

**Mathematik I und II**  
für Studierende des Ingenieurwesens

**Vorlesungsskript**

**Wintersemester 2011/12**

**Prof. Dr. Annette Huber-Klawitter**

**Sommersemester 2012**

**Prof. Dr. Sebastian Goette**

Literatur:

K. Meyberg, P. Vachnauer: Höhere Mathematik Bd 1, Springer Verlag.

Version vom 28. Juli 2012



# Inhaltsverzeichnis

<b>1</b>	<b>Grundbegriffe und Zahlbereiche</b>	<b>1</b>
1.1	Mengen und Abbildungen . . . . .	1
1.2	Reelle Zahlen . . . . .	4
1.3	Die Ebene und der Raum . . . . .	11
1.4	Komplexe Zahlen . . . . .	13
<b>2</b>	<b>Funktionen, Folgen, Reihen und Grenzwerte</b>	<b>19</b>
2.1	Grundbegriffe . . . . .	19
2.2	Polynome und rationale Funktionen . . . . .	21
2.3	Trigonometrische Funktionen . . . . .	27
2.4	Zahlenfolgen und Grenzwerte . . . . .	33
2.5	Unendliche Reihen . . . . .	38
2.6	Potenzreihen . . . . .	44
<b>3</b>	<b>Stetigkeit und Differentiation</b>	<b>49</b>
3.1	Funktionengrenzwerte, Stetigkeit . . . . .	49
3.2	Die Ableitung . . . . .	55
3.3	Anwendungen der Differentiation . . . . .	61
3.4	Umkehrfunktionen . . . . .	67
3.5	Exponential- und Logarithmusfunktion . . . . .	71
3.6	Beispiele . . . . .	75
<b>4</b>	<b>Integration</b>	<b>83</b>
4.1	Das bestimmte Integral . . . . .	83
4.2	Integrationsregeln . . . . .	90
4.3	Integration rationaler Funktionen . . . . .	95
4.4	Uneigentliche Integrale . . . . .	97
4.5	Kurven-, Längen- und Flächenmessung . . . . .	99

<b>5</b>	<b>Reihenentwicklungen</b>	<b>109</b>
5.1	Wiederholung und Ergänzung . . . . .	109
5.2	Darstellbarkeit durch Potenzreihen und Taylorentwicklung . .	111
5.3	Fourier-Reihen . . . . .	124
<b>6</b>	<b>Gewöhnliche Differentialgleichungen</b>	<b>127</b>
6.1	Allgemeines . . . . .	127
6.2	Lineare DGLs 1. Ordnung . . . . .	129
6.3	Existenz und Eindeutigkeit der Lösung . . . . .	133
6.4	Potenzreihenansatz . . . . .	134
6.5	Systeme von Differentialgleichungen . . . . .	140
<b>7</b>	<b>Lineare Algebra</b>	<b>143</b>
7.1	Lineare Gleichungssysteme und Matrizen . . . . .	143
7.2	Matrizenmultiplikation . . . . .	151
7.3	Vektorräume und Basen . . . . .	159
7.4	Matrizen und lineare Abbildungen . . . . .	168
7.5	Determinanten und Eigenwerte . . . . .	173
7.6	Skalarprodukt und orthogonale Abbildungen . . . . .	187
7.7	Symmetrische Matrizen . . . . .	191
<b>8</b>	<b>Differentiation von Funktionen in mehreren Variablen</b>	<b>195</b>
8.1	Kurven im $\mathbb{R}^n$ . . . . .	195
8.2	Reellwertige Funktionen mehrerer Veränderlicher . . . . .	198
8.3	Anwendungen . . . . .	205
8.4	Vektorwertige Funktionen . . . . .	215
<b>9</b>	<b>Integration von Funktionen mehrerer Variablen</b>	<b>227</b>
9.1	Parameterintegrale . . . . .	227
9.2	Kurvenintegrale . . . . .	230
9.3	Integration über ebene Bereiche . . . . .	241
9.4	Integration über Flächen im Raum . . . . .	253
9.5	Integration über dreidimensionale Bereiche . . . . .	264



# Kapitel 1

## Grundbegriffe und Zahlbereiche

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

### 1.1 Mengen und Abbildungen

**1.1.1.** *Eine Menge ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. (Georg Cantor, ca. 1875)*

**Bezeichnungen:**

- Objekte der Mengen heißen **Elemente**.

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

- Zwei Mengen sind gleich  $A = B$ , wenn sie dieselben Elemente haben.
- Beschreibung einer Menge durch Aufzählen ihrer Elemente:  
 $A = \{1, 2, 3\}$  hat genau die Elemente 1, 2 und 3. Es gilt

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

denn auf Reihenfolge und Nummerierung kommt es nicht an.

- Beschreibung einer Menge  $X$  durch Angeben einer definierenden Eigenschaft

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

- Die **leere Menge**  $\emptyset$  enthält kein Element.
- $B$  heißt **Teilmenge** von  $A$ , wenn jedes Element von  $B$  auch ein Element von  $A$  ist
- $B$  heißt **echte Teilmenge** von  $A$ , in Zeichen  $B \subsetneq A$ , wenn  $B \subseteq A$ , und es ein  $x \in A$  gibt mit  $x \notin B$ .

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

- Aussagen kann man miteinander verknüpfen. Diese Verknüpfungen werden in **Wahrheitstabellen** definiert.
- **Negation**  $\neg A$ : Es gilt genau das Gegenteil von  $A$ .
- **Konjunktion**  $A \wedge B$ : Es gilt  $A$  und  $B$  (beide gleichzeitig!)
- **Alternative** :  $A \vee B$ : Es gilt  $A$  oder  $B$  (oder beide!)
- **Implikation**  $A \Rightarrow B$ : Wenn  $A$  gilt, dann auch  $B$ . (Wenn  $A$  nicht gilt, dann kann  $B$  gelten oder auch nicht.)
- **Äquivalenz**:  $A$  gilt genau dann, wenn  $B$  gilt.  $A \Leftrightarrow B$

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	W	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

- **Existenzquantor**  $\exists$

$$\exists x \in M : E$$

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

- **Generalisationsquantor**  $\forall$

$$\forall x \in M : E$$

heißt: für alle  $x$  aus  $M$  gilt die Eigenschaft  $E$ .

In dieser Notation bedeutet  $A \subseteq B$ :

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

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

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

- 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.

- $A \times B$  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  $A_i, i = 1, \dots, n$ ,

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

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

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

**1.1.4 Definition.** 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 **Funktionswert von  $f$  an der Stelle  $x$** . Wir nennen  $x$  auch Argument.

- $A$  heißt **Definitionsbereich** von  $f$ , in Zeichen  $D(f) := A$ .
- 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 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:

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

- 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:**

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

Die **reellen Zahlen** werden axiomatisch eingeführt. Man betrachtet eine Menge auf der zwei Operationen  $+, \cdot$  definiert sind. Weiterhin müssen die Körperaxiome, die Ordnungsaxiome und das Vollständigkeitsaxiom erfüllt sein. Außerdem ist  $\mathbb{R}$  archimedisch (s.u.) Die Körperaxiome sichern, dass man wie gewohnt addieren, subtrahieren, multiplizieren und dividieren kann. Die Ordnungsaxiome sichern, dass man zwei beliebige reelle Zahlen der Größe nach vergleichen kann, d.h. es gilt immer genau eine der Möglichkeiten:

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

Man schreibt

$$x \leq y \Leftrightarrow (x < y) \vee (x = y).$$

Die Ordnungsrelation  $\leq$  ist kompatibel mit den Operationen  $\cdot, +$  und es gilt:

$$\begin{aligned} x \leq y, a \leq b &\Rightarrow a + x \leq b + y, \\ x \leq y, a \geq 0 &\Rightarrow a \cdot x \leq a \cdot y, \end{aligned} \quad (2.1)$$

$$\begin{aligned} x \leq y &\Leftrightarrow -y \leq -x, \\ 0 < x \leq y &\Leftrightarrow 0 < \frac{1}{y} \leq \frac{1}{x}. \end{aligned} \quad (2.2)$$

**1.2.1 Definition.** 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.

**1.2.2 Axiom. (Vollständigkeit)** Jede nicht-leere nach oben beschränkte Menge reeller Zahlen besitzt eine kleinste obere Schranke.

- Die kleinste obere Schranke heißt **Supremum**. Sei  $S \subseteq \mathbb{R}$  nach oben beschränkt. Dann setzt man

$$s := \sup S \Leftrightarrow \forall \varepsilon > 0 \exists x \in S: s - \varepsilon < x \leq s$$

- Die größte untere Schranke, nennt man **Infimum**.

**1.2.3 Axiom.** (*Archimedisches Axiom*) Zu jeder reellen Zahl  $x$  gibt es eine ganze Zahl  $n$ , die größer als  $x$  ist, also  $x < n$ .

**1.2.4 Definition.** 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:

$$\begin{aligned} -|a| \leq a \leq |a|, \\ |-a| = |a|, \end{aligned} \tag{2.3}$$

$$\begin{aligned} |ab| &= |a||b|, \\ \left| \frac{a}{b} \right| &= \frac{|a|}{|b|} \quad \text{falls } b \neq 0, \end{aligned} \tag{2.4}$$

$$|a| \leq b \quad \Leftrightarrow \quad -b \leq a \leq b. \tag{2.5}$$

- Aus (2.3) und (2.5) folgt die **Dreiecksungleichung**:  $\forall a, b \in \mathbb{R}$  gilt:

$$|a + b| \leq |a| + |b|. \tag{2.6}$$

- 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  $a, x \in \mathbb{R}, \varepsilon > 0$ :

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

**1.2.5 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  $(a, b], [a, b)$

- Zum Vermeiden von Fallunterscheidungen führen wir  $-\infty, \infty$  ein und legen fest:  $-\infty < x < \infty \quad \forall x \in \mathbb{R}$ . Somit gilt:

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

Also ist  $\mathbb{R}$  ein offenes Intervall.

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

**1.2.6 vollständige Induktion.** Aussagen oder Eigenschaften können von Elementen einer Menge  $A$  abhängen, z.B.

$$\forall n \in A: E(n).$$

Falls  $A = \{n_0, n_0 + 1, \dots\} \subseteq \mathbb{Z}$  kann man den Wahrheitsgehalt 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.

**1.2.7 Beispiel. (Bernoulli - Ungleichung:)**

$$\forall h \in (-1, \infty), \forall n \in \mathbb{N}: (1 + h)^n \geq 1 + nh \quad (2.7)$$

*Beweis.* Sei  $h \in (-1, \infty)$  fest. Für  $n = 1$  gilt

$$(1 + h)^1 = 1 + h = 1 + 1 \cdot h$$

Angenommen die Aussage gilt für  $n_0 \geq 1$ , d.h.

$$(1 + h)^{n_0} \geq 1 + n_0 h .$$

Wegen  $1 + h > 0$  folgt hieraus

$$(1 + h)^{n_0}(1 + h) \geq (1 + n_0 h)(1 + h) = 1 + n_0 h + h + n_0 h^2$$

Wegen  $h^2 \geq 0$  ist die rechte Seite  $\geq 1 + (n_0 + 1)h$ . Insgesamt folgt also

$$(1 + h)^{n_0+1} \geq 1 + (n_0 + 1)h .$$

Mit vollständiger Induktion gilt die Aussage für alle  $n_0 \in \mathbb{N}$ . □

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  $a_i \in \mathbb{R}$ ,  $m \leq n \in \mathbb{N}_0$ ,  $i = m, \dots, n$ ,  $m \leq n \in \mathbb{N}_0$ . 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 gilt die Ungleichung:

$$\left| \sum_{i=m}^n a_i \right| \leq \sum_{i=m}^n |a_i|. \quad (2.8)$$

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

$$\sum_{i=0}^n q^i = \frac{1 - q^{n+1}}{1 - q} \quad \text{falls } q \neq 1. \quad (2.9)$$

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

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

- Eine **Permutation** ist eine bijektive Abbildung einer endlichen Menge  $A$  auf sich selbst.

**1.2.9 Lemma.** *Zu jeder Menge mit  $n$  Elementen gibt es genau  $n!$  Permutationen.*

*Beweis.* Wir verteilen  $n$  Kugeln mit auf  $n$  Fächer. Wir zeigen mit vollständiger Induktion, dass es hierfür  $n!$  Möglichkeiten gibt.

Die Aussage gilt für  $n = 1$ . Sei nun die Aussage wahr für  $n$  Kugeln. Wir betrachten  $n + 1$  Kugeln. Für das erste Fach gibt es  $n + 1$  Möglichkeiten. Danach müssen wir  $n$  Kugeln auf  $n$  Fächer verteilen. Hierfür gibt es nach Induktionsannahme  $n!$  Möglichkeiten. Zusammen ergibt dies

$$(n + 1)n! = (n + 1)! .$$

□

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

*Beweis.* Wie im Beweis des Lemmas gibt es  $n(n - 1) \dots (n - k + 1)$  viele Möglichkeiten, aus  $n$  Kugeln  $k$ -viele auf Fächer zu verteilen. Hierbei werden jedoch verschiedene Anordnungen mehrfach gezählt. Es gibt  $k!$  Möglichkeiten,  $k$  Kugeln anzuordnen. Insgesamt also

$$\frac{n(n - 1) \dots (n - k + 1)}{k!} = \binom{n}{k}$$

□

- Für Binomialkoeffizienten gilt die Rekursionsformel:

$$\binom{n + 1}{k} = \binom{n}{k - 1} + \binom{n}{k}, \quad (2.10)$$

und folgende Rechenregeln:

$$\binom{n}{k} = \binom{n}{n - k}, \quad \binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{n - 1} = \binom{n}{1} = n.$$



## 1.3 Die Ebene und der Raum

Wir identifizieren die Ebene mit  $\mathbb{R}^2$ , indem wir das **kartesische Koordinatensystem** einführen.

- Man gibt den Punkt 0 in der Ebene vor,
- nimmt eine Zahlengerade, die  **$x$ -Achse**, so dass ihr Nullpunkt mit dem Punkt 0 übereinstimmt.
- Nun dreht man die  $x$ -Achse gegen den Uhrzeigersinn um  $90^\circ$  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

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

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

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  $f$** .

**1.3.1 Winkel.** Ein Winkel  $\alpha$  entsteht durch Drehung eines Zeigers um einen Punkt der Ebene.

- Die Länge des zugehörigen Einheitskreisbogens sei  $\ell$ . Wir nennen  $\ell$  das **Bogenmaß von  $\alpha$** , wenn die Drehung in positiver Richtung (gegen Uhrzeigersinn) erfolgte. Falls die Drehung in Urzeigersinn erfolgte ist  $-\ell$  das Bogenmaß.
- Ein Winkel  $\alpha$  habe das Gradmaß  $(\alpha_0)^\circ$ . Das Bogenmaß  $\ell$  des Winkels  $\alpha$  errechnet sich durch:

$$\ell = \frac{\pi}{180^\circ} (\alpha_0)^\circ$$

- Im Weiteren werden wir immer mit dem Bogenmaß arbeiten.

Genauso kann man den  $\mathbb{R}^3$  und den Anschauungsraum  $R$  miteinander identifizieren.

- **Kartesische Koordinaten** im Raum bestehen aus drei sich in einem Punkt 0 rechtwinklig schneidenden Zahlengeraden, die ein Rechtssystem bilden. Man nennt sie die  $x$ -,  $y$ -,  $z$ -Achsen.
- Die **Koordinatenebenen** sind die durch zwei Achsen aufgespannten Ebenen. Man nennt sie auch die  $(x, y)$ -,  $(x, z)$ - und  $(y, z)$ -Ebenen.

**1.3.2.** Gleichzeitig sind  $\mathbb{R}^2$  und  $\mathbb{R}^3$  **Vektorräume**, d.h. man kann die Elemente addieren und mit reellen Zahlen multiplizieren:

Seien  $\vec{v} = (x, y, z)$  und  $\vec{w} = (x', y', z')$  in  $\mathbb{R}^3$ ,  $\lambda \in \mathbb{R}$ :

$$\begin{aligned}\vec{v} + \vec{w} &= (x + x', y + y', z + z') \\ \vec{v} - \vec{w} &= (x - x', y - y', z - z') \\ a\vec{v} &= (ax, ay, cz)\end{aligned}$$

Die Elemente des Vektorraums heißen auch **Vektoren**. Man schreibt  $\vec{v}$  oder  $\underline{v}$ , aber auch einfach  $v$ . Das Element  $0 := \vec{0} := (0, 0, 0)$  heißt **Nullvektor**.

**1.3.3 Bemerkung.** Einen Vektor  $\vec{v} = (a, b, c)$  kann man auch als eine bijektive Abbildung vom Raum in sich selbst auffassen.  $\vec{v}$  ist die Parallelverschiebung um  $a$  in  $x$ -Richtung,  $b$  in  $y$ -Richtung und  $c$  in  $z$ -Richtung. Üblich ist auch die Notation  $\overrightarrow{PQ}$  für die Verschiebung, die den Punkt  $P$  auf den Punkt  $Q$  abbildet.

**1.3.4.** Seien  $\vec{a}, \vec{b} \in \mathbb{R}^3$ ,  $\lambda, \mu \in \mathbb{R}$ . Dann gelten die folgenden Rechenregeln.

$$\vec{a} + \vec{0} = \vec{a}, \quad (3.1)$$

$$\vec{a} + (-\vec{a}) = \vec{0}, \quad (3.2)$$

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}, \quad (3.3)$$

$$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}, \quad (3.4)$$

$$\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}, \quad (3.5)$$

$$(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}, \quad (3.6)$$

$$1\vec{a} = \vec{a}, \quad (3.7)$$

$$0\vec{a} = \vec{0}. \quad (3.8)$$

*Beweis.* Nachrechnen mit den Rechenregeln für  $\mathbb{R}$ . Z.B. für  $\vec{a} = (x, y, z)$

$$\vec{a} + \vec{0} = (x, y, z) + (0, 0, 0) = (x + 0, y + 0, z + 0) = (x, y, z) = \vec{a}$$

□

**1.3.5 Definition.** Die **Länge** bzw. die **Norm** eines Vektors  $\vec{a} = (x, y, z)$  ist die Länge der Strecke von  $(0, 0, 0)$  nach  $(x, y, z)$  im Raum

$$\|\vec{a}\| := \sqrt{x^2 + y^2 + z^2} .$$

**1.3.6.** Sei  $\vec{a}, \vec{b} \in \mathbb{R}^3$ ,  $\alpha \in \mathbb{R}$ . Dann gelten die folgenden Rechenregeln:

$$\|\alpha \vec{a}\| = |\alpha| \|\vec{a}\| \tag{3.9}$$

$$\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\| \tag{3.10}$$

*Beweis.* Die erste Aussage ist ganz leicht. Die zweite ist etwas mühsamer. Konzeptionell wird sie in der linearen Algebra bewiesen. Wir werden wenigstens den zweidimensionalen Fall noch sehen, aber in der Sprache der komplexen Zahlen. □

## 1.4 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

$$1 := (1, 0) \quad i = (0, 1)$$

und damit den Punkt  $z = (x, y)$  als

$$z := x + iy . \tag{4.1}$$

Wir 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

$$\mathbb{C} := \{x + iy; x, y \in \mathbb{R}\} \tag{4.2}$$

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

$$z = w \quad \iff \quad (x = u) \wedge (y = v) . \tag{4.3}$$

**1.4.1 Grundrechenarten in  $\mathbb{C}$ .** Seien  $z = x + iy$ ,  $w = u + iv$  mit  $x, y, u, v \in \mathbb{R}$  zwei komplexe Zahlen.

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

$$\begin{aligned} z + w &:= (x + u) + i(y + v), \\ z - w &:= (x - u) + i(y - v). \end{aligned} \quad (4.4)$$

- Die **Multiplikation** mit reellen Zahlen ist die Skalarmultiplikation von Vektoren:

$$\lambda z := \lambda x + i\lambda y.$$

Die Multiplikation zweier komplexer Zahlen leitet sich hieraus her mit den üblichen Rechengesetzen und der Regel

$$i^2 = -1. \quad (4.5)$$

Wir erhalten also:

$$zw := (x + iy)(u + iv) = xu + ixv + iyu + i^2yv \quad (4.6)$$

$$= (xu - yv) + i(xv + yu). \quad (4.7)$$

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

$$z^0 := 1, \quad z^n := zz^{n-1}, \quad n \in \mathbb{N}. \quad (4.8)$$

- Sei  $w \neq 0$ . Dann ist in  $\mathbb{C}$  die **Division** durch  $w$  möglich:

$$\frac{z}{w} = \frac{x + iy}{u + iv} := \frac{x + iy}{u + iv} \frac{u - iv}{u - iv} \quad (4.9)$$

$$= \frac{(x + iy)(u - iv)}{(u + iv)(u - iv)} \quad (4.10)$$

$$= \frac{(xu + yv) + i(xv - yu)}{u^2 + v^2} \quad (4.11)$$

$$= \frac{xu + yv}{u^2 + v^2} + i \frac{yu - xv}{u^2 + v^2}. \quad (4.12)$$

In  $\mathbb{C}$  gelten die Körperaxiome, also die üblichen Rechenregeln für  $+, -, \cdot, \div$ . Wegen  $i^2 = -1$  gibt es jedoch keine Anordnung  $<$  auf  $\mathbb{C}$ , die das Anordnungsaxiom erfüllt.

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

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

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

$$\begin{aligned} \overline{z + w} &= \bar{z} + \bar{w}, \\ \overline{zw} &= \bar{z}\bar{w}, \\ \overline{\left(\frac{z}{w}\right)} &= \frac{\bar{z}}{\bar{w}}, \quad \text{falls } w \neq 0, \end{aligned} \tag{4.13}$$

$$\begin{aligned} \overline{\bar{z}} &= z, \\ \operatorname{Re} z &= \frac{1}{2}(z + \bar{z}), \\ \operatorname{Im} z &= \frac{1}{2i}(z - \bar{z}). \end{aligned} \tag{4.14}$$

- Der **Betrag**  $|z|$  von  $z = x + iy$  ist die Norm als Vektor, also:

$$|z| := \sqrt{x^2 + y^2}. \tag{4.15}$$

Es gelten die Rechenregeln:

$$\begin{aligned} |z| &= \sqrt{z\bar{z}}, \\ |zw| &= |z||w|, \\ \left|\frac{z}{w}\right| &= \frac{|z|}{|w|}, \quad \text{falls } w \neq 0, \\ |z| &= |\bar{z}|. \end{aligned} \tag{4.16}$$

**1.4.3.** Es gilt die Dreiecksungleichung

$$|z + w| \leq |z| + |w|. \tag{4.17}$$

*Beweis.* Sei  $\alpha = x + iy$  eine komplexe Zahl. Dann gilt

$$|\alpha| = \sqrt{x^2 + y^2} \geq \sqrt{x^2} = |x| \geq x = \operatorname{Re} \alpha.$$

Wir schreiben dies um als

$$\sqrt{\alpha\bar{\alpha}} \geq \frac{1}{2}(\alpha + \bar{\alpha}).$$

Wir setzen  $\alpha = z\bar{w}$  ein:

$$2\sqrt{z\bar{z}w\bar{w}} \geq z\bar{w} + \bar{z}w \Rightarrow z\bar{z} + 2\sqrt{z\bar{z}w\bar{w}} + w\bar{w} \geq z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \Leftrightarrow$$

$$(\sqrt{z\bar{z}} + \sqrt{w\bar{w}})^2 \geq (z + w)(\bar{z} + \bar{w})$$

Durch Ziehen der Quadratwurzel erhalten wir also

$$|z| + |w| \geq |z + w|$$

□

Sie kann durch Induktion auf  $n$  Summanden verallgemeinert werden

$$\left| \sum_{i=1}^n z_i \right| \leq \sum_{i=1}^n |z_i|. \quad (4.18)$$

Nach Konstruktion hat  $-1$  eine Quadratwurzel in  $\mathbb{C}$ .

