{\delta \rightarrow 0} \delta^{-3} \left(\int_{F_1} + \dots + \int_{F_6} \right) (\mathbf{V} \cdot \mathbf{N}) d\sigma = \operatorname{div} \mathbf{V}(\mathbf{x}).$$

Die Divergenz kann also als Flussdichte des Vektorfeldes interpretiert werden.

8.5 Die Integralsätze

Zu den wichtigsten Ergebnissen der mehrdimensionalen Analysis gehören die Integralsätze von Gauß, Green und Stokes, die es gestatten, gewisse Kurven- und Oberflächenintegrale ineinander umzuwandeln.

8.5.1 (Green'scher Satz) (auch Stokes'scher Satz in der Ebene. George Green, 1739–1841; Georg Gabriel Stokes, 1819–1903).

Seien $A \subseteq \mathbb{R}^2$ offen und $B \subseteq A$ ein abgeschlossener Normalbereich. Die Randkurve K von B sei stückweise glatt und entgegen dem Uhrzeigersinn durchlaufen.

$\mathbf{V} = (V_1, V_2)^T : A \rightarrow \mathbb{R}^2$ sei ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_K \mathbf{V} \cdot d\mathbf{x} = \int_B \left(\frac{\partial V_2}{\partial x_1} - \frac{\partial V_1}{\partial x_2} \right) d(x_1, x_2).$$

- In den Spezialfällen $V_1(x_1, x_2) = 0$, $V_2(x_1, x_2) = x_1$ bzw. $V_1(x_1, x_2) = -x_2$, $V_2(x_1, x_2) = 0$ ergibt sich für den Flächeninhalt F von B

$$F = \int_K (0, x_1) \cdot d\mathbf{x} = - \int_K (x_2, 0) \cdot d\mathbf{x}.$$

8.5.2 (Gaußscher Satz) .

Seien $A \subseteq \mathbb{R}^3$ offen und $B \subseteq A$ ein abgeschlossener Normalbereich. Der Rand von B lasse sich in endlich viele Flächenstücke zerlegen, die so parametrisiert sind, dass die Normalenvektoren nach außen zeigen.

$\mathbf{V} = (V_1, V_2, V_3)^T : A \rightarrow \mathbb{R}^3$ sei ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_{\partial B} (\mathbf{V} \cdot \mathbf{N}) d\sigma = \int_B \operatorname{div} \mathbf{V} d(x_1, x_2, x_3).$$

8.5.3 (Satz von Stokes) (in vereinfachter Form). $A \subseteq \mathbb{R}^3$ offen. Sei F ein Flächenstück, das zusammen mit seiner Randkurve K ganz in A liegt. K sei stückweise glatt und so durchlaufen, dass die Normalenvektoren auf F „nach oben“ zeigen.

Sei $\mathbf{V} = (V_1, V_2, V_3)^T : A \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_K \mathbf{V} \cdot d\mathbf{x} = \int_F (\text{rot } \mathbf{V} \cdot \mathbf{N}) d\sigma.$$

Folgerung 8.5.4 (zum Satz von Stokes.) $A \subseteq \mathbb{R}^3$ offen und konvex (konvex: zu je zwei Punkten aus A gehört die ganze Verbindungsstrecke zu A).

$\mathbf{V} = (V_1, V_2, V_3)^T \in C^1(A, \mathbb{R}^3)$. Dann sind folgende Aussagen äquivalent

- 1) \mathbf{V} ist Potentialfeld.
- 2) $\text{rot } \mathbf{V} = \mathbf{0}$ auf A .