**1.4.4 Satz** Quadratische Gleichungen. *In  $\mathbb{C}$  ist jede quadratische Gleichung*

$$ax^2 + bx + c = 0 \quad a, b, c \in \mathbb{C}$$

*lösbar.*

*Beweis.* Wegen der bekannten Lösungsformel

$$x_{1,2} = -\frac{b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac}, \quad (4.19)$$

genügt es, die Existenz von Quadratwurzeln, d.h. Lösungen von

$$x^2 = d$$

herzuleiten.

Für  $d = 0$  ist die Behauptung klar. Sei zunächst  $d > 0$  reell. Wir betrachten

$$s = \sup S, \quad S = \{x \in \mathbb{R}; x^2 \leq d\}.$$

Das Supremum existiert nach dem Vollständigkeitsaxiom, da  $0^2 \leq d$ , d.h. die Menge ist nicht leer und  $n^2 > d$  für eine natürliche Zahl  $n > d$  (d.h.

die Menge ist nach oben beschränkt). Diese existiert wiederum nach dem archimedischen Axiom. Wir wollen nun zeigen:

$$s^2 = d$$

Angenommen,  $s^2 > d$ . Sei

$$\varepsilon = \frac{s^2 - d}{2s} > 0 .$$

Dann ist

$$(s - \varepsilon)^2 = s^2 - 2s\varepsilon + \varepsilon^2 = s^2 - s^2 + d + \varepsilon^2 > d .$$

Damit ist  $s - \varepsilon$  ebenfalls eine obere Schranke von  $S$ , ein Widerspruch dazu, dass  $s$  die kleinste obere Schranke ist.

Angenommen,  $s^2 < d$ . Wir betrachten  $\delta = d - s^2 > 0$ . Wir wählen  $\varepsilon > 0$  mit

$$\varepsilon < 2s, \quad \varepsilon < \frac{\delta}{4s}$$

Dies ist möglich, denn nach dem archimedischen Axiom gibt es eine natürliche Zahl  $n > 1/2s$  und gleichzeitig  $n > 4s/\delta$ . Dann erfüllt  $\varepsilon = 1/n$  die beiden Bedingungen.

Damit folgt

$$(s + \varepsilon)^2 = s^2 + 2s\varepsilon + \varepsilon^2 < s^2 + 2s\varepsilon + 2s\varepsilon < s^2 + 4s\frac{\delta}{4s} = s^2 + d - s^2 = d$$

Also ist  $s + \varepsilon \in S$ . Dies ist ein Widerspruch dazu, dass  $s$  eine obere Schranke von  $S$  ist.

Es bleibt nur die Möglichkeit  $s^2 = d$ , d.h.  $s = \sqrt{d}$ . Ist  $d < 0$  reell, so ist  $i\sqrt{-d}$  die gesuchte Quadratwurzel.

Wir gehen zu komplexem  $d = u + iv \in \mathbb{C} \setminus \mathbb{R}$  über. Gesucht ist eine Lösung von

$$u + iv = (x + iy)^2 = x^2 - y^2 + 2ixy \Leftrightarrow u = x^2 - y^2, v = 2xy .$$

Wegen  $d \notin \mathbb{R}$  ist  $v \neq 0$ .

Wir betrachten

$$h = \frac{u}{2} + \sqrt{\frac{u^2 + v^2}{4}} .$$

Die Zahl unter der Wurzel ist positiv, die Wurzel also reell. Damit ist  $h \in \mathbb{R}$ . Angenommen,  $h = 0$ , d.h.

$$\frac{u}{2} = \sqrt{\frac{u^2 + v^2}{4}} \Rightarrow \frac{u^2}{4} = \frac{u^2 + v^2}{4} \Rightarrow v = 0$$

Die ist ein Widerspruch zur Voraussetzung. Daher dürfen wir setzen

$$x = \sqrt{h}, \quad y = \frac{v}{2x}$$

Wir überprüfen:

$$\begin{aligned} h(x^2 - y^2) &= hx^2 - \frac{v^2}{4} = h^2 - \frac{v^2}{4} \\ &= \frac{u^2}{4} + 2 \frac{u}{2} \sqrt{\frac{u^2 + v^2}{4}} + \frac{u^2 + v^2}{4} - \frac{v^2}{4} \\ &= \frac{u^2}{2} + u \sqrt{\frac{u^2 + v^2}{4}} \\ &= uh \end{aligned}$$

und daher

$$x^2 - y^2 = u, \quad 2xy = v .$$

□

**1.4.5 Bemerkung.** Beim Suchen nach einem Beweis geht man natürlich genau umgekehrt vor: Zuerst die Gleichungen lösen und dabei feststellen, welche Bedingungen nötig sind.

# Kapitel 2

## Funktionen, Folgen, Reihen und Grenzwerte

In diesem Kapitel werden grundlegende Begriffe wie Funktionen und Grenzwerte eingeführt. Unter anderem werden Standardbeispiele für Funktionen, wie Polynome und Kreisfunktionen 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**

$$f : D \rightarrow \mathbb{R} : x \mapsto f(x) \text{ für } D \subseteq \mathbb{R} . \quad (1.1)$$

Meist wird  $D$  ein Intervall sein.

**Beispiele:**

1 Eine **lineare Funktion** ist gegeben durch

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

für festes  $a, b \in \mathbb{R}$  mit dem Definitionsbereich  $D(f) = \mathbb{R}$ .

2 Eine **quadratische Funktion** ist definiert als

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

für festes  $a, b, c \in \mathbb{R}$ , 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  $D(f) = \mathbb{R}_0^+ = \{x \in \mathbb{R}, x \geq 0\}$ . Dies ist die Umkehrfunktion  $f^{-1}$  der bijektiven Funktion  $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  mit  $f(x) = x^2$ .

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

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

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

**2.1.3** Rechnen mit Funktionen.

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

$$\begin{aligned} (f \pm g)(x) &:= f(x) \pm g(x) \\ (f \cdot g)(x) &:= f(x) \cdot g(x) \\ \left(\frac{f}{g}\right)(x) &:= \frac{f(x)}{g(x)}, \quad \text{falls } g(x) \neq 0 \end{aligned} \tag{1.2}$$

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:

$$(f \circ g)(x) := f(g(x)). \tag{1.3}$$

## 2.2 Polynome und rationale Funktionen

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

$$f(x) = a_0 + a_1x + \dots + a_nx^n = \sum_{i=0}^n a_i x^i. \quad (2.1)$$

- Die  $a_i$  heißen **Koeffizienten** des Polynoms.
- $f(x) = 0$  hat keinen Grad, ist aber in der Sprechweise: „Polynome von Grad  $n$ “ eingeschlossen.

**2.2.2 Horner Schema.** Sei  $x_0$  konkreter Wert. Um den Funktionswert  $f(x_0)$  in Darstellung (2.1) zu berechnen, braucht man  $2n-1$  Multiplikationen und  $n$  Additionen. Es gibt andere Darstellungen als (2.1), z. B. die **Horner-Darstellung**:

$$f(x) = ((\dots((a_nx + a_{n-1})x + a_{n-2})x + \dots + a_1)x + a_0. \quad (2.2)$$

Hier beträgt der Aufwand  $n$  Multiplikationen und  $n$  Additionen.

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

$$\begin{aligned} f(b) &= c_0 \\ f(x) &= (x - b) \sum_{i=1}^n c_i x^{i-1} + c_0, \end{aligned} \quad (2.3)$$

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

**2.2.4 Definition.** Als **Nullstelle** einer Funktion  $f : D \rightarrow \mathbb{R}$  bezeichnet man jede Lösung der Gleichung

$$f(x) = 0 \quad \text{in } D.$$

- Für Polynome erhält man aus (2.3):

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

d.h.  $f(x)$  ist teilbar durch 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:  $f(x) = (x - b)^2 h_1(x)$ .

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

$$f(x) = (x - b)^k g(x) \quad \text{und} \quad g(b) \neq 0. \quad (2.4)$$

**2.2.6 Satz.**

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

$$f(x) = \prod_{i=1}^r (x - b_i)^{l_i} q(x) \quad (2.5)$$

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

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

*Beweis.* Vollständige Induktion nach dem Grad des Polynoms. Der Induktionsschritt ist die Bemerkung nach Definition 2.2.4.  $\square$

**2.2.7 Komplexe Polynome.** Eine Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  heißt komplexes Polynom vom Grad  $n$ , wenn es Zahlen  $a_0, \dots, a_n \in \mathbb{C}$  gibt mit  $a_n \neq 0$ , so dass

$$f(z) = \sum_{i=0}^n a_i z^i \quad (2.6)$$

Auch  $f(z) = 0$  wird als Polynom aufgefasst.

- 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 komplexe Zahlen war das Polynom  $x^2 + 1$ , welches in  $\mathbb{R}$  keine Nullstelle hat. Wir haben gesehen, dass über  $\mathbb{C}$  jede quadratische Funktion eine Nullstelle hat (Satz 1.4.4). Dies gilt sogar allgemein, wie der folgende Satz zeigt.

**2.2.8 Theorem** (Fundamentalsatz der Algebra). *Zu jedem Polynom der Form (2.6), mit  $n \geq 1$  und  $a_n \neq 0$ , gibt es eine Zahl  $w \in \mathbb{C}$  mit  $f(w) = 0$ .*

Der Beweis ist alles andere als einfach.

Die einzigen komplexen Polynome ohne komplexe Nullstelle sind also die konstanten Polynome. Wie im reellen Fall erhält man also:

**2.2.9 Satz.** *Jedes komplexe Polynom der Form (2.6) mit  $n \geq 1, a_n \neq 0$  besitzt eine **Faktorisierung über  $\mathbb{C}$**  der Form*

$$f(z) = a_n(z - w_1)^{l_1} \cdots (z - w_k)^{l_k} \quad (2.7)$$

mit verschiedenen Nullstellen  $w_i \in \mathbb{C}$  der Vielfachheit  $l_i$  mit  $\sum_{i=1}^k l_i = n$ .

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

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

*Beweis.* Sei

$$f(z) = \sum_{i=0}^n a_i z^i .$$

Nach Voraussetzung ist  $a_i \in \mathbb{R}$ . Sei  $w \in \mathbb{C}$  eine Nullstelle von  $f$ . Dann gilt

$$f(\bar{w}) = \sum_{i=0}^n a_i \bar{w}^i = \overline{\sum_{i=0}^n a_i w^i}$$

denn komplexe Konjugation vertauscht mit Multiplikation und Addition, und es gilt  $a_i = \bar{a}_i$ . Es gilt also

$$f(\bar{w}) = \overline{f(w)} = \bar{0} = 0 .$$

□

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

$$f(x) = a_n(x - b_1)^{l_1} \cdots (x - b_r)^{l_r} (x^2 + c_1 x + d_1)^{k_1} \cdots (x^2 + c_s x + d_s)^{k_s} ,$$

mit reellen Nullstellen  $b_i \in \mathbb{R}$  der Vielfachheit  $l_i$  und quadratischen Polynomen  $x^2 + c_i x + d_i$ , die keine reelle Nullstelle haben.

*Beweis.* Sei

$$f(z) = \sum_{i=0}^n a_i z^i$$

mit  $a_n \neq 0$ . Nach dem Lemma treten die Nullstellen von  $f$  paarweise auf. Ist eine Nullstelle  $w \in \mathbb{C} \setminus \mathbb{R}$ , so enthält  $f$  den Faktor

$$(z - w)(z - \bar{w}) = z^2 - (w + \bar{w})z + w\bar{w} = z^2 - 2\operatorname{Re}(w)z + |w|^2 .$$

Dies ist ein reelles Polynom.

Wir benutzen die Faktorsierung über den komplexen Zahlen

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

wobei die  $w_i$  die verschiedenen komplexen Nullstellen von  $f$  sind. Wir sortieren die  $w_i$  so, dass  $w_1, \dots, w_r$  reell sind. Dies sind dann genau die  $b_i$  aus der Formulierung des Satzes.

Die Nullstellen  $w_{r+1}, \dots, w_k$  sind nicht reell. Seien  $w_{r+1}, \dots, w_{r+s}$  paarweise nicht komplex konjugiert,  $w_{r+s+i} = \bar{w}_{r+i}$  die restlichen. (Es gilt also  $r + 2s = k$ .) Dann ist

$$z^2 + c_i z + d_i = (z - w_{r+i})(z - w_{r+s+i}) .$$

□

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

*Beweis.* Eindeutigkeit: Seien  $p_n, p'_n$  zwei solche Polynome. Dann gilt

$$(p_n - p'_n)(x_i) = y_i - y_i = 0 \quad i = 0, \dots, n$$

d.h.  $p_n - p'_n$  ist ein Polynom vom Grad kleiner gleich  $n$  mit mindestens  $n + 1$  Nullstellen. Nach Satz 2.2.6 folgt hieraus, dass  $p_n - p'_n = 0$ . Der Existenzbeweis bedeutet das Lösen eines linearen Gleichungssystems. Es lohnt sich, das Verfahren festzuhalten, s.u. □

**2.2.13. Newton Interpolationsverfahren.** Man suche das Polynom  $p_n$  aus Satz 2.2.12 in der Form

$$\begin{aligned} p_n(x) = & \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)(x - x_1) + \cdots + \\ & + \alpha_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}) \end{aligned} \quad (2.8)$$

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

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

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

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

Hierbei ist

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

Man kann nachrechnen, dass

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

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

**2.2.14 Folgerung.** *Zwei Polynome sind genau dann gleich, wenn alle ihre Koeffizienten übereinstimmen, d.h.*

$$\sum_{i=0}^n a_i x^i = \sum_{i=0}^n b_i x^i \quad \Leftrightarrow \quad a_i = b_i \quad i = 1, \dots, n.$$

*Beweis.* Seien  $p = \sum_{i=0}^n a_i x^i$  und  $q = \sum_{j=0}^n b_j x^j$  Polynome vom Grad kleiner gleich  $n$  mit  $p = q$  als Funktionen auf  $\mathbb{R}$ . Dann stimmen sie in  $n + 1$  verschiedenen Stützpunkten überein. Nach Satz 2.2.12 sind sie dadurch eindeutig bestimmt.  $\square$

**2.2.15. Polynomdivision.** *Seien  $p$  und  $q$  Polynome,  $q \neq 0$ . Dann gibt es eindeutig bestimmte Polynome  $h$  und  $r$  mit  $\text{Grad}(r) < \text{Grad } q$  so dass*

$$p = hq + r .$$

Der mathematische Beweis wird für festes  $q$  mit vollständiger Induktion nach dem Grad von  $p$  geführt. Wichtiger ist der Algorithmus, der dem der schriftlichen Division von Zahlen (mit Rest) entspricht.

**2.2.16 Beispiel.**

$$\begin{aligned} x^4 - x^3 + x^2 - x + 1 & : x^2 + 2x = x^2 & \text{Rest } -3x^3 + x^2 - x + 1 \\ -3x^3 + x^2 - x + 1 & : x^2 + 2x = -3x & \text{Rest } 8x^2 - x + 1 \\ 8x^2 - x + 1 & : x^2 + 2x = 8 & \text{Rest } -17x + 1 \end{aligned}$$

und daher

$$h = x^2 - 3x + 8, r = -17x + 1 .$$

**2.2.17 Definition.** *Der Quotient zweier Polynome*

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

heißt **rationale Funktion**.

Haben Zähler und Nenner eine Nullstelle gemeinsam, so kann man einen Linearfaktor kürzen, ohne die Funktion zu verändern.

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

$$D(f) = \{x \in \mathbb{R} : q(x) \neq 0\}. \quad (2.9)$$

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

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

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

**2.2.18 Satz.** Jede rationale Funktion  $\frac{p(x)}{q(x)}$  lässt sich darstellen als

$$\frac{p(x)}{q(x)} = h(x) + \frac{r(x)}{q(x)} \quad (2.10)$$

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

*Beweis.* Gesucht sind  $h$  und  $r$ , so dass

$$p = hq + r ,$$

mit  $\text{Grad } r < \text{Grad } q$ , d.h. die Aussage ist äquivalent zur Polynomdivision mit Rest.  $\square$

## 2.3 Trigonometrische Funktionen

Wir erinnern uns, dass wir alle Winkel im Bogenmaß angeben, d.h. der Vollkreis hat Maß  $2\pi$ .

**2.3.1 Sinus, Cosinus.** Auf dem Einheitskreis drehe man einen Zeiger um den Winkel  $\alpha$ . Dadurch zeigt der Zeiger auf den Punkt  $P$ . Die Koordinaten von  $P$  werden mit  $\cos \alpha$  und  $\sin \alpha$  bezeichnet.

$$P = (\cos \alpha, \sin \alpha)$$

Da der Winkel  $\alpha \in \mathbb{R}$  beliebig ist erhalten wir die Funktionen:

$$\begin{array}{ll} \cos : \mathbb{R} \rightarrow [-1, 1] : \alpha \mapsto \cos \alpha & \text{Cosinusfunktion ,} \\ \sin : \mathbb{R} \rightarrow [-1, 1] : \alpha \mapsto \sin \alpha & \text{Sinusfunktion .} \end{array}$$

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

$$Q = (r \cos \alpha, r \sin \alpha) . \quad (3.1)$$

Diese Gleichungen kann man auch schreiben als

$$\begin{array}{l} \cos \alpha = \frac{a}{r} , \\ \sin \alpha = \frac{b}{r} . \end{array} \quad (3.2)$$

Aus der Definition erhält man sofort folgende Eigenschaften:

$$\begin{aligned} -1 &\leq \cos x \leq 1, \\ \cos(-x) &= \cos x && \text{(gerade Funktion),} \\ \cos(x + 2k\pi) &= \cos x && \text{(} 2\pi\text{-periodisch),} \end{aligned} \quad (3.3)$$

$$\begin{aligned} -1 &\leq \sin x \leq 1, \\ \sin(-x) &= -\sin x && \text{(ungerade Funktion),} \\ \sin(x + 2k\pi) &= \sin x && \text{(} 2\pi\text{-periodisch),} \end{aligned} \quad (3.4)$$

$$\begin{aligned} \cos x = 0 &\Leftrightarrow \pm\frac{1}{2}\pi, \pm\frac{3}{2}\pi, \pm\frac{5}{2}\pi, \dots, \\ \sin x = 0 &\Leftrightarrow 0, \pm\pi, \pm2\pi, \dots \end{aligned} \quad (3.5)$$

$$\cos^2(x) + \sin^2(x) = 1 \quad (3.6)$$

wobei  $\sin^2(x) = (\sin(x))^2$ .

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

$$\begin{aligned} \cos(x + y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x + y) &= \sin x \cos y + \cos x \sin y. \end{aligned} \quad (3.7)$$

*Beweis.* Wir setzen die definierenden rechtwinkligen Dreiecke aufeinander. Das erste hat die Ecken  $(0, 0)$ ,  $(\cos x, 0)$  und  $(\cos x, \sin x)$ . Das zweite hat als erste Ecke  $(0, 0)$ , als zweite einen Punkt  $P$  auf der Nullpunktgeraden  $G$  durch  $(\cos x, \sin x)$  mit Abstand  $\cos y$  vom Nullpunkt. Die Gerade hat die Gleichung

$$t \mapsto \frac{\sin x}{\cos x} t.$$

Es gilt also

$$P = (\cos y \cos x, \cos y \sin x).$$

Wir bestimmen die Gerade  $G'$ , die senkrecht auf  $G$  steht und durch  $P$  geht. Sie hat die Steigung  $-\frac{\cos x}{\sin x}$  und die Gleichung

$$t \mapsto -\frac{\cos x}{\sin x} t + \left( \cos y \sin x + \frac{\cos^2 x \cos y}{\sin x} \right).$$

Die dritte Ecke  $Q$  liegt auf  $G'$  und hat von  $P$  den Abstand  $\sin(y)$ . Die Behauptung ist  $Q = Q'$  mit

$$Q' = (\cos x \cos y - \sin x \sin y, \sin x \cos y + \cos x \sin y).$$

Einsetzen zeigt, dass der Punkt auf  $G'$  liegt. Es gilt

$$\overrightarrow{PQ'} = Q' - P = (-\sin y \sin x, \cos x \sin y).$$

Dieser Vektor hat die Länge  $\sin y$  wie gewünscht.  $\square$

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

$$\begin{aligned}\sin\left(x + \frac{\pi}{2}\right) &= \cos x, \\ \cos\left(x - \frac{\pi}{2}\right) &= \sin x,\end{aligned}\tag{3.8}$$

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 aus (3.5) herleiten lassen. Hier einige Beispiele:

$$\begin{aligned}\cos(x - y) &= \cos x \cos y + \sin x \sin y, \\ \sin(x - y) &= \sin x \cos y - \cos x \sin y,\end{aligned}\tag{3.9}$$

$$\begin{aligned}\sin x + \sin y &= 2 \sin \frac{x+y}{2} \cos \frac{x-y}{2}, \\ \cos x + \cos y &= 2 \cos \frac{x+y}{2} \cos \frac{x-y}{2},\end{aligned}\tag{3.10}$$

$$\begin{aligned}\cos(2x) &= \cos^2 x - \sin^2 x = 2 \cos^2 x - 1, \\ \sin 2x &= 2 \sin x \cos x,\end{aligned}\tag{3.11}$$

$$\begin{aligned}1 + \cos x &= 2 \cos^2 \frac{x}{2}, \\ 1 - \cos x &= 2 \sin^2 \frac{x}{2},\end{aligned}\tag{3.12}$$

**2.3.3 Polardarstellung komplexer Zahlen.** *Man kann jede komplexe Zahl  $z = x + iy$  auch als „Zeiger“ auffassen und deshalb ist  $z$  eindeutig bestimmt durch einen Winkel  $\varphi$  und den Betrag  $r = |z|$ .*

- Es gilt dann  $x = r \cos \varphi$ ,  $y = r \sin \varphi$  und somit erhalten wir

$$z = r(\cos \varphi + i \sin \varphi).\tag{3.13}$$

- Diese Form der Darstellung von  $z$  heißt **Polardarstellung**. Man nennt  $\varphi$  das **Argument** von  $z$ ,  $\varphi =: \arg z$ . Falls  $-\pi < \varphi \leq \pi$  heißt  $\varphi$  **Hauptwert von  $\arg z$**

- Die folgende abkürzende Schreibweise ist sinnvoll

$$e^{i\varphi} := \cos \varphi + i \sin \varphi, \quad \varphi \in \mathbb{R}. \quad (3.14)$$

Sie heißt **Euler-Formel**.

Die Begründung für diese Notation liefern die Rechenregeln. Rechnung mit Sinus und Cosinus vereinfachen sich in der Notation der komplexen Zahlen.

**2.3.4 Satz** (De Moivre). *Für alle  $\varphi, \psi \in \mathbb{R}$  gilt,*

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

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

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

*Beweis.* Wir betrachten die erste Aussage:

$$\begin{aligned} e^{i\varphi} e^{i\psi} &= (\sin \varphi + i \cos \varphi)(\sin \psi + i \cos \psi) \\ &= \sin \varphi \sin \psi - \cos \varphi \cos \psi + i(\cos \varphi \sin \psi + \sin \varphi \cos \psi) \end{aligned}$$

Nach der Additionsformel 2.3.2 ist dies

$$\cos(\varphi + \psi) + i \sin(\varphi + \psi) .$$

□

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

$$\begin{aligned} z \cdot w &= |z||w| e^{i(\varphi+\psi)}, \\ \frac{z}{w} &= \frac{|z|}{|w|} e^{i(\varphi-\psi)}, \quad \text{falls } w \neq 0. \end{aligned} \quad (3.15)$$

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

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

- Wir erhalten sofort folgende Eigenschaften:

$$\begin{aligned}\tan(-x) &= -\tan x, & (\text{ungerade Funktion}), \\ \cot(-x) &= -\cot x, & (\text{ungerade Funktion}),\end{aligned}\quad (3.16)$$

$$\begin{aligned}\tan(x + \pi) &= \tan x, & (\pi\text{-periodisch}), \\ \cot(x + \pi) &= \cot x, & (\pi\text{-periodisch}),\end{aligned}\quad (3.17)$$

$$\tan(x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}, \quad \text{für „erlaubte“ } x, y. \quad (3.18)$$

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

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

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

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

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

- Die **Überlagerung (Superposition)**  $s(t)$  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

$$\begin{aligned} z(t) &:= A \cos(\omega t + \alpha) + iA \sin(\omega t + \alpha) \\ &= Ae^{i(\omega t + \alpha)} \\ &= ae^{i\omega t}. \end{aligned} \tag{3.19}$$

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

Wir überlagern zwei harmonische Schwingungen mit gleicher Kreisfrequenz, d.h.

$$\begin{aligned} s_1(t) &= A_1 \cos(\omega t + \alpha_1), & s_2(t) &= A_2 \cos(\omega t + \alpha_2), \\ s(t) &= s_1(t) + s_2(t). \end{aligned}$$

Dies ist der Realteil von

$$\begin{aligned} A_1 e^{i(\omega t + \alpha_1)} + A_2 e^{i\omega t + \alpha_2} &= A_1 e^{i\alpha_1} e^{i\omega t} + A_2 e^{i\alpha_2} e^{i\omega t} \\ &= (A_1 e^{i\alpha_1} + A_2 e^{i\alpha_2}) e^{i\omega t}. \end{aligned}$$

Der erste Faktor ist eine komplexe Zahl, also von der Form  $u + iv = Ae^{i\alpha}$  mit

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

und dann  $A = \sqrt{u^2 + v^2}$ ,  $\cos \alpha = \frac{u}{A}$ ,  $\sin \alpha = \frac{v}{A}$ . Hieraus ergibt sich

$$A_1 e^{i(\omega t + \alpha_1)} + A_2 e^{i\omega t + \alpha_2} = Ae^{i(\omega t + \alpha)}.$$

Damit haben wir hergeleitet:

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

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

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

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

*mit  $A$  und  $\alpha$  wie oben.*

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

$$a_1 e^{i\omega_1 t} + a_2 e^{i\omega_2 t} = \left( a_1 e^{i\frac{\omega_1 - \omega_2}{2} t} + a_2 e^{i\frac{\omega_2 - \omega_1}{2} t} \right) e^{i\frac{\omega_1 + \omega_2}{2} t} \quad (3.20)$$

darstellen. Man nennt dies **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 doch 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

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

**2.4.1 Definition.** *Unter einer Folge reeller Zahlen versteht man eine auf  $\mathbb{N}_0$  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**

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

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

wenn es zu jeder 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 von  $\varepsilon$  ab!) liegen in einer  $\varepsilon$ -Umgebung von  $a$ .

- Falls der Grenzwert existiert heißt die Folge **konvergent**, ansonsten **divergent**.
- Falls  $a = 0$  heißt die Folge **Nullfolge**.

**2.4.3 Definition.** Man sagt, eine Folge  $(a_n)_{n \geq 0}$  **divergiert** gegen  $\infty$ , in Zeichen

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

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

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

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

*Beweis.* Sei  $x > 1$ . Sei  $K \in \mathbb{N}$ . Wir setzen  $x = 1 + y$  mit  $y > 0$ . Sei  $n_0 \geq K/y$  eine natürliche Zahl. Für  $n \geq n_0$  folgt mit der Bernoulli-Ungleichung (Beispiel 1.2.7):

$$(1 + y)^n \geq 1 + ny > \frac{K}{y}y = K.$$

Die Folge divergiert gegen  $\infty$ .

Sei nun  $|x| < 1$ ,  $\varepsilon > 0$ . Dann ist  $|x|^{-1} > 1$ . Nach dem ersten Teil gibt es  $n_0$ , so dass für alle  $n \geq n_0$  gilt

$$(|x|^{-1})^n > \varepsilon^{-1} \Rightarrow |x^n| < \varepsilon .$$

Es handelt sich um eine Nullfolge.

Sei zuletzt  $x < -1$ . Die Folge hat abwechselnd Werte kleiner  $-1$  und größer  $1$ . Sie kann nicht konvergieren.  $\square$

Seien  $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$  Folgen. Dann sind  $a_n \pm b_n, a_n \cdot b_n$ , und  $a_n/b_n$  neue Folgen. Zur Grenzwertberechnung sollte man zuerst überprüfen, ob die zu betrachtende Folge aus einfacheren Folgen zusammengesetzt ist.

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

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

*Beweis.* Alle Beweise gehen ähnlich. Wir zeigen die erste Aussage. Sei  $\lim_{n \rightarrow \infty} a_n = a$ ,  $\lim_{n \rightarrow \infty} b_n = b$ . Wir wollen zeigen, dass  $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b$ . Sei  $\varepsilon > 0$ . Dann gibt es  $n_0$ , so dass

$$|a_n - a| < \frac{\varepsilon}{2}, \quad |b_n - b| < \frac{\varepsilon}{2} \quad \text{für alle } n \geq n_0 .$$

Dann gilt

$$|a_n + b_n - a - b| \leq |a - a_n| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \quad \text{für alle } n \geq n_0 .$$

$\square$

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

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

*Beweis.* (überspringen) Sei  $\varepsilon > 0$ . Für  $n \geq n_0$  gilt

$$|c - b_n| < \varepsilon, \quad |c - a_n| < \varepsilon .$$

Es folgt

$$|a_n - b_n| = |a_n - c + c - b_n| \leq |a_n - c| + |c - b_n| < 2\varepsilon .$$

Für  $n \geq \max(n_0, n_1)$  folgt  $|a_n - c_n| < 2\varepsilon$  und damit

$$|c - c_n| = |c - b_n + b_n - a_n + a_n - c_n| \leq |c - b_n| + |b_n - a_n| + |a_n - c_n| < 5\varepsilon .$$

□

**2.4.8 Definition.** Eine Folge heißt **beschränkt**, wenn es Konstanten  $K_1, K_2$  gibt, so dass

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

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

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

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

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

*Beweis.* Sei  $\varepsilon > 0$ . Nach Voraussetzung gibt es  $n_0$ , so dass

$$|a - a_n| < \varepsilon, \quad |b - a_n| < \varepsilon \quad \text{für alle } n \geq n_0 .$$

Hieraus folgt mit der Dreiecksungleichung

$$|a - b| = |a - a_n + a_n - b| < 2\varepsilon .$$

Da  $\varepsilon$  beliebig klein werden kann, folgt  $|a - b| = 0$  d.h.  $a = b$ .

Für die zweite Aussage wählen wir wieder  $\varepsilon > 0$ . Dann gibt es  $n_0$  so dass  $|a - a_n| < \varepsilon$  für alle  $n \geq n_0$ . D.h. es gilt

$$a_n \in (a - \varepsilon, a + \varepsilon) \quad \text{für alle } n \geq n_0 .$$

Also gilt

$$\min(a_0, \dots, a_{n_0-1}, a - \varepsilon) \leq a_n \leq \max(a_0, \dots, a_{n_0-1}, a + \varepsilon) \quad \text{für alle } n \geq 0 .$$

□

- 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.  $a_{n+1} \leq a_n$ ).

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

*Beweis.* Sei  $(a_n)_{n \geq 0}$  nach oben beschränkt und monoton wachsend. Wir setzen  $a = \sup(\{a_n | n \geq 0\})$ . Wir zeigen

$$a = \lim_{n \rightarrow \infty} a_n .$$

Es existiert nach dem Vollständigkeitsaxiom. Sei  $\varepsilon > 0$ . Nach Definition des Supremums gibt es ein  $n_0$  mit  $a_{n_0} \in (a - \varepsilon, a]$ . Für jedes  $n \geq n_0$  ist  $a_n \geq a_{n_0}$ , da die Folge monoton wächst. Da  $a$  obere Schranke ist, folgt  $a_n \leq a$ . Insgesamt also  $a_n \in (a - \varepsilon, a]$ .  $\square$

**2.4.11 Beispiel.** Sei  $0 < x < 1$ . Wir betrachten die Folge  $a_n = x^n$  für  $n \geq 0$ . Die Folge ist monoton fallend, hat also einen Grenzwert  $a$ . Die Folge  $b_n = x^{n+1}$  hat denselben Grenzwert. Es folgt aus den Rechenregeln für Grenzwerte

$$a = \lim_{n \rightarrow \infty} b_n = x \lim_{n \rightarrow \infty} a_n = a^2 \Rightarrow a = 0, 1 .$$

Wegen  $a < a_1 = x < 1$  ist also  $a = 0$ .

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

**2.4.13 Satz.** *Jede beschränkte Folge  $(a_n)_{n \geq 0}$  hat eine konvergente Teilfolge.*

*Beweis.* (überspringen). Sei  $b_n = \sup(a_k | k \geq n)$ . Die Folge ist beschränkt und monoton fallend. Also hat sie einen Grenzwert  $b$ . Wir bestimmen induktiv die Indices der Teilfolge  $n_m$ . Für jedes  $m$  gibt es einen Index  $n_m > n_{m+1}$ , so dass

$$|b - b_{n_m}| < \frac{1}{m} .$$

Dies ist die gesuchte Teilfolge.  $\square$

## 2.5 Unendliche Reihen

**2.5.1 Definition.** Sei  $(a_k)_{k \geq 0}$  eine Folge reeller Zahlen. Durch die Festsetzung

$$S_n := \sum_{k=0}^n a_k \quad (5.1)$$

wird eine Zahlenfolge  $(S_n)_{n \geq 0}$  definiert, die die zu  $(a_k)_{k \geq 0}$  gehörende **unendliche Reihe** heißt und mit

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

bezeichnet wird. Die Zahlen  $a_k$  heißen **Glieder** der Reihe, die Summen  $S_n$  heißen **Partialsommen**. Wenn  $(S_n)_{n \geq 0}$  konvergiert (bzw. divergiert) sagt man, die Reihe **konvergiert** (bzw. **divergiert**). Falls

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k = S$$

mit  $S \in \mathbb{R} \cup \{\pm\infty\}$  nennt man  $S$  die **Summe** der Reihe und schreibt

$$S = \sum_{k=0}^{\infty} a_k. \quad (5.2)$$

Das Symbol  $\sum_{k=0}^{\infty} a_k$  wird sowohl für die Folge der Partialsommen als auch für ihren Grenzwert benutzt (wenn letzterer existiert).

**2.5.2 Beispiel** Geometrische Reihe. Sei  $x \in \mathbb{R}$ . Die geometrische Reihe ist die Reihe

$$\sum_{i=0}^{\infty} x^i.$$

Ihre Partialsommen sind die endlichen geometrische Reihen:

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

Es können folgende Fälle auftreten:

$$\sum_{n=0}^{\infty} x^n := \lim_{N \rightarrow \infty} \sum_{n=0}^N x^n = \begin{cases} \frac{1}{1-x}, & \text{falls } |x| < 1, \\ \infty, & \text{falls } x \geq 1, \\ \text{divergiert,} & \text{falls } x < -1. \end{cases} \quad (5.3)$$

**2.5.3 Satz** (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  $N_\varepsilon$  gibt, so dass für alle  $n, m \geq N_\varepsilon$*

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

*gilt.*

*Beweis.* (überspringen) Sei  $S_n \rightarrow S$ , d.h. für alle  $\varepsilon > 0$  gibt es  $n_0$ , so dass für alle  $n \geq n_0$ :

$$|S_n - S| \leq \frac{\varepsilon}{2} .$$

Für  $m, n \geq n_\varepsilon$  folgt

$$|S_n - S_m| \leq |S_n - S| + |S - S_m| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

Sei umgekehrt das Cauchy-Kriterium erfüllt. Sei  $\varepsilon > 0$ . Dann gibt es  $n_0$ , so dass  $|S_n - S_m| < \varepsilon$  für alle  $n, m \geq n_0$ . Insbesondere ist  $S_n \in (S_{n_0} - \varepsilon, S_{n_0} + \varepsilon)$  für alle  $n \geq n_0$ . Daher ist die Folge beschränkt. Nach Satz 2.4.13 gibt es eine konvergente Teilfolge  $(S_{n_k})_{k \geq 0}$ . Sei  $S$  deren Grenzwert. Dann gibt es  $k_0$ , so dass  $|S_{n_k} - S| < \varepsilon$  für alle  $k \geq k_0$ . Ohne Einschränkung ist  $n_{k_0} \geq n_0$ . Sei  $n \geq n_{k_0}$ . Dann folgt

$$|S - S_n| \leq |S - S_{n_{k_0}}| + |S_{n_{k_0}} - S_n| < \varepsilon + \varepsilon .$$

Daher ist  $S_n \rightarrow S$ . □

**2.5.4 Satz.** *Die Glieder einer konvergenten Reihe bilden eine Nullfolge.*

*Beweis.* Wähle  $m = n + 1$  im Cauchy-Kriterium. Oder explizit: Sei

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

Wir setzen  $S_n = \sum_{i=0}^n a_i$ . Sei  $\varepsilon > 0$ . Dann gibt es  $n_0$ , so dass

$$|S - S_n| < \varepsilon \quad \text{für alle } n \geq n_0 .$$

Es folgt

$$|a_n| = |S_n - S_{n-1}| < |S_n - S| + |S - S_{n-1}| < 2\varepsilon \quad \text{für alle } n \geq n_0 + 1 .$$

□

**2.5.5 Beispiel** Harmonische Reihe. Die harmonische Reihe ist Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

Wir betrachten die Folge  $(\frac{1}{n})_{n \geq 1}$ . Sie divergiert gegen  $\infty$ .

*Beweis.* Sei  $K \in \mathbb{N}_0$ ,  $m = 2K + 2$ . Für jedes  $n \geq 2^m$  gilt dann:

$$\begin{aligned} s_n &\geq s_{2^m} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{m-1}+1} + \dots + \frac{1}{2^m}\right) \\ &\geq 1 + \frac{1}{2} + 2\frac{1}{4} + 4\frac{1}{8} + \dots + 2^{m-1}\frac{1}{2^m} \\ &= 1 + m\frac{1}{2} = 1 + \frac{2K-1}{2} > K \end{aligned}$$

□

**2.5.6 Definition.** Eine Reihe heißt **alternierend**, wenn aufeinander folgende Reihenglieder unterschiedliche Vorzeichen haben, d.h.  $a_n \cdot a_{n+1} < 0$  für alle  $n \in \mathbb{N}_0$  gilt.

**2.5.7 Satz** (Leibnitz - Kriterium). Für jede monoton fallende Nullfolge  $(a_n)_{n \geq 0}$  konvergiert die alternierende Reihe  $\sum_{k=0}^{\infty} (-1)^k a_k$ .

Beweis:

*Beweis.* (überspringen) Sei  $S_n = \sum_{i=0}^n (-1)^i a_i$  eine Partialsumme. Nach Voraussetzung ist

$$\begin{aligned} S_{2n+2} - S_{2n} &= (-1)^{2n+2}(a_{2n+2} - a_{2n+1}) > 0, \\ S_{2n+1} - S_{2n-1} &= (-1)^{2n+1}(a_{2n+1} - a_{2n}) < 0. \end{aligned}$$

Die Folge der  $S_{2n}$  ist streng monoton steigend und nach oben beschränkt durch  $S_1$ . Sei  $S$  der Grenzwert. Die Folge der  $S_{2n+1}$  ist streng monoton fallend und beschränkt durch  $S_0$ . Sei  $S'$  der Grenzwert. Da  $|S_{2n} - S - 2n - 1| = |a_{2n}|$  eine Nullfolge ist, folgt leicht  $S = S'$  und die Reihe konvergiert. □

**2.5.8 Beispiel.**  $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$  ist konvergent.

**2.5.9 Satz.** Sei  $c \in \mathbb{R}$  und seien  $\sum_{k=0}^{\infty} a_k = a$ ,  $\sum_{k=0}^{\infty} b_k = b$  konvergente Reihen. Dann gilt

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

*Beweis.* Satz 2.4.5. □

Elementare Operationen wie Klammern Setzen oder Weglassen und Umordnen der Glieder, die für endliche Reihen erlaubt sind, sind im Allgemeinen für unendliche Reihen nicht harmlos.

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

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

**2.5.11 Definition.** 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**.

**2.5.12 Satz.** Jede absolut konvergente Reihe ist konvergent.

*Beweis.* Dies ist eine Anwendung des Cauchy-Kriterium und der Beobachtung

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

□

**2.5.13 Satz.** Die Reihe  $\sum_{k=0}^{\infty} a_k$  ist genau dann absolut konvergent, wenn die Folge der Partialsummen  $S_n = \sum_{k=0}^n |a_k|$  beschränkt ist.

*Beweis.* Satz 2.4.10, denn die Folge der  $S_n$  ist monoton wachsend. □

**2.5.14 Definition.** 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  $\sum_{k=0}^{\infty} a_k$

- 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  $b_{\sigma^{-1}(k)} = a_k$ .

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

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

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

Wir wenden uns weiteren Konvergenzkriterien zu.

**2.5.17 Satz** (Vergleichskriterium). Besteht für die Reihenglieder die Abschätzung

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

dann gilt:

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

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

*Beweis.* Wir wenden das Cauchy Kriterium an und beachten

$$\left| \sum_{k=n}^m a_k \right| \leq \sum_{k=n}^m |a_k| \leq \sum_{k=n}^m b_k.$$

□

Wir wenden dieses Konvergenzkriterium auf die eine Reihe, deren Konvergenz wir verstanden haben, nämlich die geometrische Reihe:

**2.5.18 Satz** (Quotientenkriterium). Sei  $a_k \neq 0$  für alle  $k \geq k_0$  und sei  $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = a$ . Dann gilt:

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

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

*Beweis.* Sei  $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < 1$ . Dann gibt es  $q \in (0, 1)$  mit  $\left| \frac{a_{k+1}}{a_k} \right| \leq q$  für alle  $k \geq k_1$ .

$$|a_{k_1+l}| \leq q |a_{k_1+l-1}| \leq q^2 |a_{k_1+l-2}| \leq \cdots \leq q^l |a_{k_1}|$$

Wir wenden nun das Vergleichskriterium an auf

$$b_k := |a_{k_1}| q^{k-k_1} .$$

Die Reihe  $\sum_{k=0}^{\infty} b_k$  konvergiert als geometrische Reihe. □

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

**2.5.19 Beispiel.** Wir betrachten die Reihe

$$\sum_{k=0}^n \frac{1}{k!} .$$

Es gilt  $a_{k+1}/a_k = 1/(k+1)$ . Dies ist eine Nullfolge. Nach dem Quotientenkriterium konvergiert sie, und man definiert die **Eulersche Zahl**  $e$  durch

$$e := \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n \rightarrow \infty} \left( \sum_{k=0}^n \frac{1}{k!} \right) . \quad (5.4)$$

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.

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n . \quad (5.5)$$

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

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

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

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

*Beweis.* Ebenfalls Vergleich mit einer geometrischen Reihe. □

## 2.6 Potenzreihen

**2.6.1 Definition.** Sei  $(f_n)_{n \geq 0}$  eine auf dem Intervall  $I \subseteq \mathbb{R}$  definierte Funktionenfolge. Falls für alle  $x \in I$  die Zahlenfolge  $(f_n(x))_{n \geq 0}$  den Grenzwert  $f(x)$  besitzt, sagt man, dass die Funktionenfolge  $(f_n)_{n \geq 0}$  **punktweise gegen  $f$  konvergiert**.

Es gibt also für jedes  $x \in I$  und jedes  $\varepsilon > 0$  ein  $n_0$ , so dass für alle  $n \geq n_0$  gilt  $|f(x) - f_n(x)| < \varepsilon$ . Der Index  $n_0$  hängt also bei punktweiser Konvergenz von  $x$  ab.

**2.6.2 Definition.** Eine Funktionenfolge  $(f_n)_{n \geq 0}$  **konvergiert gleichmäßig** auf  $I$  gegen die Funktion  $f : I \rightarrow \mathbb{R}$ , wenn es für alle  $\varepsilon > 0$  einen Index  $n_0$  gibt, so dass für alle  $x \in I$  und alle  $n \geq n_0$  gilt:

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

Mit diesem Konvergenzbegriff übertragen sich schöne Eigenschaften der Folge wie Stetigkeit auf den Grenzwert.

**2.6.3 Satz.** 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.

*Beweis.* Genau wie Satz 2.5.17, also mit dem Cauchy Kriterium. □

Eine Funktionenreihe  $\sum_{k=0}^{\infty} f_k(x)$  mit den speziellen Funktionen  $f_k(x) = a_k x^k$  heißt **Potenzreihe**.

**2.6.4 Beispiel.** (geometrische Reihe)

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \dots$$

konvergiert nach Beispiel 2.5.2 in  $x \in (-1, 1)$  punktweise gegen  $1/(1-x)$ . Sie konvergiert dort aber nicht gleichmäßig. Wir werden die korrekte Aussage gleich allgemein formulieren.

**2.6.5 Definition.** 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:

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

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

- $R = 0 \Leftrightarrow$  Reihe konvergiert nur für  $x = 0$ .
- Ist  $R > 0$  und  $\varrho \in (0, R)$ , dann konvergiert die Reihe  $\sum_{k=0}^{\infty} a_k x^k$  absolut und gleichmäßig auf  $|x| \leq \varrho$ .
- Für alle  $x$  mit  $|x| > R$  ist die Reihe  $\sum_{k=0}^{\infty} a_k x^k$  divergent.

*Beweis.* Der interessante Teil ist b). Sei  $\sum_{k=0}^{\infty} a_k x^k$  konvergent in einem Punkt  $x_0$  mit  $R \geq |x_0| > \varrho$ . Insbesondere ist  $|a_k x_0^k|$  eine Nullfolge, also beschränkt durch ein  $c > 0$ . Sei nun  $|x| \leq \varrho$ . Für jedes  $k \in \mathbb{N}_0$  folgt

$$|a_k x^k| = |a_k| |x_0|^k \left| \frac{x}{x_0} \right|^k \leq c q^k$$

mit  $q := \left| \frac{x}{x_0} \right|^k < 1$ . Nach Kriterium Satz 2.6.3 konvergiert die Reihe absolut und gleichmäßig auf  $[-\varrho, \varrho]$ .  $\square$

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

**2.6.7** Berechnung des Konvergenzradius. *Der Konvergenzradius kann immer durch folgende Formel berechnet werden:*

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

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

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  $a = 0$ .

**2.6.8** Beispiel.

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

haben Konvergenzradius 1.

**2.6.9 Satz.** *Die Potenzreihe*

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

hat Konvergenzradius  $\infty$ . Die hierdurch definierte **Exponentialfunktion** erfüllt die Funktionalgleichung

$$\exp(x) \exp(y) = \exp(x+y).$$

*Beweis.* Mit  $a_k = 1/k!$  gilt  $\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = \lim_{k \rightarrow \infty} \frac{1}{k} = 0$ . Insbesondere konvergiert die Reihe für jede reelle Zahl absolut. Wir multiplizieren mit dem Cauchy-Produkt Satz 2.5.16:

$$\begin{aligned} \exp(x) \exp(y) &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \sum_{l=0}^{\infty} \frac{y^l}{l!} \\ &= \sum_{n=0}^{\infty} \sum_{k+l=n} \frac{x^k y^l}{k! l!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k+l=n} \binom{n}{k} x^k y^l \\ &= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y) . \end{aligned}$$

□

Die Exponentialfunktion definiert die Potenzen der Eulerschen Zahl  $e$  (siehe Beispiel 2.5.19) bezüglich aller reellen Zahlen  $x$ .

**2.6.10 Bemerkung.** Tatsächlich gelten alle Sätze dieses Abschnitts wörtlich und mit denselben Beweisen für Folgen  $(a_n)_{n \geq 0}$  komplexer Zahlen, Folgen  $(f_n)_{n \geq 0}$  von komplexwertigen Funktionen und für Potenzreihen  $\sum_{k=0}^{\infty} a_k z^k$  mit komplexen Koeffizienten. Insbesondere definiert die Exponentialreihe eine Funktion

$$\exp : \mathbb{C} \rightarrow \mathbb{C} .$$

Auch die Funktionen Sinus und Cosinus lassen sich als Potenzreihen schreiben, wie wir später zeigen werden:

$$\begin{aligned} \sin(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} , \\ \cos(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} . \end{aligned}$$

Wegen der Potenzreihenidentität

$$\exp(iz) = \cos(z) + i \sin(z)$$

wird aus der Schreibweise  $e^{i\varphi} = \sin \varphi + i \cos \varphi$  nun ein echter Satz.



# Kapitel 3

## Stetigkeit und 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 Funktionengrenzwerte, Stetigkeit

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

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

$$f(x) \xrightarrow{x \rightarrow a^+} \alpha$$

(ebenso für  $x \rightarrow a^-$ ,  $x \rightarrow a$ ).

**3.1.2 Bemerkung.** Diese Definition von  $\lim_{x \rightarrow a}$  weicht von der Standarddefinition in Mathematikbüchern ab. Dort wird zusätzlich  $f(a) = \lim_{x \rightarrow a} f(x)$  verlangt. Wir folgen weiterhin unserer Hauptquelle Meyberg-Vachenauer.

**3.1.3 Beispiel.** 1. Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  gegeben durch

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} .$$

Dann gilt

$$\lim_{x \rightarrow a} f(x) = \begin{cases} 0 & a < 0 \\ 1 & a > 0 \end{cases}$$

und

$$\lim_{x \rightarrow 0^-} f(x) = 0, \quad \lim_{x \rightarrow 0^+} f(x) = 1 .$$

Der jeweilige Funktionswert  $f(a)$  spiele keine Rolle in den Berechnungen!

2. Wir betrachten  $f = \cos$ ,  $a = 0$ .

$$\lim_{x \rightarrow 0} \cos x = 1 .$$

Da die Funktion gerade ist, genügt es  $x \rightarrow 0^+$  zu betrachten. Wir behandeln zunächst die Nullfolge  $2^{-n}$  für  $n \geq 1$ . Sie ist streng monoton fallend. Aus der Dreieckskonstruktion ist klar, dass  $\cos$  auf  $[0, \pi/2]$  streng monoton fallend ist. Daher ist  $\cos 2^{-n}$  streng monoton wachsend. Gleichzeitig ist  $\cos$  beschränkt, daher hat die Folgen einen Grenzwert  $\alpha$ . Nach den Rechenregeln für Cosinus (Gleichung 3.11) ( $\cos 2x = 2 \cos^2 x - 1$ ) erfüllt er

$$\alpha = 2\alpha^2 - 1 \Rightarrow \alpha = 1, -\frac{1}{2} .$$

Da  $\cos 2^{-n} \geq 0$  für alle  $n$ , folgt  $\alpha = 1$ .

Sei  $(x_n)_{n \geq 0}$  mit  $x_n > 0$  eine beliebige Nullfolge. Sei  $\varepsilon > 0$ . Wir wählen  $N$ , so dass  $1 - \cos 2^{-N} < \varepsilon$ . Dies ist möglich, da  $\alpha = 1$ . Da  $(x_n)_{n \geq 0}$  eine Nullfolge ist, gibt es zu  $2^{-N}$  ein  $n_0$ , so dass  $x_n < 2^{-N}$  für alle  $n \geq n_0$ . Wegen der Monotonie von  $\cos$  folgt

$$\cos x_n \geq \cos 2^{-N} \Rightarrow 1 - \cos x_n \leq 1 - \cos 2^{-N} \leq \varepsilon .$$

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

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

*Beweis.* Wir beweisen beispielhaft die erste Aussage. Sei  $\lim_{x \rightarrow a^+} f(x) = \alpha,$   
 $\lim_{x \rightarrow a^+} g(x) = \beta.$  Wir betrachten eine Folge  $(x_n)_{n \geq 0}$  mit  $x_n \in I, x_n > a.$  Nach Voraussetzung gilt

$$\lim_{n \rightarrow \infty} f(x_n) = \alpha, \quad \lim_{n \rightarrow \infty} g(x_n) = \beta.$$

Nach den Rechenregeln für Grenzwerte in Satz 2.4.5 folgt

$$\lim_{n \rightarrow \infty} (f + g)(x_n) = \lim_{n \rightarrow \infty} f(x_n) + \lim_{n \rightarrow \infty} g(x_n) = \alpha \pm \beta.$$

Nach Definition ist also

$$\lim_{x \rightarrow a^+} (f + g)(x) = \alpha \pm \beta.$$

□

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

*Beweis.* Folgt aus Satz 2.4.6. □

**3.1.6 Beispiel.** Es gelten:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1, \tag{1.1}$$

$$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x} = 0. \tag{1.2}$$

*Beweis.* Wir behandeln die erste Gleichung: Wir betrachten den Kreissektor des Einheitskreises mit Winkel  $x$ . Er hat den Flächeninhalt

$$\frac{x}{2\pi} 1^2 \pi = \frac{x}{2}.$$

Er wird nach unten abgeschätzt durch die Fläche des rechtwinkligen Dreiecks mit Kantenlänge  $\cos x$ , nach oben durch das Dreieck mit Kantenlänge 1. Also

$$\frac{1}{2} \cos x \sin x \leq \frac{x}{2} \leq \frac{1}{2} 1 \tan x.$$

Hieraus folgt

$$\frac{1}{\cos x} \geq \frac{\sin x}{x} \leq \cos x.$$

Wegen  $\lim_{x \rightarrow 1} = 1$  folgt die Behauptung.  $\square$

### 3.1.7 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  $\infty$  oder  $-\infty$  streben.
- 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.
- 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  $x \rightarrow -\infty$ .

**3.1.8 Definition.** Sei  $I \subseteq \mathbb{R}$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  eine Funktion. Die Funktion  $f$  heißt **im Punkt**  $x_0 \in I$  **stetig**, wenn  $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ .

- Falls  $x_0$  ein Randpunkt des Intervalls ist, so ist der Grenzwert nur einseitig zu verstehen.
- Die Funktion  $f$  heißt **auf**  $I$  **stetig**, wenn sie in jedem Punkt  $x_0 \in I$  stetig ist.

### 3.1.9 Satz.

- 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.

*Beweis.* Folgt aus Satz 3.1.4. □

**3.1.10 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  $q(x) \neq 0$ .

*Beweis.* Aus den Rechenregeln und der Stetigkeit der Identität  $x \mapsto x$ . □

**3.1.11 Satz.** Für jede auf einem abgeschlossenen beschränkten Intervall  $a \leq x \leq b$  stetige Funktion gilt:

a) **Schranksatz:** 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 stets 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  $[a, b]$  immer 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  $\varepsilon$  unterscheiden, sobald die Argumente weniger als  $\delta$  voneinander entfernt sind, d.h.

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

Ein besonders wichtiger Fall des Zwischenwertsatzes ist dieser:

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

Wir beweisen den Zwischenwertsatz, da der Beweis auch einen sehr wichtigen Algorithmus enthält, die **Intervallhalbierung** oder **Bisektion**.

*Beweis des Zwischenwertsatzes:*. Zunächst zeigen, wir dass der Zwischenwertsatz aus dem Korollar folgt.

Sei  $I = [a, b]$ ,

$$\inf_{x \in I} f(x) \leq c \leq \sup_{x \in I} f(x) .$$

Wir ersetzen  $f$  durch  $f - c$ . Nach Voraussetzung gibt es  $x_0$  und  $x_1$  mit  $f(x_0) - c < 0 < f(x_1) - c$ . Wir wenden das Korollar auf das Intervall  $[x_0, x_1]$  (bzw.  $[x_1, x_0]$ ) an. Es existiert also eine Nullstelle  $\bar{x}$  von  $f - c$ . Es gilt  $f(\bar{x}) = c$ . Sei also nun  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion so, dass  $f(b) \cdot f(a) < 0$ . Sei ohne Einschränkung  $f(a) < 0, f(b) > 0$ . Wir konstruieren zwei Folgen  $a_n < b_n$  mit  $f(a_n) < 0$  und  $f(b_n) > 0$ .

Sei  $a_0 = a, b_0 = b$ . Wir betrachten  $c = \frac{a+b}{2}$ . Falls  $f(c) = 0$ , so haben wir eine Nullstelle gefunden und sind fertig. Falls  $f(c) > 0$ , so setzen wir  $a_1 = a, b_1 = c$ . Falls  $f(c) < 0$ , so setzen wir  $a_1 = c, b_1 = b_0$ . Wir wiederholen das Verfahren iterativ.

Seien  $a_n$  und  $b_n$  konstruiert. Sei  $c_n = \frac{a_n+b_n}{2}$ . Falls  $f(c_n) = 0$ , so ist die Nullstelle gefunden. Falls  $f(c_n) > 0$ , so setzen wir  $a_{n+1} = a_n, b_{n+1} = c_n$ . Falls  $f(c_n) < 0$ , so setzen wir  $a_{n+1} = c_{n+1}, b_{n+1} = b_n$ .

Die Folge  $(a_n)_{n \geq 0}$  ist nach oben beschränkt und monoton wachsend. Sei  $\bar{x}$  der Grenzwert. Die Folge  $(b_n)_{n \geq 0}$  ist nach unten beschränkt und monoton fallend. Sei  $\bar{y}$  der Grenzwert. Wegen

$$|a_{n+1} - b_{n+1}| = \frac{1}{2}|a_n - b_n| = \frac{1}{2^n}|a - b|$$

stimmen die beiden Grenzwerte überein. Da  $f$  stetig ist, gilt

$$f(\bar{x}) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} f(b_n) .$$

Wegen  $f(a_n) < 0$  für alle  $n$  gilt

$$f(\bar{x}) = \lim_{n \rightarrow \infty} f(a_n) \leq 0 .$$

Wegen  $f(b_n) > 0$  für alle  $n$  gilt

$$f(\bar{x}) = \lim_{n \rightarrow \infty} f(b_n) \geq 0 .$$

Damit ist  $\bar{x}$  die gesuchte Nullstelle. □

Bisektion erlaubt also die Bestimmung einer Nullstelle mit jeder gewünschten Genauigkeit. Das Verfahren ist schnell.

## 3.2 Die Ableitung

**3.2.1 Definition.** 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

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (2.1)$$

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:  $\frac{df(x_0)}{dx}$ ,  $\frac{d}{dx} f$  In der Physik nennt man die Variable oft  $t$  (Zeit) und schreibt dann  $\dot{f}$ .
- 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  $f$** .
- Falls die Ableitung  $f'$  existiert und stetig ist, so heißt  $f$  **stetig differenzierbar**.

**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 $x > 0$

(2.2)

**3.2.2 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  $y = f(x)$  hat im Punkt  $(x_0, f(x_0))$  die Formel

$$y = f'(x_0)(x - x_0) + f(x_0). \quad (2.3)$$

**3.2.3 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 (2.3) ist die „beste“ lineare Approximation von  $f(x)$  nahe  $(x_0, f(x_0))$ .

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

*Beweis.* Sei  $f : I \rightarrow \mathbb{R}$  differenzierbar in  $x_0$ , also existiert

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Sei  $h(x) = f(x) - f(x_0) - f'(x_0)(x - x_0)$ . Es gilt dann wie in 3.2.3

$$\lim_{x \rightarrow x_0} \frac{h(x)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) = 0.$$

Hieraus folgt

$$h(x) = \frac{h(x)}{x - x_0} (x - x_0) \xrightarrow{x \rightarrow x_0} 0 \cdot 0 = 0,$$

also

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + h(x) \xrightarrow{x \rightarrow x_0} f(x_0) + 0 + 0.$$

□

• Stetigkeit von  $f$  ist **nicht** hinreichend für Differenzierbarkeit!

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

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

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

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

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

*Beweis.* Wir zeigen beispielhaft die zweite Aussage im Punkt  $x_0$ . Nach Voraussetzung ist

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x), \quad g(x) = g(x_0) + g'(x_0)(x - x_0) + s(x)$$

mit

$$\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{s(x)}{x - x_0} = 0.$$

Es folgt

$$f(x)g(x) = f(x_0)g(x_0) + [f(x_0)g'(x_0) + f'(x_0)g(x_0)](x - x_0) + t(x)$$

mit

$$t(x) = f'(x_0)g'(x_0)(x - x_0)^2 + r(x)[f(x_0) + f'(x_0)(x - x_0) + s(x)] + s(x)[g(x_0) + g'(x_0)(x - x_0)]$$

mit

$$\frac{t(x)}{x - x_0} \xrightarrow{x \rightarrow x_0} 0.$$

□

**3.2.6 Beispiel.** • Für Polynome  $p(x) = \sum_{i=0}^n a_i x^i$  gilt:

$$p'(x) = \sum_{i=1}^n i a_i x^{i-1}. \quad (2.4)$$

• Wir haben folgenden Spezialfall von Satz 3.2.5 d)

$$\left( \frac{1}{x^n} \right)' = \frac{-n}{x^{n+1}} \quad n \in \mathbb{N}, x \neq 0. \quad (2.5)$$

**3.2.7 Satz.** Die Sinus- und die Cosinusfunktionen sind auf  $\mathbb{R}$  differenzierbar, und es gilt:

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

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

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

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

*Beweis.* Wir behandeln  $\sin x$ . Das Additionstheorem des Sinus (Satz 2.3.2) besagt

$$\sin(x+h) = \sin(x)\cos h + \sin h \cos x.$$

Es folgt mit Beispiel 3.1.6

$$\frac{\sin(x+h) - \sin x}{h} = \sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h} \xrightarrow{h \rightarrow 0} \sin x \cdot 0 + \cos x \cdot 1.$$

□

**3.2.8 Bemerkung.** Insbesondere sind Polynome, rationale Funktionen und die trigonometrischen Funktionen stetig im Definitionsbereich.

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

$$(f(g(x)))' = f'(g(x)) \cdot g'(x). \quad (2.6)$$

*Beweis.* Es gilt

$$\frac{f(g(x)) - f(g(x_0))}{x - x_0} = \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)} \frac{g(x) - g(x_0)}{x - x_0}.$$

Nach Definition konvergiert der zweite Faktor gegen  $g'(x_0)$ . Sei  $(x_n)_{n \geq 1}$  eine gegen  $x_0$  konvergente Folge. Da  $g$  stetig ist, konvergiert  $g(x_n)$  gegen  $g(x_0)$  und daher

$$\lim_{n \rightarrow \infty} \frac{f(g(x_n)) - f(g(x_0))}{g(x_n) - g(x_0)} = f'(g(x_0)).$$

□

**3.2.10 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 (3.14) und Satz 3.2.7 kann man zeigen, dass

$$\frac{d}{dt} e^{iwt} = iw e^{iwt}, \quad t \in \mathbb{R}. \quad (2.7)$$

**3.2.11 Satz.** Sei  $(f_n)_{n \geq 0}$  eine gleichmäßig konvergente Folge von stetigen Funktionen auf einem Intervall  $I$ . Dann ist die Grenzfunktion  $f$  stetig.

*Beweis.* Sei  $x \in I$ ,  $(x_n)_{n \geq 0}$  eine Folge, die gegen  $x$  konvergiert. Zu zeigen ist

$$\lim_{n \rightarrow \infty} f(x_n) = f(x).$$

Sei  $\varepsilon > 0$ . Da die Funktionenfolge gleichmäßig konvergiert gibt es ein  $m_0$ , so dass

$$|f_m(y) - f(y)| < \varepsilon \quad \text{für alle } y \in I, m \geq m_0.$$

Dies gilt insbesondere für  $y = x$  und  $y = x_n$ ,  $n \geq 0$ . Sei  $m \geq m_0$  fest. Die Funktion  $f_m$  ist stetig, d.h.

$$\lim_{n \rightarrow \infty} f_m(x_n) = f_m(x).$$

Es gibt also  $n_0$ , so dass

$$|f_m(x_n) - f_m(x)| < \varepsilon \quad \text{für alle } n \geq n_0.$$

Wir setzen zusammen für  $n \geq n_0$ :

$$|f(x_n) - f(x)| \leq |f(x_n) - f_m(x_n)| + |f_m(x_n) - f_m(x)| + |f_m(x) - f(x)| \leq 3\varepsilon.$$

□

**3.2.12 Satz.** Sei  $(f_n)_{n \geq 0}$  eine punktweise konvergente Folge von stetig differenzierbaren Funktionen auf einem Intervall  $I$  mit Grenzfunktion  $f$ . Konvergiert  $(f'_n)_{n \geq 0}$  auf  $I$  gleichmäßig, dann ist auch die Grenzfunktion  $f$  stetig differenzierbar und es gilt

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

Die gleichmäßige Konvergenz der Ableitungen ist eine notwendige Voraussetzung.

**3.2.13 Beispiel.** Sei  $f_n = \frac{1}{n} \sin nx$ . Die Folge konvergiert gleichmäßig gegen 0. Die Folge der Ableitungen ist  $f'_n = \cos nx$ . Sie konvergiert nicht.

Wir wissen aus Satz 2.6.6, dass Potenzreihen auf abgeschlossenen Intervallen innerhalb des Konvergenzradius gleichmäßig konvergieren.

**3.2.14 Satz.** Sei  $\sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R > 0$ . Dann konvergiert für alle  $\varrho \in (0, R)$  die Reihe  $\sum_{k=1}^{\infty} k a_k x^{k-1}$  absolut und gleichmäßig auf  $|x| \leq \varrho$  gegen die Ableitung

$$\left( \sum_{k=0}^{\infty} a_k x^k \right)' = \sum_{k=1}^{\infty} k a_k x^{k-1} .$$

*Beweis.* Sei  $f_n(x) = \sum_{k=0}^n a_k x^k$ . Die Reihe konvergiert nach Satz 2.6.6. Es gilt

$$f'_n(x) = \sum_{k=1}^n k a_k x^{k-1} .$$

Sei  $g(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}$ . Wie im Beweis von Satz 2.6.6 kann  $g$  abgeschätzt werden gegen die konvergente Reihe

$$\sum_{k=1}^{\infty} c k q^k \quad q < 1 .$$

Die Potenzreihe  $g$  konvergiert, also konvergiert die Folge der  $f'_n$  gleichmäßig. Nun können wir Satz 3.2.12 anwenden.  $\square$

**3.2.15 Satz.** Die Exponentialfunktion ist differenzierbar auf  $\mathbb{R}$ , und es gilt

$$(\exp)' = \exp .$$

*Beweis.* Es gilt auf ganz  $\mathbb{R}$ :

$$\left( \sum_{k=0}^{\infty} \frac{x^k}{k!} \right)' = \sum_{k=1}^{\infty} k \frac{x^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} .$$

$\square$

Insbesondere ist auch die Exponentialfunktion stetig.

### 3.3 Anwendungen der Differentiation

**3.3.1 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  $\varepsilon$ -Umgebung  $(b - \varepsilon, b + \varepsilon)$  von  $b$  gibt, so dass  $f(x) \leq f(b)$  für alle  $x \in D \cap (b - \varepsilon, b + \varepsilon)$ . Analog definiert man **globales Minimum**, **lokales Minimum**. Jedes Minimum und jedes Maximum heißt **Extremum**.*

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

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

*Beweis.* Sei  $x_0$  ein lokales Maximum, also ein globales Maximum auf einem Intervall  $(x_0 - \varepsilon, x_0 + \varepsilon)$ , also  $f(x_0 + h) - f(x_0) \leq 0$  für alle  $h \in (x_0 - \varepsilon, x_0 + \varepsilon)$ . Wir betrachten den Differenzenquotienten in  $x_0$ :

$$\frac{f(x_0 + h) - f(x_0)}{h} \begin{cases} \leq 0 & h > 0 \\ \geq 0 & h < 0 \end{cases}.$$

Der Grenzwert  $h \rightarrow 0$  ist also größer gleich 0 für  $h > 0$  und kleiner gleich 0 für  $h < 0$ . Insgesamt gilt

$$\Rightarrow f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = 0.$$

□

• Aus Satz 3.3.2 folgt, dass

- 1) Randpunkte von  $D$ ,
  - 2) Punkte, wo  $f$  nicht differenzierbar ist, und
  - 3) **stationäre Punkte**, d.h. Punkte wo  $f'(x) = 0$  ist,
- (3.1)

Kandidaten für Extrema sind.

**3.3.3 Satz (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*

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

*Beweis.* Wir betrachten die Funktion

$$F(x) = f(x) - (x - b) \frac{f(b) - f(a)}{b - a}.$$

Die Funktion ist in  $[a, b]$  differenzierbar, also stetig. Sie hat nach dem Satz von Maximum und Minimum (Satz 3.1.11) ein Maximum und ein Minimum. Es gilt  $F(a) = F(b) = f(b)$ , daher liegt ein Extremum  $x_0$  im Inneren. Nach Satz 3.3.2 gilt

$$F'(x_0) = 0 = f'(x_0) - \frac{f(b) - f(a)}{b - a}.$$

□

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

- a)  $f'(x) > 0$  auf  $I \Rightarrow f$  ist auf  $I$  streng monoton wachsend,
- b)  $f'(x) < 0$  auf  $I \Rightarrow f$  ist auf  $I$  streng 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  $I$ .

*Beweis.* Wir behandeln a). Seien  $a < b$  Elemente von  $I$ . Nach dem Mittelwertsatz gibt es  $x_0$  mit

$$\frac{f(b) - f(a)}{b - a} = f'(x_0) > 0.$$

Hieraus folgt  $f(b) > f(a)$ . □

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

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

a)  $f''(x_0) < 0 \Rightarrow f$  hat in  $x_0$  ein lokales Maximum,

b)  $f''(x_0) > 0 \Rightarrow f$  hat in  $x_0$  ein lokales Minimum.

*Beweis.*  $f''(x_0) < 0$  bedeutet, dass  $f'(x_0)$  streng monoton fällt. Daher ist  $f'(x) > f'(x_0)$  für  $x < x_0$  und  $f'(x) < f'(x_0)$  für  $x > x_0$ . Mit dem letzten Satz folgt die Aussage.  $\square$

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

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

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

• Die Punkte  $\lambda x_1 + (1 - \lambda)x_2$  liegen auf der Verbindungsstrecke zwischen  $x_1$  und  $x_2$ . Die Punkte  $\lambda f(x_1) + (1 - \lambda)f(x_2)$  liegen auf der Verbindungsstrecke zwischen  $f(x_1)$  und  $f(x_2)$ . Die Funktion ist konvex, wenn die Funktionswerte auf dem Intervall unterhalb dieser Verbindungsstrecke liegen.

**3.3.8 Satz.** 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.

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

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

*Beweis.* Satz 3.3.5 anwenden auf  $f'$ .  $\square$

**3.3.11 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.

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

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

*Beweis.* Wegen  $g'(x) \neq 0$  auf  $(a, b)$  gilt nach dem Mittelwertsatz  $g(b) \neq g(a)$ . Damit ist die Funktion

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g(x)$$

wohldefiniert auf  $[a, b]$  und differenzierbar. Wir wenden den Mittelwertsatz auf  $F$  an. Es gilt  $F(a) = F(b)$ , also gibt es ein  $x_0 \in (a, b)$  mit  $F'(x_0) = 0$ . Dies ist der gesuchte Punkt.  $\square$

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

a)  $f(x) \rightarrow 0, g(x) \rightarrow 0$  oder  $f(x) \rightarrow \infty, g(x) \rightarrow \infty$  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  $x \rightarrow \pm\infty$ .

*Beweis.* Wir behandeln  $f(x) \rightarrow 0$ ,  $g(x) \rightarrow 0$ . Beide Funktionen setzen sich durch 0 zu einer stetigen Funktion auf  $(a, b]$  fort. Nach dem verallgemeinerten Mittelwertsatz gibt es zu jedem  $a' \in (a, b)$  ein  $x_0$  mit

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

Mit  $a' \rightarrow b^-$  gilt auch  $x_0 \rightarrow b^-$ , also

$$\lim_{a' \rightarrow b^-} \frac{f(a')}{g(a')} = \lim_{x_0 \rightarrow b^-} \frac{f'(x_0)}{g'(x_0)},$$

wobei der rechte Grenzwert existiert, da der linke existiert.  $\square$

### 3.3.14 Beispiel.

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

Man beachte, dass wir diese Regel benutzt haben, um die Differenzierbarkeit von  $\sin x$  herzuleiten. A posteriori, lässt sie sich leicht wieder bestätigen.

Wir wollen nun zwei Verfahren zum Bestimmen von Nullstellen vorstellen. 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  $f(x_n) \rightarrow 0$ .

**3.3.15 Definition.** Sei  $f : [a, b] \rightarrow \mathbb{R}$ . Dann heißt  $x \in [a, b]$  **Fixpunkt** von  $f$ , wenn  $f(x) = x$ .

Das Finden von Nullstellen von  $f$  ist äquivalent zum Finden von Fixpunkten von  $g(x) = f(x) + x$ .

**3.3.16 Satz** Banachscher Fixpunktsatz. *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

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}_0 \quad (3.2)$$

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

3) Es gilt die Abschätzung:

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

Die Bedeutung von 3) ist es, dass die Genauigkeit der Approximation kontrolliert werden kann, ohne den Fixpunkt exakt zu kennen.

*Beweis.* Die Existenz des Fixpunktes wird vom Zwischenwertsatz für die Funktion  $f(x) - x$  garantiert. Angenommen,  $x_1 < x_2$  sind zwei Fixpunkte. Dann gibt es nach dem Mittelwertsatz für  $f(x)$  ein  $x_0 \in (x_1, x_2)$  mit

$$f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{x_2 - x_1}{x_2 - x_1} = 1.$$

Dies ist ein Widerspruch zur Voraussetzung  $f'(x) < 1$  auf  $[a, b]$ .

Sei ab jetzt  $x^*$  der eindeutige Fixpunkt. Sei  $x_0 \in [a, b]$  beliebig und  $(x_n)_{n \geq 0}$  wie im Satz. Es gilt

$$x_n - x^* = f(x_{n-1}) - f(x^*).$$

Nach dem Mittelwertsatz gibt es  $y$  zwischen  $x_{n-1}$  und  $x^*$  mit

$$f'(y) = \frac{f(x_{n-1}) - f(x^*)}{x_{n-1} - x^*}.$$

Zusammen:

$$|x_n - x^*| = |f(x_{n-1}) - f(x^*)| = |f'(y)| |x_{n-1} - x^*| \leq K |x_{n-1} - x_n|.$$

Hieraus folgt induktiv

$$|x_n - x^*| \leq K^n |x_0 - x^*| \Rightarrow \lim_{n \rightarrow \infty} x_n = x^*$$

da  $|K| < 1$ .

Aus der Abschätzung folgt auch

$$|x_n - x^*| \leq K |x_{n-1} - x_n + x_n - x^*| = K |x_{n-1} - x_n| + K |x_n - x^*|$$

und damit 3). □

**3.3.17** 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

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 \in [a, b]. \quad (3.3)$$

**3.3.18 Satz.** Sei  $x^* \in [a, b]$  eine Nullstelle der zweimal stetig differenzierbaren Funktion  $f$ . Falls  $f'(x^*) \neq 0$  ist, gibt es ein kleines Intervall  $I$ , welches  $x^*$  enthält, so dass die in (3.3) 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  $M = \frac{\max\{|f''(x)|; x \in I\}}{\min\{|f'(x)|; x \in I\}}$ .

*Beweis.* Die Voraussetzungen des Fixpunktsatzes sind für  $f/f'$  in einem kleinen Intervall erfüllt.  $\square$

## 3.4 Umkehrfunktionen

Wir erinnern uns, dass eine Funktion  $f : A \rightarrow B$  bijektiv heißt, wenn es zu jedem  $b \in B$  genau ein  $a \in A$  gibt mit  $f(a) = b$ . In dieser Situation existiert die Umkehrfunktion  $f^{-1} : B \rightarrow A$ , die jedem  $b$  genau das  $a$  mit  $f(a) = b$  zuordnet. Es gilt also

$$f \circ f^{-1} = \text{id}, f^{-1} \circ f = \text{id}.$$

**3.4.1 Definition.** Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt **umkehrbar** auf  $D \subset I$ , wenn

$$f|_D : D \rightarrow f(D)$$

bijektiv ist.

- 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  $y = x$ .

- 3.4.2 Satz.** a) Jede strikt monotone Funktion  $f : D \rightarrow \mathbb{R}$  ist umkehrbar.
- b) Jede auf einem Intervall stetige strikt monotone Funktion  $f$  ist umkehrbar auf diesem Intervall. Die Umkehrfunktion ist ebenfalls stetig.
- c) Jede über einem Intervall  $I$  stetig differenzierbare Funktion mit  $f'(x) \neq 0$  für alle  $x \in I$  ist über  $I$  umkehrbar.
- d) 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  $x \in f(I)$  mit  $f'(f^{-1}(x)) \neq 0$  differenzierbar und es gilt

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}. \quad (4.1)$$

*Beweis.* Seien  $x < x'$  in  $D$ . Dann gilt  $f(x) < f(x')$  (bzw.  $f(x) > f(x')$ ), also insbesondere  $f(x) \neq f(x')$ . Die Funktion ist also injektiv. Ist zusätzlich  $f$  stetig und auf einem Intervall definiert, so ist nach Zwischenwertsatz und dem Satz von Maximum und Minimum der Wertebereich ebenfalls ein Intervall. (Stetigkeit übergehen)

Ist  $f$  stetig differenzierbar und  $f'(x) \neq 0$  auf ganz  $D$ , so ist (Zwischenwertsatz für  $f'$ )  $f'(x) < 0$  auf ganz  $D$  oder  $f'(x) > 0$  auf ganz  $D$ . Hieraus folgt nach Satz 3.3.4 die strenge Monotonie.

Um die Differenzierbarkeit von  $\varphi = f^{-1}$  zu zeigen, betrachten wir eine konvergente Folge  $y_n \rightarrow y$  in  $f(D)$ . Sei  $x_n = \varphi(y_n)$  und  $x = \varphi(y)$ . Wegen der Stetigkeit von  $\varphi$  gilt  $\lim_{n \rightarrow \infty} x_n = x$ . Es folgt

$$\lim_{n \rightarrow \infty} \frac{\varphi(y_n) - \varphi(y)}{y_n - y} = \lim_{n \rightarrow \infty} \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x)}{x_n - x}} = \frac{1}{f'(x)}.$$

□

**3.4.3  $n$ -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  $x^n = (-x)^n$ . Aber für  $x \in \mathbb{R}_0^+$  gilt  $f'(x) > 0$  und somit ist  $x^{2k}$  auf  $\mathbb{R}_0^+$  umkehrbar. Die Umkehrfunktion heißt  **$n$ -te Wurzel**, in Zeichen  $\sqrt[n]{\cdot} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ , d.h.  $x = \sqrt[n]{y}$  genau dann wenn  $x^n = y$ .

- 2) Falls  $n$  ungerade ist, d.h.  $n = 2k + 1$ , ist  $f$  auf ganz  $\mathbb{R}$  strikt monoton wachsend, denn  $f'(x) = (2k + 1)x^{2k} > 0$ , für  $x \neq 0$ . Damit ist die  $n$ -te **Wurzel** auf ganz  $\mathbb{R}$  definiert. Zusammenfassend haben wir:

$$y = \sqrt[n]{x} \Leftrightarrow y^n = x \begin{cases} \text{für } x \geq 0, & n \in \mathbb{N} \text{ gerade,} \\ \text{für } x \in \mathbb{R}, & n \in \mathbb{N} \text{ ungerade.} \end{cases} \quad (4.2)$$

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

$$\begin{aligned} x^{\frac{1}{n}} &:= \sqrt[n]{x}, \\ x^{\frac{m}{n}} &:= \left(x^{\frac{1}{n}}\right)^m. \end{aligned} \quad (4.3)$$

Aus der Definition ergeben sich sofort die Rechenregeln

$$\begin{aligned} x^\alpha x^\beta &= x^{\alpha+\beta}, \\ (x^\alpha)^\beta &= x^{\alpha\beta}, \\ x^\alpha y^\alpha &= (xy)^\alpha \end{aligned} \quad (4.4)$$

und die Formel für die Ableitung

$$(x^\alpha)' = \alpha x^{\alpha-1}, \quad x > 0. \quad (4.5)$$

**3.4.4 Arcusfunktionen.** *Wir betrachten jetzt 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 monoton wachsend, denn  $\sin'(x) = \cos x > 0$  falls  $x \in (-\pi/2, \pi/2)$ . Die Umkehrfunktion heißt **Arcussinus**, in Zeichen  $\arcsin$ , und ist definiert durch

$$\begin{aligned} \arcsin : [-1, 1] &\rightarrow [-\pi/2, \pi/2] \\ y = \arcsin x &\Leftrightarrow (\sin y = x) \wedge (y \in [-\pi/2, \pi/2]). \end{aligned} \quad (4.6)$$

Für die Ableitung gilt:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \quad \text{für } -1 < x < 1. \quad (4.7)$$

*Beweis.* Aus der allgemeinen Formel folgt zunächst

$$\frac{d}{dx} \arcsin x = \frac{1}{\cos(\arcsin x)} .$$

Wegen  $\cos y = \sqrt{1 - \sin^2 y}$  folgt

$$\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1 - \sin^2 \arcsin x}} = \frac{1}{\sqrt{1 - x^2}} .$$

□

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

$$\begin{aligned} \arccos : [-1, 1] &\rightarrow [0, \pi] \\ y = \arccos x &\Leftrightarrow (\cos y = x) \wedge (y \in [0, \pi]). \end{aligned} \quad (4.8)$$

Es gilt

$$(\arccos x)' = \frac{-1}{\sqrt{1-x^2}}, \quad \text{für } -1 < x < 1 \quad (4.9)$$

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

$$\begin{aligned} \arctan : \mathbb{R} &\rightarrow (-\pi/2, \pi/2) \\ \arctan x = y &\Leftrightarrow (\tan y = x) \wedge (y \in (-\pi/2, \pi/2)) \end{aligned} \quad (4.10)$$

$$\begin{aligned} \operatorname{arccot} : \mathbb{R} &\rightarrow (0, \pi) \\ \operatorname{arccot} x = y &\Leftrightarrow \cot y = x \wedge y \in (0, \pi) \end{aligned} \quad (4.11)$$

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

$$\begin{aligned} (\arctan x)' &= \frac{1}{1+x^2}, \\ (\operatorname{arccot} x)' &= \frac{-1}{1+x^2} \end{aligned} \quad (4.12)$$

## 3.5 Exponential- und Logarithmusfunktion

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

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

b) Für alle  $x \in \mathbb{R}$  gilt:

$$\frac{d}{dx} \exp(x) = \exp(x). \quad (5.1)$$

Somit ist insbesondere  $e^x$  stetig differenzierbar.

c) Es gelten:

$$\begin{aligned} \exp(x+y) &= \exp(x)\exp(y), & \forall x, y \in \mathbb{R}, \\ \exp(-x) &= \frac{1}{\exp x}, & \forall x \in \mathbb{R}. \end{aligned} \quad (5.2)$$

d) Die  $e$ -Funktion ist strikt monoton wachsend und konvex.

e)  $\lim_{x \rightarrow \infty} \exp x = \infty$ ,  $\lim_{x \rightarrow -\infty} \exp x = 0$

f)  $\lim_{x \rightarrow \infty} \frac{\exp x}{x^n} = \infty$ ,  $\forall n \in \mathbb{N}$ .

*Beweis.* Die Aussagen b) und c) kennen wir schon, vergleiche Satz 2.6.9, Satz 3.2.15. Für  $x \geq 0$  ist die Exponentialfunktion nach Definition positiv. Wegen 5.2 folgt dies auch für negative  $x$ .

d) folgt aus den Eigenschaften von  $\exp'$ .

Wegen der strengen Monotonie gilt

$$\lim_{x \rightarrow \infty} \exp x = \lim_{n \rightarrow \infty} \exp n = \lim_{n \rightarrow \infty} e^n.$$

Wegen  $e > 1$  divergiert diese Folge gegen  $\infty$ .

Formel f) folgt induktiv mit der Regel von de l'Hospital (Satz 3.3.13):

$$\lim_{x \rightarrow \infty} \frac{\exp(x)}{x^n} = \lim_{x \rightarrow \infty} \frac{\frac{d}{dx} \exp(x)}{\frac{d}{dx} x^n} = \lim_{x \rightarrow \infty} \frac{\exp(x)}{nx^{n-1}}.$$

□

Als streng monotone Funktion hat  $\exp : \mathbb{R} \rightarrow \mathbb{R}^+$  eine Umkehrfunktion.

**3.5.2** Der natürliche Logarithmus. Die Exponentialfunktion wächst strikt monoton, also existiert nach Satz 3.4.2 auf  $(0, \infty)$  eine Umkehrfunktion. Diese heißt der **natürliche Logarithmus** und ist definiert durch:

$$\begin{aligned} \ln : (0, \infty) &\rightarrow \mathbb{R} \\ y = \ln x &\Leftrightarrow \exp(y) = x. \end{aligned} \quad (5.3)$$

Insbesondere gilt:

$$\begin{aligned} \ln(\exp x) &= x, & x \in \mathbb{R}, \\ \exp(\ln x) &= x, & x > 0. \end{aligned} \quad (5.4)$$

$$\begin{aligned} \ln e &= 1, \\ \ln 1 &= 0, \end{aligned} \quad (5.5)$$

$$\begin{aligned} x \in (0, 1) &\Rightarrow \ln x < 0, \\ x \in (1, \infty) &\Rightarrow \ln x > 0, \end{aligned} \quad (5.6)$$

$$\begin{aligned} \lim_{x \rightarrow 0^+} \ln x &= -\infty, \\ \lim_{x \rightarrow \infty} \ln x &= \infty. \end{aligned} \quad (5.7)$$

In der Mathematik schreibt man meist  $\log$  und meint den natürlichen Logarithmus  $\ln$ .

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

$$\frac{d}{dx} \ln x = \frac{1}{x}, \quad (5.8)$$

$$\begin{aligned} \ln(xy) &= \ln x + \ln y, \\ \ln \frac{x}{y} &= \ln x - \ln y, \end{aligned} \quad (5.9)$$

wobei  $x, y > 0$ .

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

$$\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[n]{x}} = 0. \quad (5.10)$$

*Beweis.* Die Formel für die Ableitung folgt aus Satz 5.2:

$$\frac{d}{dx} \ln x = \frac{1}{\frac{d}{dx} \exp(\ln(x))} = \frac{1}{\exp \ln x} = \frac{1}{x}.$$

Hieraus und den Eigenschaften der Exponentialfunktion folgen alle anderen Aussagen.  $\square$

### 3.5.4 Allgemeine Exponentialfunktionen und Logarithmen.

Sei  $a > 0$ . Für  $r \in \mathbb{Q}$  und  $x \in \mathbb{R}$  gilt

$$(\exp(x))^r = \exp(rx).$$

Die Wahl  $x = \ln a$  liefert also

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

Auf Grund der Stetigkeit der Exponentialfunktion definiert man daher

$$a^x := \exp(x \ln a) \quad \text{für alle } x \in \mathbb{R}, a > 0 \quad (5.11)$$

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

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

$$\begin{aligned} a^x a^y &= a^{x+y}, \\ (ab)^x &= a^x b^x, \end{aligned} \quad (5.12)$$

$$\begin{aligned} (a^x)^y &= a^{xy}, \\ \ln(a^x) &= x \ln a, \end{aligned} \quad (5.13)$$

$$\frac{d}{dx} a^x = a^x \ln a. \quad (5.14)$$

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

$$x^\alpha = e^{\alpha \ln x}. \quad (5.15)$$

Man berechnet sofort

$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \frac{\alpha}{x} = \alpha x^{\alpha-1}. \quad (5.16)$$

- 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:

$$\begin{aligned} \ln_a : (0, \infty) &\rightarrow \mathbb{R} \\ y = \ln_a(x) &\Leftrightarrow a^y = x. \end{aligned} \quad (5.17)$$

Es gelten :

$$\begin{aligned} \ln_a(xy) &= \ln_a(x) + \ln_a(y), \\ \frac{d}{dx} \ln_a(x) &= \frac{1}{x \ln a}. \end{aligned} \quad (5.18)$$

**3.5.6 Definition.** 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 3.5.6 und den Eigenschaften der  $e$ -Funktion erhält man folgende Rechenregeln:

$$\begin{aligned} \sinh(-x) &= -\sinh(x) \\ \cosh(-x) &= \cosh(x) \end{aligned} \quad (5.19)$$

$$\begin{aligned} \sinh(x+y) &= \sinh(x) \cosh(y) + \cosh(x) \sinh(y), \\ \cosh(x+y) &= \cosh(x) \cosh(y) + \sinh(x) \sinh(y), \end{aligned} \quad (5.20)$$

$$(\cosh(x))^2 - (\sinh(x))^2 = 1. \quad (5.21)$$

Für die Ableitungen gilt:

$$\begin{aligned} \sinh(x)' &= \cosh(x), \\ \cosh(x)' &= \sinh(x), \\ \tanh(x)' &= \frac{1}{\cosh^2(x)}. \end{aligned} \quad (5.22)$$

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

$$\begin{aligned}\operatorname{arsinh}(x) &= \ln\left(x + \sqrt{x^2 + 1}\right), & x \in \mathbb{R}, \\ \operatorname{arcosh}(x) &= \ln\left(x + \sqrt{x^2 - 1}\right), & x \geq 1.\end{aligned}\tag{5.23}$$

- Die Kettenregel und (4.33) liefern:

$$\begin{aligned}\frac{d}{dx} \operatorname{arsinh}(x) &= \frac{1}{\sqrt{1+x^2}}, & x \in \mathbb{R}, \\ \frac{d}{dx} \operatorname{arcosh}(x) &= \frac{1}{\sqrt{x^2-1}}, & x > 1.\end{aligned}\tag{5.24}$$

## 3.6 Beispiele

**3.6.1 Beispiel.** Eine Lampe wird im Punkt  $B = (0, h)$  befestigt. Ihre Leuchtstärke im Punkt  $A = (a, 0)$  ist gegeben durch

$$y(h) = h(a^2 + h^2)^{-3/2}.$$

Wir wollen maximale Helligkeit in  $A$  erreichen.

Es gilt  $y(0) = 0$ ,

$$\lim_{h \rightarrow \infty} \frac{h}{(a^2 + h^2)^{3/2}} = \lim_{h \rightarrow \infty} \frac{1}{3/2(a^2 + h^2)^{1/2}2h} = 0.$$

Wegen  $h(1) \neq 0$  hat die Funktion wenigstens eine Maximalstelle in  $(0, \infty)$ .

Wir bestimmen die stationären Punkte:

$$\frac{d}{dh} y = (a^2 + h^2)^{-3/2} + h(-3/2)(a^2 + h^2)^{-5/2}2h = \frac{a^2 + h^2 - 3h^2}{(a^2 + h^2)^{5/2}}$$

hat als einzige Nullstelle in  $h > 0$

$$h = \frac{a}{\sqrt{2}}.$$

Dies ist demnach das eindeutige Maximum.

**3.6.2 Beispiel.** (Thermodynamik) Die Molwärme (Ableitung der enthaltenen Energiemenge nach der Temperaturänderung je Mol) eines zweiatomigen Gases ist bei festem Volumen als Funktion der absoluten Temperatur  $T$  gegeben durch

$$c(T) = R \frac{(T_0/T)^2 e^{T_0/T}}{(e^{T_0/T} - 1)^2}$$

mit der Gaskonstanten  $R$  und der charakteristischen Temperatur  $T_0$ . Wir betrachten die Grenzwerte  $T \rightarrow 0$  und  $T \rightarrow \infty$ . Zur Vereinfachung setzen wir  $x = T_0/T$ . Mit der Regel von Hospital:

$$\begin{aligned} \lim_{T \rightarrow 0^+} c(T) &= \lim_{x \rightarrow \infty} R \frac{x^2 e^x}{(e^x - 1)^2} \\ &= \lim_{x \rightarrow \infty} R \frac{(2x + x^2)e^x}{2(e^x - 1)e^x} \\ &= \lim_{x \rightarrow \infty} R \frac{2 + 2x}{2e^x} \\ &= \lim_{x \rightarrow \infty} R \frac{2}{2e^x} = 0 \\ \lim_{T \rightarrow \infty} c(T) &= \lim_{x \rightarrow 0^+} R \frac{x^2 e^x}{(e^x - 1)^2} \\ &= \lim_{x \rightarrow \infty} R \frac{(2x + x^2)}{2(e^x - 1)} \\ &= \lim_{x \rightarrow \infty} R \frac{1 + x}{e^x} = R \end{aligned}$$

**3.6.3 Beispiel.** Wir wollen

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

herleiten. Wir betrachten die Funktion

$$f(x) = \left(1 + \frac{1}{x}\right)^x = \exp\left(x \ln\left(1 + \frac{1}{x}\right)\right).$$

Wir wollen gerne

$$\lim_{n \rightarrow \infty} f(n) = \lim_{x \rightarrow \infty} f(x)$$

ausnutzen. Dies gilt, wenn der rechte Grenzwert existiert. Wegen der Stetigkeit der Exponentialfunktion genügt es,

$$\lim_{x \rightarrow \infty} x \ln \left( 1 + \frac{1}{x} \right)$$

zu berechnen. Der erste Faktor streben gegen  $\infty$ , der zweite gegen 0. Wir wende die Regel von Hospital an:

$$\lim_{x \rightarrow \infty} \frac{\ln \left( 1 + \frac{1}{x} \right)}{x^{-1}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{1+\frac{1}{x}} \cdot \frac{-1}{x^2}}{-x^{-2}} = 1 .$$

Es folgt

$$\lim_{x \rightarrow \infty} f(x) = \exp(1) = e .$$

**3.6.4 Beispiel** Fixpunktiteration. Gesucht wird eine Lösung von

$$x = \exp(x^2 - 2) .$$

Sei  $f(x)$  die rechte Seite. Es gilt

$$f'(x) = 2x \exp(x^2 - 2) .$$

Daher hat  $f$  ein Minimum in  $x$  und wächst streng monoton für  $x > 0$ . Es gilt

$$f''(x) = (2 + 4x^2) \exp(x^2 - 2) .$$

Die zweite Ableitung ist überall größer 0, daher wächst die Ableitung streng monoton auf ganz  $\mathbb{R}$ . Wir suchen ein Intervall, in dem  $|f'(x)| < 1$ . Da die Gleichung nicht explizit lösbar ist, experimentieren wir:

$x$	$f'(x)$
0	0
1	$2e^{-1}$
2	$4e^2$

Innerhalb  $[-1, 1]$  ist also  $|f'(x)| < 1$ . Es gilt  $f(1) = f(-1) = e^{-1} < 1$ . In diesem Intervall sind also die Voraussetzung des Fixpunktsatzes erfüllt mit

$$K = f'(1) = e^{-1} \Rightarrow \frac{K}{1-K} = \frac{1}{e-1} < 0,6 .$$

In diesem Intervall hat die Funktion also einen eindeutigen Fixpunkt. Wir arbeiten mit dem Startwert 0:

$n$	$x_n$	$x_n - x_{n-1}$
0	0	
1	$e^{-2} = 0,1353$	0,1353
2	$e^{-1,9817} = 0,1378$	0,002
3	$e^{-1,9810} = 0,1379$	0,0001

Damit sind die ersten drei Dezimalstellen bestimmt, die vierte mit Genauigkeit kleiner 0,6, also liegt der genaue Fixpunkt in  $[0,1378, 0,1380]$ .

**3.6.5 Bemerkung** Geometrische Interpretation des Newton-Verfahrens. Die Iterationsvorschrift lautet:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Dies hat eine geometrische Interpretation: Die Kurventangente

$$y = f'(x_n)(x - x_n) + f(x_n)$$

schneidet die  $x$ -Achse in  $x_{n+1}$ . Im günstigsten Fall liegt  $x_{n+1}$  näher an der Nullstelle als  $x_n$ .

**3.6.6 Beispiel.** Wir wollen mit dem Newton-Verfahren die Zahl  $\sqrt{5}$  bestimmen. Es ist eine Nullstelle von  $x^2 - 5$ . Die Ableitung  $2x$  ist positiv für  $x > 0$ . Es gibt also ein kleines Intervall um  $\sqrt{5}$ , in dem das Verfahren funktioniert. Die Iterationsvorschrift lautet

$$x_{n+1} = x_n - \frac{x_n^2 - 5}{2x_n} = \frac{x_n^2 + 5}{2x_n}.$$

Wir beginnen mit  $x = 1$ .

$n$	$x_n$
0	1
1	$\frac{6}{2} = 3$
2	$\frac{14}{6} = 2,333$
3	$\frac{49/9 + 45/9}{14/3} = \frac{47}{21} = 2,2381$

Die Folge scheint zu konvergieren.

Alternativ: Wir suchen eine Nullstelle von  $1 - \frac{5}{x^2}$ . Die Ableitung  $10/x^3$  ist positiv für  $x > 0$ . Die Iterationsvorschrift lautet

$$\begin{aligned} x_{n+1} &= x_n - \frac{1 - \frac{5}{x_n^2}}{\frac{10}{x_n^3}} \\ &= x_n - \frac{x_n^2 - 5}{x_n^2} \frac{x_n^3}{10} \\ &= x_n \frac{15 - x_n^2}{10} \end{aligned}$$

Wir probieren den Startwert 3:

$n$	$x_n$
0	3
1	$\frac{18}{10} = \frac{9}{5} = 1,8$
2	2,1
3	2,22
4	2,23601

Diese Folge konvergiert anscheinend.

Tatsächlich ist eine hinreichende Bedingung:

1.  $f$  hat in  $[a, b]$  eine Nullstelle  $x^*$
2.  $f'(x)$  hat in  $[a, b]$  keine Nullstelle
3.  $f$  ist konvex oder konkav
4. Die Iterationswerte  $x_1$  zu  $x_0 = a, b$  liegen im Intervall.

Dann konvergiert für jeden Startwert in  $[a, b]$  die Rekursionsfolge gegen  $x^*$  und

$$|x^* - x_k| \leq \frac{M}{2} |x_k - x_{k-1}|^2$$

mit  $M$  wie in 3.3.17.

Für die Funktion  $x^2 - 5$  können wir  $a = 1, b = 3$  wählen. Es gilt

$$f(1) = -4 < 0, f(3) = 4 > 0 .$$

$x_1$  für 1 und 3 haben wir oben berechnet, und sie liegen im Intervall. Wegen  $f'' = 2$  ist die Funktion konvex. Die Folge konvergiert.

Für die Funktion  $1 - \frac{5}{x^2}$  können wir  $a = 1$ ,  $b = 3$  wählen. Es gilt

$$f(1) = 1 - 5 = -4 < 0, \quad f(3) = 1 - \frac{5}{9} > 0.$$

$x_1$  für 3 haben wir oben berechnet. Für den Startwert 1 erhalten wir

$$x_1 = \frac{15 - 1}{10} = \frac{14}{10} \in [1, 3].$$

$f'$  ist positiv und  $f''(x) = -30x^{-4}$  ist negativ. Die Folge konvergiert tatsächlich.

**3.6.7 Beispiel** Radioaktiver Zerfall. Experimentell findet man heraus, dass die Anzahl der radioaktiven Zerfälle unter  $N(t)$  Teilchen in einem kleinen Intervall  $\Delta t$  proportional ist zu  $N(t)$ , d.h. die Zerfallsrate ist konstant. Unter der Annahme, dass  $N(t)$  differenzierbar ist ergibt sich

$$\dot{N}(t) = -kN(t).$$

Dies führt also auf die Funktion

$$N(t) = N(0) \exp(-kt).$$

Die **Halbwertszeit**  $T$  ist definiert als die Zeit, in der die Hälfte des Materials zerfallen ist, also  $N(T) = N(0)/2$ . Wir lösen die Gleichung

$$\frac{1}{2} = \exp(-kT) \Leftrightarrow 2 = \exp(kT) \Leftrightarrow \ln 2 = kT.$$

Es gilt also

$$T = \frac{1}{k} \ln 2.$$

**3.6.8 Beispiel** Höhenformel. Der Zusammenhang zwischen Druck  $p$  und Höhe  $h$  in einem gasgefüllten Behälter wird durch

$$p = p_0 \left( 1 - \frac{\gamma - 1}{\gamma} \frac{\rho_0}{p_0} gh \right)^{\frac{\gamma}{\gamma - 1}}$$

beschrieben. Dabei hängt  $\gamma$  vom verwendeten Gas. Wir betrachten die Annäherung an ein ideales Gas  $\gamma \rightarrow 1$ . Sei  $x = \frac{\gamma}{\gamma - 1}$ . Es ist  $\lim_{\gamma \rightarrow 1} x \rightarrow \infty$ .

Sei außerdem zur Vereinfachung  $a = -\frac{\rho_0}{p_0} gh$ . Wir berechnen

$$\lim_{x \rightarrow \infty} \left( 1 + \frac{a}{x} \right)^x = \lim_{x \rightarrow \infty} \exp(x \ln(1 + a/x)).$$

Wie in Beispiel 3.6.3 finden wir

$$\lim_{x \rightarrow \infty} x \ln \left( 1 + \frac{a}{x} \right) = a$$

und daher wegen Stetigkeit von  $\exp$

$$\lim_{\gamma \rightarrow 1} p = p_0 \exp \left( -\frac{\rho_0}{p_0} gh \right) .$$



# Kapitel 4

## Integration

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

### 4.1 Das bestimmte Integral

• **Motivation:** Sei  $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$  eine Zerlegung von  $[a, b]$ , dann existieren nach dem Mittelwertsatz (Kapitel 3 Satz 3.3.3)  $\xi_i \in (x_{i-1}, x_i)$  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  $\xi_i$  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 Definition.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine Funktion. 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

$$Z_n := \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \quad (1.1)$$

die **Riemannsche Summe von  $f$** . Die Funktion  $f$  heißt **(Riemann) integrierbar**, wenn 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 und der Grenzwert unabhängig von der gewählten Zerlegung des Intervalls  $[a, b]$  und der Zwischenpunkte  $\xi_i$  ist.

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.

$$\int_a^b f(x) dx := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}). \quad (1.2)$$

- Um Fallunterscheidungen zu vermeiden setzen wir

$$\begin{aligned} \int_a^a f(x) dx &:= 0, \\ \int_b^a f(x) dx &:= - \int_a^b f(x) dx, \quad b > a. \end{aligned} \quad (1.3)$$

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

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

Genau genommen wird erst durch das Integral der Flächeninhalt definiert.

**4.1.3 Definition.** 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**.

**4.1.4 Satz.** *Stückweise stetige Funktionen sind integrierbar.*

*Beweis.* Wir behandeln den Fall  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Sei  $x_0 < x_1 < \dots < x_n$  eine Zerlegung. Wir können die Riemannsche Summe  $Z_n$  nach unten abschätzen durch die Untersumme

$$U_n = \sum_{i=1}^n \min_{x \in [x_{i-1}, x_i]} f(x)(x_i - x_{i-1})$$

und nach oben durch die Obersumme

$$O_n = \sum_{i=1}^n \max_{x \in [x_{i-1}, x_i]} f(x)(x_i - x_{i-1})$$

Wir zeigen

$$\lim_{n \rightarrow \infty} U_n = \lim_{n \rightarrow \infty} O_n$$

und damit auch die Konvergenz von  $\lim_{n \rightarrow \infty} Z_n$  gegen den selben Grenzwert. Die Folge der Obersummen ist monoton fallend, die der Untersummen monoton wachsend. Wegen  $U_n < O_n$  sind beide Folgen beschränkt, also konvergent.

Sei  $\varepsilon > 0$ . Nach Satz 3.1.10 d) ist  $f$  auf  $[a, b]$  gleichmäßig stetig, d.h. zu  $\varepsilon$  gibt es ein  $\delta > 0$ , so dass

$$|f(x) - f(x')| \leq \varepsilon \quad \text{für alle } |x - x'| \leq \delta .$$

Sei  $n$  groß genug, so dass  $|x_i - x_{i-1}| \leq \delta$  für alle  $i$ . Dann gilt

$$|f(x) - f(x')| \leq \varepsilon \quad \text{für alle } x, x' \in [x_{i-1}, x_i] .$$

Es folgt

$$\max_{x \in [x_{i-1}, x_i]} f(x) - \min_{x \in [x_{i-1}, x_i]} f(x) \leq \varepsilon$$

und daher

$$O_n - U_n \leq \sum_{i=1}^n \varepsilon(x_i - x_{i-1}) = \varepsilon(b - a) .$$

Damit ist die Gleichheit der Grenzwerte gezeigt.

Wir betrachten nun eine andere Folge von Zerlegungen. Die Untersumme zu einer Zerlegung ist stets kleiner als die Obersumme zur anderen (und umgekehrt). Daher stimmen die Grenzwerte auch für verschiedene Zerlegungen überein.  $\square$

**4.1.5 Bemerkung.** Da  $f$  auf jedem Teilintervall sein Maximum und Minimum annimmt, sind Obersumme und Untersumme spezielle Riemannsche Summen.

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

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

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

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

*Beweis.* Wir betrachten a). Wir wählen eine Folge von Zerlegungen

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

und eine Folge von Zwischenwerten. Nach Definition gilt

$$\begin{aligned} \int_a^b (\alpha f(x) + \beta g(x)) dx &= \lim_{n \rightarrow \infty} \sum_{i=1}^n (\alpha f(\xi_i) + \beta g(\xi_i))(x_i - x_{i-1}) \\ \int_a^b f(x) dx &= \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \\ \int_a^b g(x) dx &= \lim_{n \rightarrow \infty} \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) \end{aligned}$$

Die Aussage folgt nun aus den Rechenregeln für Grenzwerte von Folgen. Die übrigen Aussagen folgen ähnlich direkt aus der Definition.  $\square$

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

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

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

b) Es gilt die Ungleichung

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

*Beweis.* Wie Satz 4.1.6 aus der Definition und den Abschätzungen für Grenzwerte von Folgen.  $\square$

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

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

*Beweis.* Sei  $m$  das Minimum und  $M$  das Maximum von  $f$  auf  $[a, b]$ . Auf  $[a, b]$  gilt

$$mg(x) \leq f(x)g(x) \leq Mg(x) \Rightarrow m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx.$$

Daher ist

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

mit einer Zahl  $c \in [m, M]$ . Nach dem Zwischenwertsatz gibt es ein  $\xi \in [a, b]$  mit  $f(\xi) = c$ .  $\square$

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

$$\int_a^b f(x) dx = f(\xi) (b - a). \quad (1.4)$$

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

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

a) Die durch

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

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

$$\frac{d}{dx} \left( \int_a^x f(t) dt \right) = f(x). \quad (1.5)$$

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

b) Mit einer beliebigen Stammfunktion  $F$  von  $f$  gilt:

$$\int_a^b f(x) dx = F(x) \Big|_a^b := F(b) - F(a). \quad (1.6)$$

*Beweis.* Wir bilden den Differenzenquotienten für  $F_a$ . Nach den Rechenregeln für Integrale und dem Mittelwertsatz 4.1.9 gilt

$$\begin{aligned} F_a(x+h) - F_a(x) &= \int_a^{x+h} f(t) dt - \int_a^x f(t) dt \\ &= \int_x^{x+h} f(t) dt \\ &= f(\xi_h)h \end{aligned}$$

mit  $\xi_h \in [x, x+h]$ . Mit  $h \rightarrow 0$  gilt  $\xi_h \rightarrow x$ . Da  $f$  stetig ist, folgt

$$F'_a(x) = \lim_{h \rightarrow 0} \frac{F_a(x+h) - F_a(x)}{h} = \lim_{h \rightarrow 0} f(\xi_h) = f(x).$$

Ist  $F$  eine andere Stammfunktion, so gilt

$$(F - F_a)' = F' - F'_a = f - f = 0.$$

Es ist ein einfache Konsequenz des Mittelwertsatzes der Differentialrechnung 3.3.3, dass dann  $F - F_a$  konstant ist.

In b) ist nun  $F_a = F - c$ . Es folgt also

$$\int_a^b f(x) dx = F_a(b) - F_a(a) = F(b) - c - F(a) + c = F(b) - F(a).$$

□

• Für die Berechnung des bestimmten Integrals ergibt sich damit das Verfahren:

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

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

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

• Nach Satz 4.1.10 a) gilt

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

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

$$\int f(x) dx = F + c \quad \Leftrightarrow \quad F' = f. \quad (1.7)$$

- Aus den Differentiationsformeln folgen somit folgende Integrationsformeln:

$F(x)$	$f(x) = F'(x)$	Bemerkungen
$\frac{1}{n+1} x^{n+1}$	$x^n$	$n \neq -1$
$\frac{1}{a+1} x^{a+1}$	$x^a$	$a \in \mathbb{R}, x > 0$
$\ln  x $	$x^{-1}$	$x \neq 0$
$-\cos x$	$\sin x$	
$\sin x$	$\cos x$	
$\tan x$	$\frac{1}{\cos^2 x}$	$x \neq (k + \frac{1}{2})\pi, k \in \mathbb{Z}$
$\cot x$	$\frac{-1}{\sin^2 x}$	$x \neq k\pi, k \in \mathbb{Z}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$ x  < 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

Alle Rechenregeln für die Ableitung werden nun zu Rechenregeln für das Integral.

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

$$\int (af(x) + bg(x)) dx = a \int f(x) dx + b \int g(x) dx. \quad (2.1)$$

*Beweis.* Nach dem Hauptsatz der Differential und Integralrechnung berechnen wir die Stammfunktion als  $\int_a^x$ . Die Aussage folgt dann aus der Rechenregel für das bestimmte Integral Satz 4.1.6.  $\square$

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

$$\int u'(x)v(x) dx = uv - \int u(x)v'(x) dx. \quad (2.2)$$

Für das bestimmte Integral erhalten wir

$$\int_a^b u'v dx = u(x)v(x)\Big|_a^b - \int_a^b uv' dx. \quad (2.3)$$

**4.2.3** Beispiel. Aus (2.2) folgt mit  $u' = 1$

$$\int v(x) dx = xv - \int xv'(x) dx, \quad (2.4)$$

insbesondere also:

$$\int \ln x dx = x \ln x - \int x \frac{1}{x} dx = x(\ln x - 1) + c. \quad (2.5)$$

**4.2.4.** Für Integrale der Form

$$\begin{aligned} S_n &:= \int_a^b (\sin x)^n dx, & C_n &:= \int_a^b (\cos x)^n dx, \\ A_n &:= \int_a^b x^n \sin x dx, & B_n &:= \int_a^b x^n \cos x dx, \\ E_n &:= \int_a^b x^n e^x dx, & L_n &:= \int_a^b (\ln x)^n dx \end{aligned} \quad (2.6)$$

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

*Beweis.* Wir behandeln beispielhaft  $S_2$ .

Mit  $u' = \sin x$ ,  $v = \sin x$  erhalten wir  $u = -\cos x$ ,  $v' = \cos x$

$$\begin{aligned} \int \sin^2 x dx &= -\cos x \sin x - \int -\cos x \cos x dx \\ &= -\cos x \sin x + \int \cos^2 x \\ &= -\cos x \sin x + \int (1 - \sin^2 x) dx \\ &= -\cos x \sin x + x - \int \sin^2 x dx . \end{aligned}$$

Diese Gleichung lösen wir auf zu

$$\int \sin^2 x dx = -\frac{1}{2}(x - \cos x \sin x) .$$

□

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

$$\int f(g(x)) g'(x) dx = F(g(x)) + c \quad (2.7)$$

und für das bestimmte Integral

$$\int_a^b f(g(x)) g'(x) dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(t) dt. \quad (2.8)$$

- Ein konkreter Fall ist

$$\int f^k(x) f'(x) dx = \frac{1}{k+1} f^{k+1} \quad k \neq -1 .$$

Besonders interessant ist der Fall  $k = 1$ :

$$\int f f' dx = \frac{1}{2} f^2 .$$

- Es gibt zwei Möglichkeiten (2.8) anzuwenden. Beide nutzen die suggestive Schreibweise  $g' = \frac{dg}{dx}$  und lösen nach  $dx$  bzw.  $dg$  auf.

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  $\int f(x) dx = H(h(x)) + c$

**4.2.6 Bemerkung.** Anders als bei der Differentiation gibt es keine klare Vorschrift, wie ein gegebenes kompliziertes Integral auf eines der elementaren wohlbekanntesten Integrale zurückzuführen ist. Umgekehrt lassen sich jedoch gefundene Formeln leicht durch Differentiation verifizieren.

**4.2.7 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

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

**4.2.8 Lemma** (Orthogonalitätsbeziehungen). 1. Für  $k \in \mathbb{Z} \setminus \{0\}$  gilt

$$\begin{aligned} \int_0^{2\pi} \sin(kx) dx &= 0, \\ \int_0^{2\pi} \cos(kx) dx &= 0. \end{aligned} \tag{2.9}$$

2. Sei  $m, n \in \mathbb{N}_0$ . Dann gilt:

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

$$\int_0^{2\pi} \cos(mx) \cos(nx) dx = \begin{cases} 2\pi & \text{falls } m = n = 0, \\ 0 & \text{falls } m \neq n, \\ \pi & \text{falls } m = n \neq 0, \end{cases} \quad (2.10)$$

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

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

$$\int_0^{2\pi} \sin(kx + \varphi) dx = -\frac{1}{k} \cos(kx + \varphi) \Big|_0^{2\pi} = 0. \quad (2.11)$$

Somit erhält man für  $\varphi = 0$  oder  $\varphi = -\pi$  die Formeln in (2.9).

Mit Hilfe der Additionstheoreme erhalten wir:

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

Also folgt für das Integral

$$\int_0^{2\pi} \sin(mx) \sin(nx) = \frac{1}{2} \int_0^{2\pi} \cos((m-n)x) dx - \frac{1}{2} \int_0^{2\pi} \cos((m+n)x) dx .$$

Für  $m \neq \pm n$  verschwinden beide Summanden.

Wegen  $m, n \geq 0$  tritt der Fall  $m+n=0$  nur im trivialen Fall  $m=n=0$  auf.

Für  $m=n$  ist  $m-n=0$  und  $m+n \neq 0$  (außer im trivialen Fall  $m=n=0$ ), daher verschwindet der zweite Summand. Der erste vereinfacht sich zu

$$\int_0^{2\pi} \sin^2(nx) = -\frac{1}{2} \int_0^{2\pi} dx = \pi .$$

□

## 4.3 Integration rationaler Funktionen

**4.3.1 Partialbruchzerlegung.** Sei  $f(x) = \frac{p(x)}{q(x)}$  eine rationale Funktion mit Grad  $p < \text{Grad } q$  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

$$\frac{A_1}{x - b}, \quad \frac{A_2}{(x - b)^2}, \quad \dots, \quad \frac{A_k}{(x - b)^k},$$

$$\frac{B_1x + C_1}{Q(x)}, \quad \frac{B_2x + C_2}{(Q(x))^2}, \quad \dots, \quad \frac{B_lx + C_l}{(Q(x))^l}$$

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:

$$f(x) = \sum_{i=1}^r \sum_{j=1}^{k_i} \frac{A_j}{(x - b_i)^j} + \sum_{i=1}^s \sum_{j=1}^{l_i} \frac{B_jx + C_j}{(Q_i)^j}$$

• Die dabei auftretenden Integrale berechnen sich wie folgt:

$$\int \frac{dx}{x \pm a} = \ln |x \pm a| + c, \quad (3.1)$$

$$\int \frac{dx}{(x \pm a)^k} = \frac{1}{1 - k} (x \pm a)^{1-k} + c, \quad k \in \mathbb{Z} \setminus \{1\}, \quad (3.2)$$

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

$$\int \frac{dx}{x^2 + px + q} = \frac{2}{\sqrt{4q - p^2}} \arctan \frac{2x + p}{\sqrt{4q - p^2}} + c, \quad (3.3)$$

$$\int \frac{ax + b}{x^2 + px + q} dx = \frac{a}{2} \ln |x^2 + px + q| + \left(b - \frac{ap}{2}\right) \int \frac{dx}{x^2 + px + q}, \quad (3.4)$$

$$\int \frac{dx}{(x^2 + px + q)^k} = \frac{2x + p}{(k-1)(4q - p^2)(x^2 + px + q)^{k-1}} + \frac{2(2k-3)}{(k-1)(4q - p^2)} \int \frac{dx}{(x^2 + px + q)^{k-1}}, \quad (3.5)$$

$$\int \frac{ax + b}{(x^2 + px + q)^k} dx = \frac{a}{2(k-1)(x^2 + px + q)^{k-1}} + \left(b - \frac{ap}{2}\right) \int \frac{dx}{(x^2 + px + q)^k}. \quad (3.6)$$

**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:

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

## 4.4 Uneigentliche Integrale

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

**4.4.1 Definition.** 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)

$$\int_1^\infty \frac{dx}{x} = \lim_{c \rightarrow \infty} \int_1^c \frac{dx}{x} = \lim_{c \rightarrow \infty} \ln c = \infty, \quad (\text{divergiert}). \quad (4.1)$$

$$\int_0^1 \frac{dx}{x} = \lim_{c \rightarrow 0^+} -\ln c = \infty, \quad (\text{divergiert}). \quad (4.2)$$

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

$$\int_1^\infty \frac{dx}{x^\alpha} = \lim_{c \rightarrow \infty} \frac{1}{\alpha-1} \left(1 - \frac{1}{c^{\alpha-1}}\right) = \begin{cases} \frac{1}{\alpha-1}, & \text{falls } \alpha > 1, \\ \infty, & \text{falls } \alpha < 1. \end{cases} \quad (4.3)$$

$$\int_0^1 \frac{dx}{x^\alpha} = \lim_{c \rightarrow 0^+} \frac{1}{\alpha-1} \left(\frac{1}{c^{\alpha-1}} - 1\right) = \begin{cases} \infty, & \text{falls } \alpha > 1 \\ \frac{1}{1-\alpha} & \text{falls } \alpha < 1. \end{cases} \quad (4.4)$$

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

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

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

**4.4.3 Definition.** 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.** Es kann passieren, dass die beiden uneigentlichen Integrale aus 4.4.4 divergieren, aber dass der Grenzwert

$$\lim_{\varepsilon \rightarrow 0^+} \left( \int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right) \quad (4.5)$$

existiert. In diesem Fall wird der Grenzwert in (4.5) **Cauchyscher Hauptwert** genannt und mit

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

bezeichnet.

## 4.5 Kurven-, Längen- und Flächenmessung

Unter einer **Kurve** versteht man eine differenzierbare Abbildung eines Intervalls  $I$  in die Ebene.

**4.5.1 Definition.** Sei der  $\mathbb{R}^2$  mit einem festen kartesischen Koordinatensystem versehen. Dann nennt man die vektorwertige Funktion

$$\vec{r} : [a, b] \rightarrow \mathbb{R}^2 : t \mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad (5.1)$$

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**.

**4.5.2 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.

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

$$\dot{\vec{r}}(t) := \lim_{h \rightarrow 0} \frac{1}{h} (\vec{r}(t+h) - \vec{r}(t)) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix}, \quad (5.2)$$

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  $90^\circ$  in positive Richtung, d.h.

$$\vec{n}(t) := \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}. \quad (5.3)$$

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

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

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

**4.5.4 Definition** Die Bogenlänge. Sei  $a = t_0 < t_1 < \dots < t_n = b$  eine äquidistante Zerlegung von  $[a, b]$ , d.h.  $t_i - t_{i-1} = \Delta t$ ,  $\forall i = 1, \dots, n$ . Die Kurve

$K$  wird durch die Sekanten, die  $(x(t_i), y(t_i))$  und  $(x(t_{i+1}), y(t_{i+1}))$  verbinden, angenähert. Die Länge dieser Approximation ist

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

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

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

**4.5.6 Satz.** 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

$$L = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt. \quad (5.4)$$

*Beweis.* Wir benutzen die Definition und zerlegen das Intervall durch äquidistante Punkte  $t_i$ . Die Länge der Sekante zwischen  $\vec{r}(t_{i-1})$  und  $\vec{r}(t_i)$  ist mit dem Mittelwertsatz der Differentialrechnung

$$\Delta s_i = \sqrt{\Delta y_i^2 + \Delta x_i^2} = \Delta t \sqrt{\dot{y}(\zeta_i)^2 + \dot{x}(\eta_i)^2}$$

für geeignete  $\eta_i, \zeta_i \in (t_i, t_{i+1})$ . Beim Grenzübergang  $\Delta t_i \rightarrow 0$  gilt dann  $\eta_i, \zeta_i \rightarrow t_i$ .

Durch Aufsummieren erhalten wir

$$P_n = \sum_i \Delta s_i.$$

Nach Definition ist  $\lim_{n \rightarrow \infty} P_n = L$ . Gleichzeitig sind die  $P_n$  nach Definition Riemansche Summen der Funktion  $\sqrt{\dot{x}^2 + \dot{y}^2}$ . Diese sind stetig, also integrierbar, und wir erhalten.

$$L = \lim_{n \rightarrow \infty} P_n = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2} dt.$$

□

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

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx. \quad (5.5)$$

**4.5.8 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  $\Delta s$  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

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

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

$$K(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}^2(t) + \dot{y}^2(t))^{\frac{3}{2}}}. \quad (5.6)$$

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

$$K(x) = \frac{f''(x)}{\left(\sqrt{1 + (f')^2(x)}\right)^3}. \quad (5.7)$$

• 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:

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

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

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

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

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

der Kurve mit Parameter  $\varphi$  heißt **Polardarstellung**, wobei der Winkel  $\varphi$  von der positiven  $x$ -Achse aus gemessen wird. Aus der Polardarstellung  $r(\varphi)$ ,  $\varphi \in [\alpha, \beta]$  erhält man folgende Parameterdarstellung mit dem Polarwinkel  $\varphi$  als Parameter

$$x = r(\varphi) \cos \varphi, \quad y = r(\varphi) \sin \varphi, \quad \varphi \in [\alpha, \beta]. \quad (5.8)$$

Aus Satz 4.5.6 erhält man für die Länge der Kurve  $r(\varphi)$ ,  $\varphi \in [\alpha, \beta]$

$$L = \int_{\alpha}^{\beta} \sqrt{r^2(\varphi) + \left(\frac{dr(\varphi)}{d\varphi}\right)^2} d\varphi.$$

• Nach Definition misst das Integral einer positiven Funktion die Fläche unter dem Graphen. Dies lässt sich auch auf Flächen mit komplizierter Berandung verallgemeinern.

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

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

*Beweis.* Wir betrachten den wesentlichen Fall  $0 < g < f$ . Dann ist die Fläche zwischen  $f$  und  $g$  die Differenz der Fläche unter dem Graphen von  $f$  und dem Graphen von  $g$ , also

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

□

**4.5.13 Satz.** Sei  $r : [\alpha, \beta] \rightarrow \mathbb{R}_0^+$  stetig,  $\alpha < \beta$ . Wir fassen  $(r(\varphi), \varphi)$  als Kurve in Polarkoordinaten auf. Dann beträgt der Inhalt der von den drei Kurven  $r = r(\varphi)$ ,  $\varphi = \alpha$ ,  $\varphi = \beta$  berandeten Sektorfläche

$$F = \frac{1}{2} \int_a^b r^2(\varphi) d\varphi$$

*Beweis.* Die Sektorfläche zwischen  $\varphi$  und  $\varphi + \Delta\varphi$  hat angenähert den Flächeninhalt eines Kreissektors mit Radius  $r(\varphi)$ , d.h.

$$\Delta F = \frac{\Delta\varphi}{2\pi} (r(\varphi))^2 \pi = \frac{1}{2} r(\varphi)^2 \Delta\varphi .$$

Wir summieren und können wieder die Summe als Riemannsumme interpretieren. Im Grenzübergang  $\Delta\varphi \rightarrow 0$  erhalten wir die Aussage. □

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

$$F = \frac{1}{2} \left| \int_a^b (x(t)y'(t) - y(t)x'(t)) dt \right| .$$

*Beweis.* Mit  $\varphi = \arctan \frac{y(t)}{x(t)}$  und Substitutionsregel folgt dies aus Satz 4.5.13. □

**4.5.15 Beispiel.** Der Inhalt der Ellipse  $x = a \cos t$ ,  $y = b \sin t$  beträgt

$$F = \frac{1}{2} \int_0^{2\pi} (ab \cos^2 t + ab \sin^2 t) dt = ab\pi .$$

**4.5.16 Satz.** Die Formel aus Satz 4.5.14 gilt auch für geschlossene überschneidungsfreie Kurven mit stückweise stetig differenzierbarer Parameterdarstellung  $x = x(t), y = y(t), t \in [a, b]$ .

**4.5.17 Redeweisen und Bezeichnungen.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige positive Funktion. Der Flächeninhalt  $F$  unter dem Graphen von  $f$  ist gegeben durch

$$F = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) \Delta x,$$

wobei wir eine äquidistante Zerlegung gewählt haben. Mit der Bezeichnung  $\Delta F_i := f(\xi_i) \Delta x$  haben wir also

$$F = \lim_{n \rightarrow \infty} \sum_{i=1}^n \Delta F_i.$$

Es ist üblich eine Differenz  $\Delta x$  mit  $dx$  zu bezeichnen, wenn man im Verlauf der Rechnung den Grenzübergang  $\Delta x \rightarrow 0$  durchführen will. Bei diesem Grenzübergang geht das Summenzeichen  $\sum$  in das Integral  $\int$  über. In Analogie dazu bezeichnen wir auch  $\Delta F$  mit  $dF$  und ersetzen für den Grenzübergang das Summenzeichen durch das Integral. Somit erhalten wir:

$$F = \int dF \quad \text{und} \quad F = \int f(x) dx.$$

Diese Überlegungen kann man verallgemeinern. Sei  $G$  eine reelle Größe, die man in kleine „Bausteine“ oder Elemente  $dG$  zerlegen kann und die „aufsummiert“ wieder  $G$  ergeben. Die Summation wird als Integral bezeichnet, und somit haben wir

$$G = \int dG. \tag{5.9}$$

Weiterhin möchte man  $dG$  durch eine von  $x$  abhängige Größe der Form

$$dG = f(x) dx \tag{5.10}$$

annähern und erhält also

$$G = \int f(x) dx.$$

**Beispiel:** Sei  $K$  eine Kurve mit der Parameterdarstellung  $x = x(t)$ ,  $y = y(t)$ ,  $t \in [a, b]$ , dann ist die Länge  $L$  gegeben durch (Satz 4.5.6)

$$L = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2} dt.$$

Andererseits haben wir

$$L = \int ds,$$

wobei  $ds$  ein Längenelement ist. Ein Vergleich beider Formeln liefert

$$ds = \sqrt{\dot{x}^2 + \dot{y}^2} dt. \quad (5.11)$$

Insbesondere erhalten wir im Falle eines Graphen  $y = f(x)$  die Formel

$$ds = \sqrt{1 + (f')^2} dx. \quad (5.12)$$

**4.5.18** Volumen von Rotationskörpern. Sei  $K$  ein Rotationskörper, der durch Rotation einer Kurve  $y = f(x)$  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:

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

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

**4.5.20 Beispiel.** Ein **Kreiskegel** entsteht durch Rotation von  $y = \frac{r}{h}x$  um  $x$ -Achse. Daher

$$\Rightarrow V = \pi \frac{r^2}{h^2} \int_0^h x^2 dx = \pi \frac{r^2}{h^2} \frac{1}{3} h^3 = \frac{1}{3} \pi r^2 h$$

**4.5.21** 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

$$M = \int dM = 2\pi \int f(x) \sqrt{1 + (f')^2} dx. \quad (5.13)$$

**4.5.22 Beispiel.** Die **Kugeloberfläche** wird beschrieben durch die Gleichung

$$x^2 + y^2 = r^2$$

d.h., sie ist die Rotationsfläche von

$$y = \sqrt{r^2 - x^2} \quad x \in [-r, r]$$

Wir erhalten

$$y' = \frac{-2x}{\sqrt{r^2 - x^2}}$$

und daher

$$\begin{aligned} M &= 2\pi \int \sqrt{r^2 - x^2} \sqrt{1 + \frac{x^2}{r^2 - x^2}} dx \\ &= 2\pi \int \sqrt{r^2 - x^2} \sqrt{\frac{r^2 - x^2 + x^2}{r^2 - x^2}} dx \\ &= 2\pi r \int_{-r}^r dx = 2\pi r x \Big|_{-r}^r = 2\pi r^2 + 2\pi r^2 \\ &= 4\pi r^2 \end{aligned}$$



# Kapitel 5

## Reihenentwicklungen

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.

### 5.1 Wiederholung und Ergänzung

- Die Konvergenz von Reihen wurde bereits in Kapitel 2.5 behandelt. Eine Reihe konvergiert **absolut**, wenn die Reihe über die Absolutbeträge konvergiert (vergleiche Definition 2.5.11).
- Das Cauchy-Produkt zweier absolut konvergenter Reihen konvergiert gegen das Produkt der Grenzfunktionen (vergleiche Satz 2.5.16).
- Für Folgen und Reihen von Funktionen gibt es verschiedene Konvergenzbegriffe. Eine Folge von Funktionen  $(f_n)_{n \geq 0}$  auf einem Intervall  $I$  konvergiert **punktweise**, wenn für jedes  $x \in I$  die Folge  $(f_n(x))_{n \geq 0}$  konvergiert (vergleiche Definition 2.6.1).
- Sie konvergiert **gleichmäßig**, wenn es für jedes  $\varepsilon > 0$  einen Index  $n_0$  gibt, so dass für alle  $x \in I$  und  $n \geq n_0$  gilt

$$|f(x) - f_n(x)| < \varepsilon$$

(vergleiche Definition 2.6.2).

- $(f_n)_{n \geq 0}$  gleichmäßig konvergent. Sind alle  $f_n$  stetig, so ist die Grenzfunktion stetig (vergleiche Satz 3.2.11).
- Sei  $(f_n)_{n \geq 0}$  punktweise konvergent, alle  $f_n$  stetig differenzierbar und die Folge  $(f'_n)_{n \geq 0}$  gleichmäßig konvergent. Dann ist  $f$  stetig differenzierbar und  $f' = \lim_{n \rightarrow \infty} f'_n$  (vergleiche Satz 3.2.12).
- **Potenzreihen**  $\sum_{n=0}^{\infty} a_n x^n$  sind ein besonders wichtiges Beispiel. Sie haben einen Konvergenzradius  $R$ , in dessen Inneren sie absolut konvergieren, auf jeden abgeschlossenen Teilintervall sogar gleichmäßig (vergleiche Satz 2.6.6). Dort sind sie beliebig oft differenzierbar, insbesondere auch stetig.

**5.1.1 Satz.** *Konvergiert die Folge stetiger Funktionen  $f_n$ ,  $n \geq 0$ , auf dem Intervall  $I$  gleichmäßig gegen  $f : I \rightarrow \mathbb{R}$ , dann gilt für alle  $a, b \in I$*

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

*Beweis.* Sei  $\varepsilon > 0$ . Nach Voraussetzung existiert ein  $n_0$ , so dass für alle  $n \geq n_0$  und alle  $x \in I$  gilt  $|f_n(x) - f(x)| < \frac{\varepsilon}{b-a}$ . Hieraus folgt mit Satz 4.1.7

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| &\leq \int_a^b |f(x) - f_n(x)| dx \\ &\leq (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon \end{aligned}$$

□

**5.1.2 Beispiel.** Sei

$$h_n(x) = \begin{cases} n^2 x, & \text{falls } x \in [0, \frac{1}{n}], \\ 2n - n^2 x, & \text{falls } x \in [\frac{1}{n}, \frac{2}{n}], \\ 0, & \text{falls } x \in [\frac{2}{n}, 1]. \end{cases}$$

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 111

Dann konvergiert  $(h_n)_{n \geq 0}$  punktweise gegen 0, jedoch nicht gleichmäßig. Es gilt

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

Alle Sätze übertragen sich auf Reihen.

**5.1.3 Satz.** *Konvergiert die Reihe  $\sum_{k=0}^{\infty} f_k(x)$  stetiger Funktionen  $f_k$  auf  $I$  gleichmäßig gegen  $f$ , dann ist*

$$f(x) = \sum_{k=0}^{\infty} f_k(x)$$

stetig und für alle  $a, b \in I$  gilt:

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

**5.1.4 Folgerung.** *Sei  $\sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R$ . Dann gilt für jedes Teilintervall  $[a, b] \subset (-R, R)$*

$$\int_a^b \left( \sum_{k=0}^{\infty} a_k x^k \right) dx = \sum_{k=1}^{\infty} \int_a^b \frac{a_{k-1}}{k} x^k dx.$$

*Beweis.* Wir wenden den Satz an auf  $f_k = a_k x^k$  mit Integral  $a_k \frac{1}{k+1} x^{k+1}$ .  $\square$

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung

**5.2.1 Definition.** *Sei  $\sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R > 0$  und Summe  $f(x)$  für  $|x| < R$ . Man sagt,  $f$  wird auf  $(-R, R)$  durch die Potenzreihe  $\sum_{k=0}^{\infty} a_k x^k$  dargestellt.*

**5.2.2 Satz.** Eine durch eine Potenzreihe dargestellte Funktion  $f$  ist im offenen Konvergenzintervall  $(-R, R)$ ,  $R > 0$ , beliebig oft differenzierbar. Die Ableitung erhält man durch gliedweise Differentiation:

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} k a_k x^{k-1}, \\ f''(x) &= \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}, \end{aligned} \tag{2.1}$$

etc. Die abgeleiteten Reihen haben alle den Konvergenzradius  $R$ .

*Beweis.* Das ist Satz 3.2.14. Der entscheidende Punkt ist die Berechnung des Konvergenzradius von Wir behandeln den Fall, dass er sich der Konvergenzradius von  $\sum_{k=0}^{\infty} a_k x^k$  mit der Formel aus Satz 2.6.7 berechnen lässt, also

$$R^{-1} = \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

Dann ist der Konvergenzradius von  $\sum_{k=1}^{\infty} k a_k x^{k-1}$

$$\lim_{k \rightarrow \infty} \frac{(k+1)a_{k+1}}{k a_k} = R^{-1}$$

nach den Rechenregeln für Grenzwerte. □

**Beispiel:**

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

**5.2.3 Satz.** Für alle  $a, b$  aus dem offenen Konvergenzintervall  $(-R, R)$  der Potenzreihe  $f(x) = \sum_{k=0}^{\infty} a_k x^k$  gilt:

$$\int_a^b f(x) dx = \sum_{k=0}^{\infty} \int_a^b a_k x^k dx = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (b^{k+1} - a^{k+1}).$$

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 113

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  $(-R, R)$ , deren Konvergenzradius  $R$  ist.

Beweis. Satz 5.1.4

□

**5.2.4 Beispiel.** Durch Integrieren der geometrischen Reihe erhalten wir

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k} \quad |x| < 1.$$

Die Exponentialfunktion wird durch die Reihe  $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$  dargestellt.

**5.2.5 Definition.** Eine unendliche Reihe der Form

$$\sum_{k=0}^{\infty} a_k (x-a)^k \tag{2.2}$$

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 (2.2) ist der Konvergenzradius der entsprechenden Reihe mit Zentrum  $0$ .

**5.2.6 Beispiel.** Es gilt  $e^x = e^a e^{x-a}$  und daher

$$\exp(x) = \sum_{k=0}^{\infty} \frac{e^a}{k!} (x-a)^k$$

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

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

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

**5.2.8** Koeffizientenvergleich. Sei  $f$  auf Intervall  $(a - R, a + R)$  als Potenzreihe  $f(x) = \sum_{k=0}^{\infty} a_k (x - a)^k$  dargestellt. Nach Satz 5.2.2 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  $f^{(n)}(a) = n!a_n$ .

**5.2.9 Satz** (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

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

• Viel schwieriger ist die Frage, ob eine Funktion durch eine Potenzreihe dargestellt werden kann. Etwas weniger ehrgeizig versuchen wir, eine hinreichend oft differenzierbare Funktion durch Polynome zu approximieren. Die Approximation durch ein lineares Polynom hatte uns gerade auf den Ableitungsbegriff geführt.

**5.2.10 Satz** (Taylorformel). 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

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

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 115

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

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

hat.

*Beweis.* Nach Definition von Stammfunktionen gilt

$$f(x) = f(a) + \int_a^x f'(t) dt$$

Dies ist die behauptete Formel im Fall  $n = 0$ .

Wir führen nun eine partielle Integration mit  $u = t - x$ ,  $v = f'$  aus und erhalten

$$\begin{aligned} f(x) &= f(a) + \int_a^x u' f' dt \\ &= f(a) + (t-x) f'(x) \Big|_a^x - \int_a^x (t-x) f''(t) dt \\ &= f(a) + f'(a)(x-a) + \int_a^x (x-t) f''(t) dt \end{aligned}$$

Dies ist die Formel für  $n = 1$ . Wiederholung des Argumentes liefert die Formel für jedes  $n$ .

Die zweite Darstellung des Restgliedes  $R_{n+1}$  erhalten wir durch den Mittelwertsatz der Integralrechnung 4.1.8:

$$\begin{aligned} \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt &= \frac{1}{n!} f^{(n+1)}(\xi) \int_a^x (x-t)^n dt \\ &= \frac{1}{(n+1)!} f^{(n+1)}(\xi) (x-a)^{n+1} \end{aligned}$$

wobei  $\xi$  ein geeignetes Element von  $[a, x]$  ist. □

- Das Taylorpolynom  $T_n(x, a)$  ist in Umgebung von  $a$  sehr gute Approximation

$$\begin{aligned} T_n(a, a) &= f(a) && \text{gleiche Funktionswert} \\ T'_n(a, a) &= f'(a) && \text{gleiche Anstieg} \\ T''_n(a, a) &= f''(a) && \text{gleiche Krümmung} \\ &&& \text{usw.} \end{aligned}$$

- Der Fehler

$$|f(x) - T_n(x, a)| = |R_{n+1}(x, a)|$$

kann genau abgeschätzt werden, z.B. auf dem Intervall  $[a - R, a + R]$  durch

$$\sup_{[a-R, a+R]} |f^{(n+1)}(x)| \frac{R^{n+1}}{(n+1)!}$$

**5.2.11 Beispiel.** Sei  $f(x) = \sin x$ ,  $a = 0$ . Dann gilt

$$f'(x) = \cos x, f^{(2)}(x) = -\sin x, f^{(3)}(x) = -\cos x, f^{(4)}(x) = \sin x$$

und daher

$$f'(0) = 0, f^{(2)}(0) = 1, f^{(3)}(0) = 0, f^{(4)}(0) = -1.$$

Es gilt also nach der Taylorformel

$$\sin x = 0 + x + 0x^2 - \frac{1}{3!}x^3 + \frac{\sin \xi}{4!}x^4$$

mit der Abschätzung

$$|R_4(x, 0)| \leq \frac{1}{4!}|x|^4.$$

Allgemeiner:

$$\sin x = \sum_{k=0}^n (-1)^k \frac{1}{(2k+1)!} x^{2k+1} + R_{2n+2}$$

wobei

$$|R_{2n+2}| \leq \frac{1}{(2n+2)!} x^{2n+2}.$$

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 117

**5.2.12 Satz.** *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.

*Beweis.* Unter unseren Voraussetzungen besagt die Taylorformel

$$f(x) - f(a) = \frac{f^{(n)}(\xi)}{n!} (x - a)^n$$

für ein  $\xi \in [a, x]$ .

Ist  $n$  ungerade, so hat  $f(x) - f(a)$  für  $x > a$ ,  $x < a$  verschiedene Vorzeichen, und  $f$  hat in  $a$  kein Extremum.

Ist  $n$  gerade, so hat  $f(x) - f(a)$  für  $x > a$ ,  $x < a$  gleiche Vorzeichen, und  $f$  hat in  $a$  ein Extremum. Wir betrachten  $x > a$ ,  $f^{(n)}(a) < 0$ . Dann ist  $f^{(n)}(\xi) < 0$  für  $x$  (und damit  $\xi$ ) nahe bei  $a$ . Es folgt  $f(x) - f(a) < 0$ . Damit hat  $f$  in  $a$  ein Maximum. Der andere Fall folgt analog.  $\square$

**5.2.13 Definition.** *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**.

- Wenn  $f$  durch eine Potenzreihe darstellbar ist, dann lässt es sich wegen Satz 5.2.9 als Taylorreihe entwickeln.
- Auch wenn  $f$  unendlich oft differenzierbar ist, so ist es nicht immer/oftnicht/meist nicht als Taylorreihe entwickelbar!

**5.2.14 Beispiel.** Wir betrachten  $f(x) = \exp(-\frac{1}{x^2})$ . Die Funktion lässt sich durch 0 zu einer stetigen Funktion auf  $\mathbb{R}$  fortsetzen. Die Ableitungen von  $f$  haben alle die Form  $P(x)f(x)$ , wobei  $P$  eine rationale Funktion ist, z.B.

$$f'(x) = \frac{2}{x^3}f, f'' = \frac{-6}{x^4}f + \frac{2}{x^3}f' .$$

Wegen  $\lim_{x \rightarrow 0} P(x)f(x) = 0$  ist  $f$  dann auch in 0 unendlich oft differenzierbar mit  $f^{(n)}(0) = 0$ . Es ist also

$$T_f(x, 0) = 0 \neq f .$$

**5.2.15 Satz.** 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

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

*Beweis.* Nach der Taylorformel 5.2.10 gilt

$$f(x) - T_n(x, a) = R_{n+1}(x, a)$$

wobei  $T_n(x, a)$  die  $n$ -te Partialsumme von  $T_\varphi(x, a)$  ist d.h.

$$f(x) = \lim_{n \rightarrow \infty} T_n(x, a) \quad \Leftrightarrow \quad \lim_{n \rightarrow \infty} R_{n+1}(x, a) = 0$$

Sei nun

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

Es folgt

$$|R_n(x, a)| = \left| \frac{f^{(n)}(\xi)}{n!} (x - a)^n \right| \leq A \frac{B^n}{n!} \cdot |x - a|^n .$$

Da die Reihe

$$\sum \frac{B^n |x - a|^n}{n!}$$

konvergiert, ist  $A \frac{B^n}{n!} \cdot |x - a|^n$  eine Nullfolge. □

**5.2.16 Methoden der Reihenentwicklung.**

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 119

### 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 5.2.2 und 5.2.3 ermöglichen es, durch gliedweise Differentiation und Integration bekannter Reihen neue Reihenentwicklung zu erhalten (siehe Satz 5.2.17).

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:

$$\begin{aligned} \cosh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, & x \in \mathbb{R}, \\ \sinh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}, & x \in \mathbb{R}. \end{aligned} \tag{2.3}$$

### 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 2.5.16) 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

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

Die Reihenentwicklung von  $e^{e^x}$  erhalten wir mit  $f(x) = g(x) = \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ . Es ist

$$(g(x))^k = (e^x)^k = e^{kx} = \sum_{n=0}^{\infty} \frac{k^n x^n}{n!}$$

und damit

$$\begin{aligned} e^{e^x} &= \sum_{k=0}^{\infty} \frac{1}{k!} \left( \sum_{n=0}^{\infty} \frac{k^n x^n}{n!} \right) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^{\infty} \frac{k^n}{k!} \right) \frac{x^n}{n!} \\ &= e + xe + \frac{x^2}{2} \left( \sum_{k=0}^{\infty} \frac{k^2}{k!} \right) + \dots \end{aligned}$$

**5.2.17 Satz.** Wir haben folgende Potenzreihendarstellungen:

$$\begin{aligned} a) \quad e^x &= \sum_{k=0}^{\infty} \frac{1}{k!} x^k, & x \in \mathbb{R}, \\ b) \quad \sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}, & x \in \mathbb{R}, \\ c) \quad \cos x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, & x \in \mathbb{R}, \\ d) \quad \sinh x &= \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} x^{2k+1}, & x \in \mathbb{R}, \\ e) \quad \cosh x &= \sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{2k}, & x \in \mathbb{R}, \\ f) \quad \ln(1+x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}, & |x| < 1, \\ g) \quad \arctan x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}, & |x| < 1. \end{aligned}$$

*Beweis.* a) war unsere Definition.

b) war Beispiel 5.2.11, c) durch differenzieren.

d) e) (siehe oben) aus der Reihenentwicklung von  $e^x$ .

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 121

f) folgt durch die Substitution von  $-x$  in der geometrischen Reihe, gefolgt von Integration.

g) Wir benutzen

$$\arctan x = \int_{y=0}^x \frac{dy}{1+y^2}$$

Der Integrand ist die geometrische Reihe

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

die termweise integriert wird. □

**5.2.18 Satz.** 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

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

*Beweis.* Sei  $g(x)$  die rechte Seite. Wir berechnen den Konvergenzradius mit der Quotientenformel  $R^{-1} = \lim |a_{k+1}/a_k|$  und erhalten wegen

$$\lim_{k \rightarrow \infty} \frac{\alpha(\alpha-1) \cdots (\alpha-(k+1)+1)k!}{(k+1)! \alpha(\alpha-1) \cdots (\alpha-k+1)} = \lim_{k \rightarrow \infty} \frac{\alpha-k}{k+1} \rightarrow -1$$

den Konvergenzradius 1.

Sei nun  $-1 < x < 1$ . Es folgt

$$g(x) = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!} x^3 + \dots$$

$$g'(x) = \alpha + \frac{\alpha(\alpha-1)}{1 \cdot 2!} 2x + \frac{\alpha(\alpha-1)(\alpha-2)}{2!} x^2 + \dots$$

$$xg'(x) = \alpha x + \frac{\alpha(\alpha-1)}{1!} x^2 + \dots$$

$$g'(x)(1+x) = \alpha + \alpha x(\alpha-1+1) + \alpha(\alpha-1) \left(\frac{\alpha-2}{2} + 1\right) x^2 + \dots$$

$$= \alpha \left( 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots \right)$$

$$= \alpha g(x)$$

Hieraus folgt dann für

$$\begin{aligned} h(x) &:= \frac{g(x)}{(1+x)^\alpha} \\ h'(x) &= \frac{g'(x)(1+x)^\alpha - g(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \\ &= \frac{(1+x)^{\alpha-1} [g'(x)(1+x) - g(x)\alpha]}{(1+x)^{2\alpha}} = 0 \end{aligned}$$

Die Funktion  $h$  ist also konstant. Im Punkt 0 gilt  $h(0) = g(0) = 1$ , also

$$1 = \frac{g(x)}{(1+x)^\alpha}.$$

□

• Die Formel aus Satz 5.2.18 enthält als Spezialfälle:

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

$$\begin{aligned} \binom{\alpha}{k} &= \binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!} = 0 \quad \text{falls } n < k \\ \Rightarrow (1+x)^n &= \sum_{k=0}^n \binom{n}{k} x^k \quad (\text{klassische binomische Formel}) \end{aligned}$$

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

$$\begin{aligned} \binom{\frac{1}{2}}{0} &= 1, \quad \binom{\frac{1}{2}}{1} = \frac{1}{2}, \quad \binom{\frac{1}{2}}{2} = -\frac{1}{8}, \\ \binom{\frac{1}{2}}{k} &= (-1)^{k-1} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{2 \cdot 4 \cdots (2k)}, \quad k \geq 2 \\ \Rightarrow \sqrt{1+x} &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 \pm \cdots, \quad (2.4) \end{aligned}$$

falls  $|x| < 1$ .

## 5.2 Darstellbarkeit durch Potenzreihen und Taylorentwicklung 123

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

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

$$\Rightarrow \frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 \pm \dots, \quad (2.5)$$

falls  $|x| < 1$ .

- Die Taylorentwicklung hat viele verschiedenen Anwendungen.

**5.2.19** Grenzwertberechnung. *Durch Einsetzen der Reihenentwicklung der entsprechenden Funktion erhält man eine rationale Funktion in  $x$  deren Grenzwert einfacher zu berechnen ist. Diese Methode ist eine Alternative zur Regel von de l'Hospital, wenn man die Reihen kennt.*

**5.2.20** Beispiel. Wir bestimmen

$$\lim_{x \rightarrow 0} \frac{x \ln(1-x)}{\sin^2 x}$$

Nach Satz 5.2.17 gilt

$$\begin{aligned} \ln(1-x) &= \sum_{k=0}^{\infty} \frac{(-1)^k (-1)^{k+1}}{k+1} x^{k+1} \\ &= - \left( x + \frac{x^2}{2} + \frac{x^3}{3} + \dots \right) \\ \sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \\ &= x - \frac{x^3}{3!} + \dots \\ \frac{x \ln(1-x)}{\sin^2 x} &= \frac{-x \left( x + \frac{x^2}{2} + \frac{x^3}{3} + \dots \right)}{\left( x - \frac{x^3}{3!} + \dots \right) \left( x - \frac{x^3}{3!} + \dots \right)} \\ &= \frac{- \left( 1 + \frac{x}{2} + \frac{x^2}{3} \right)}{1 - \frac{2x^2}{3!} + \dots} \rightarrow -1 \quad x \rightarrow 0 \end{aligned}$$

**5.2.21** Näherungsformeln. *In komplizierten, funktionalen Zusammenhängen kann man durch Einsetzen der Reihenentwicklung die „dominanten“ Terme bestimmen.*

**5.2.22** Berechnung nicht elementar integrierbarer Integrale. *Einige Integrale besitzen keine in geschlossener Form darstellbare Stammfunktion. Durch die Reihendarstellung des Integranden erhält man oft eine Darstellung des Integrals als unendliche Reihe.*

**5.2.23** Beispiel.

$$\begin{aligned}
 F(\varphi, k) &= \int_{t=0}^{\varphi} \frac{dt}{\sqrt{1 - k^2 \sin^2 t}} \quad 0 \leq k^2 < 1 \\
 \frac{1}{\sqrt{1 - x^2}} &= 1 + \frac{x^2}{2} + \frac{1 \cdot 3}{2 \cdot 4} x^4 + \dots \\
 \int_{t=0}^{\varphi} \frac{1}{\sqrt{1 - k^2 \sin^2 t}} dt &= \int_{t=0}^{\varphi} 1 + \frac{\sin^2 t k^2}{2} + \frac{1 \cdot 3}{2 \cdot 4} k^4 \sin^4 t + \dots \\
 &= \varphi + \frac{k^2}{2} \int_{t=0}^{\varphi} \sin^2 t dt + \frac{3k^4}{8} \int_{t=0}^{\varphi} \sin^4 t dt + \dots
 \end{aligned}$$

Für konkrete Werte  $k, \varphi$  kann nun  $F(\varphi, k)$  beliebig genau berechnet werden.

## 5.3 Fourier-Reihen

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.

Die einfachsten  $P$  periodischen Funktionen sind  $\cos(2\pi/Pnx)$  und  $\sin(2\pi/Pnx)$  mit  $n \in \mathbb{N}_0$ .

**5.3.1 Definition.**  $P > 0$ . 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  $f$ .

**5.3.2 Satz.** Sei  $f$   $P$ -periodisch und beliebig oft differenzierbar. Dann konvergiert die Fourier-Reihe zu  $f$  gleichmäßig gegen  $f$ . Die Darstellung als Fourierreihe ist eindeutig.

*Beweis.* Wir behandeln nur die Eindeutigkeit. Sei  $P = 2\pi$ ,

$$f(x) = \sum_{k=0}^{\infty} a_k \cos(kx) + \sum_{k=1}^{\infty} b_k \sin(kx) .$$

Dann gilt für die Fourier-Koeffizienten nach Satz 5.1.3 (Vertausch von Summe

und Integral) und den Orthogonalitätsrelationen 4.2.8

$$\begin{aligned} \frac{1}{\pi} \int_0^{2\pi} f(x) dx &= \frac{1}{\pi} \int_0^{2\pi} \left( \sum_{k=0}^{\infty} a_k \cos(kx) + \sum_{k=1}^{\infty} b_k \sin(kx) \right) \sin(nx) dx \\ &= \frac{1}{\pi} \sum_{k=0}^{\infty} a_k \int_0^{2\pi} \cos(kx) \sin(nx) + \frac{1}{\pi} \sum_{k=1}^{\infty} b_k \int_0^{2\pi} \sin(kx) \sin(nx) \\ &= \frac{1}{\pi} \pi a_n = a_n . \end{aligned}$$

Wie in Fall der Taylorreihen folgt hieraus die Eindeutigkeit. □

**5.3.3.** *Die Voraussetzung kann abgeschwächt werden: Es genügt, dass  $I$  in Teilintervalle zerlegt werden kann, auf denen die Funktion beschränkt und beliebig oft differenzierbar ist. In den Unstetigkeitsstellen konvergiert die Fourierreihe (punktweise Konvergenz) dann gegen das arithmetische Mittel des links- und rechtsseitigen Grenzwertes.*

**5.3.4 Beispiel.** Sei  $f$  periodisch mit Periode  $2\pi$  gegeben durch

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

$f$  ist ungerade. Dann verschwinden alle  $a_n$ .

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

Die Fourierreihe ist also

$$f(x) = \sum_{n=0}^{\infty} \frac{4}{\pi(2n+1)} \sin((2n+1)x) .$$

**5.3.5 Beispiel.** Wir betrachten die  $2\pi$ -periodische Sägezahnfunktion

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

und erhalten die Fourierreihe

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

# Kapitel 6

## Gewöhnliche Differentialgleichungen

**Literatur:** Meyberg, Vachenaer: Höhere Mathematik II, Kapitel 9.

### 6.1 Allgemeines

Viele physikalische Zusammenhänge stellen einen Zusammenhang her zwischen einer Größe und ihrer Veränderung. Wir nennen dies eine Differentialgleichung. **Gewöhnliche Differentialgleichungen** hängen von einer reellen Variablen ab. Differentialgleichungen in mehreren Variablen heißen **partielle Differentialgleichungen** (da in ihnen partielle Ableitungen vorkommen).

**6.1.1 Definition.** Sei  $n \in \mathbb{N}$ ,  $D \subset \mathbb{R}^{n+2}$ ,  $F : D \rightarrow \mathbb{R}$  eine Funktion.

1. Ein Bestimmungsgleichung für  $y = y(x)$  der Form

$$F(x, y, y', \dots, y^{(n)}) = 0 \quad (1.1)$$

in der Variablen  $x$  und der gesuchten Funktion  $y$  und deren Ableitungen bis zur Ordnung  $n$  auftreten, heißt **gewöhnliche Differentialgleichung  $n$ -ter Ordnung**.

2. Eine  $n$ -mal differenzierbare Funktion  $y : I \rightarrow \mathbb{R}$  heißt **Lösung** von (1.1), wenn für alle  $x \in I$  gilt

$$\begin{aligned} (x, y(x), \dots, y^{(n)}(x)) &\in D \\ F(x, y(x), \dots, y^{(n)}(x)) &= 0 \end{aligned}$$

**6.1.2 Beispiel.**  $y' - xy^2 = 0$  ist eine Differentialgleichung 1. Ordnung. Es gilt  $F(x_1, x_2, x_3) = x_3^2 - x_1x_2^2$  auf  $D = \mathbb{R}^3$ . Meist schreiben wir

$$y' = xy^2 .$$

Eine Lösung auf  $I = \mathbb{R}$  ist

$$y(x) = \frac{-2}{1+x^2}$$

denn

$$y'(x) = (-2)(-1)(1+x^2)^{-2}(2x) = x \frac{4}{(1+x^2)^2} = xy^2 .$$

**6.1.3 Beispiel.**

$$y' = y$$

hat die Lösung  $y(x) = c \exp(x)$  für jedes  $c \in \mathbb{R}$ .

**6.1.4 Beispiel.** Sei  $f : I \rightarrow \mathbb{R}$  eine Funktion. Die Differentialgleichung

$$y' = f$$

hat als Lösung die Stammfunktion  $F$  von  $f$ . Jede weitere Lösung hat dann die Form  $F + c$  für  $c \in \mathbb{R}$ .

$$y'' = f$$

hat als allgemeine Lösung  $\int F + cx + d$ , wobei  $c, d \in \mathbb{R}$  beliebig.

Wie wir in den Beispielen sehen, legt die Differentialgleichung die Funktion nicht eindeutig fest. Für eine Gleichung der Ordnung  $n$  benötigen wir im allgemeinen  $n$  Bedingungen.

**6.1.5 Definition.** 1. Ein **Anfangswertproblem** in  $x_0$  ist eine Differentialgleichung

$$y^{(n)} = F(x, y, \dots, y^{(n-1)})$$

zusammen mit der Vorgabe

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$$

wobei  $(x_0, y_0, \dots, y_{n-1})$  im Definitionsbereich von  $F$  liegt.

2. Das Anfangswertproblem heißt **lokal lösbar**, wenn es ein  $\varepsilon > 0$  gibt, so dass auf  $(x_0 - \varepsilon, x_0 + \varepsilon)$  eine Lösung  $y = y(x)$  der DGL existiert mit den vorgegebenen Anfangsbedingungen existiert. Dieses  $y(x)$  heißt **lokale Lösung**.

**6.1.6 Beispiel.** Das Anfangswertproblem

$$y' = 1 + y^2, x(0) = 0$$

wird von  $y = \tan x$  gelöst, da  $\tan' = \frac{1}{\cos^2}$ . Die Lösung ist nur  $-\frac{\pi}{2} < x < \frac{\pi}{2}$  definiert, obwohl die DGL auf ganz  $\mathbb{R}$  erklärt ist.

• Ein Anfangswertproblem heißt **sachgemäß gestellt**, wenn folgende Eigenschaften gesichert sind:

1. Die Existenz der lokalen Lösung
2. Die Eindeutigkeit der lokalen Lösung
3. Die stetige Abhängigkeit der lokalen Lösung von den Anfangswerten

(Die letzte Bedingung stellt sicher, dass das Ergebnis physikalisch relevant ist, d.h. stabil unter kleinen Störungen der Anfangswerte, die immer nur näherungsweise bekannt sind.) Für ein solches Problem stellen sich sofort zusätzliche Probleme:

4. die globale Existenz
5. explizite Bestimmung der Lösung

**6.1.7 Beispiel.** Die DGL

$$y' = 3\sqrt[3]{y^2}$$

hat als Lösungen die kubischen Parabeln  $y = (x - c)^3$ , aber zusätzlich noch die Funktion  $y = 0$ . Damit gehen durch jeden Punkt der  $x$ -Achse zwei Lösungen. Tatsächlich gibt es weitere, die durch Einkleben eines Streifens  $y = 0$  zwischen 2 Parabeläste entstehen.

## 6.2 Lineare DGLs 1. Ordnung

Wir betrachten DGLs der Form

$$y' + a(x)y = f(x) \tag{2.1}$$

mit auf einem Intervall erklärten Funktionen  $a, f$ . Die rechte Seite der DGL heißt **Störfunktion**. Die DGL heißt **homogen**, falls  $f = 0$ , andernfalls **inhomogen**. Die (2.1) zugeordnete **homogene DGL** ist

$$y' + a(x)y = 0. \tag{2.2}$$

**6.2.1 Satz.** Sei  $a : I \rightarrow \mathbb{R}$  stetig. Dann ist die vollständige allgemeine Lösung der homogenen DGL (2.2) gegeben durch

$$y(x) = ce^{-A(x)} \quad c \in \mathbb{R} ,$$

wobei  $A$  eine Stammfunktion von  $a$  ist.

*Beweis.* Sei  $y(x)$  wie im Satz. Dann gilt mit Kettenregel und dem Hauptsatz der Differential- und Integralrechnung.

$$y' = ce^{-A(x)}(-A') = ce^{-A(x)}(-a) = -ay .$$

Die Herleitung der Eindeutigkeit beruht auf dem Verfahren der **Trennung der Variablen**:

$$\begin{aligned} \frac{dy}{dx} = -a(x)y &\Leftrightarrow \frac{dy}{y} = -a(x)dx \\ &\Rightarrow \int \frac{dy}{y} = \int -a(x)dx \\ &\Rightarrow \ln(y) = -A(x) + c \\ &\Rightarrow y = \exp(-A(x) + c) = e^c e^{-A(x)} \end{aligned}$$

□

**6.2.2 Beispiel.** Sei  $a(t) = a$  konstant. Dann hat das Anfangswertproblem

$$\dot{x} = ax, x(t_0) = x_0$$

die eindeutige Lösung

$$x(t) = x_0 e^{a(t-t_0)} .$$

Diese Differentialgleichung tritt häufig auf, etwa bei radioaktiven Zerfall oder beim Wärmetransport.

**6.2.3 Satz.** Seien  $a, f : I \rightarrow \mathbb{R}$  stetig. Dann hat die inhomogene lineare DGL (2.1) die vollständige allgemeine Lösung

$$y(x) = e^{-A(x)} \left( \int_{x_0}^x e^{A(\xi)} f(\xi) d\xi + c \right)$$

mit  $c \in \mathbb{R}$ , wobei  $A(x) = \int_{x_0}^x a(\xi) d\xi$  Stammfunktion von  $A$  und  $x_0 \in I$  fest.

*Beweis.* Das angewendete Verfahren heißt **Variation der Konstanten**. Sei  $y_h$  eine Lösung der zugehörigen homogenen DGL. Wir machen den Ansatz

$$y(x) = c(x)y_h(x) .$$

Eingesetzt in (2.2) ergibt

$$\begin{aligned} f(x) &= y' + a(x)y(x) \\ &= c'(x)y_h(x) + c(x)y_h'(x) + a(x)c(x)y_h(x) \\ &= c'(x)y_h(x) . \end{aligned}$$

da  $y_h$  die homogene Gleichung löst. Es folgt

$$c'(x) = \frac{f(x)}{y_h(x)} \Rightarrow c(x) = \int_{x_0}^x \frac{f(\xi)}{y_h(\xi)} d\xi + c .$$

Nun setzen wir  $y_h(x) = \exp(-A(x))$  ein und erhalten die behauptete Formel.  $\square$

- Für jeden beliebigen Anfangswert  $y_0 = y(x_0)$  ist das Anfangswertproblem zur inhomogenen linearen DGL 1. Ordnung (2.1) eindeutig und global auf  $I$  lösbar.
- Die vollständige allgemeine Lösung der inhomogenen DGL hat die Form

$$y(x) = y_p(x) + y_h(x)$$

wobei  $y_h(x)$  die allgemeine Lösung der zugehörigen homogenen DGL ist und  $y_p$  eine Lösung der inhomogenen.

*Beweis.* Die Differenz zweier Lösungen der inhomogenen DGL erfüllt die homogene.  $\square$

- **Superpositionsprinzip** Wir betrachten Lösungen von

$$y_1' + a(x)y_1 - 1 = f_1(x), y_2' + a(x)y_2 = f_2(x)$$

Für jedes  $\alpha, \beta \in \mathbb{R}$  ist dann  $y = \alpha y_1 + \beta y_2$  eine Lösung von

$$y' + a(x)y = \alpha f_1 + \beta f_2 .$$

**6.2.4 Beispiel.** Wir betrachten

$$y' + \frac{1}{x}y = x^3 .$$

Es gilt  $A(x) = \ln x$ , und  $y_h(x) = ce^{-\ln x} = c\frac{1}{x}$  löst die zugehörige homogene DGL. Der Ansatz (Variation der Konstanten)

$$y = c(x)\frac{1}{x}$$

führt auf

$$\begin{aligned} x^3 &= c'(x)\frac{1}{x} - c(x)\frac{1}{x^2} + \frac{1}{x}c(x)\frac{1}{x} \Rightarrow \\ c'(x) &= x^4 \Rightarrow c(x) = \frac{1}{5}x^5 + C \\ y &= \frac{1}{5}x^4 + \frac{C}{x} \end{aligned}$$

• **Integration durch Substitution** Durch Übergang zu neuen Variablen wird die gegebene DGL in eine bekannte Form gebracht.

**6.2.5 Beispiel. Die Bernoulli-Differentialgleichung**

$$y' + a(x)y = b(x)y^\alpha$$

$\alpha \in \mathbb{R}$ ,  $\alpha \neq 0, 1$  und stetigen Funktionen  $a, b$ . Wir substituieren

$$\eta(x) = y(x)^{1-\alpha}$$

überführt die DGL in die Form

$$\eta' + (1 - \alpha)a(x)\eta = (1 - \alpha)b(x) .$$

*Beweis.* Wir setzen  $\eta' = (1 - \alpha)y^{-\alpha}y'$  und  $\eta$  ein und erhalten

$$(1 - \alpha)y^{-\alpha}y' + (1 - \alpha)a(x)y(x)^{1-\alpha} = (1 - \alpha)b(x)$$

Dies unterscheidet sich von der ursprünglichen Gleichung um den Faktor  $(1 - \alpha)y^{-\alpha}$ .  $\square$

Die lineare Gleichung für  $\eta$  ist lösbar. Durch die Rücksubstitution

$$y(x) = \eta(x)^{\frac{1}{1-\alpha}}$$

wird die ursprüngliche Gleichung gelöst.

Die Differentialgleichung tritt auf als Bewegungsgleichung eines Pendels mit Luftreibung.

## 6.3 Existenz und Eindeutigkeit der Lösung

Wir betrachten DGLs 1. Ordnung

$$y' = f(x, y)$$

mit  $f : G \rightarrow \mathbb{R}^2$ , wobei  $G \subset \mathbb{R}^2$  ein offenes Rechteck  $I \times J$ .

**6.3.1 Satz.** *Sei  $G = I \times J \subset \mathbb{R}^2$  ein offenes Rechteck,  $f : G \rightarrow \mathbb{R}$  stetig differenzierbar in jeder der beiden Variablen. Dann gibt es zu jedem  $(x_0, y_0) \in G$  für das Anfangswertproblem*

$$y' = f(x, y), y(x_0) = y_0$$

eine eindeutige Lösung  $y = y(x)$ , die sich beidseitig bis an den Rand von  $G$  erstreckt.

- Die Voraussetzung ist viel zu stark. Eine bessere Formulierung benötigt aber die Sprache der mehrdimensionalen Analysis, die uns noch nicht zur Verfügung steht.  $G$  kann ersetzt werden durch eine offene, zusammenhängende Teilmenge des  $\mathbb{R}^2$ . (Eine Teilmenge heißt zusammenhängend, wenn je zwei Punkte durch einen Weg verbunden werden können.)

**6.3.2 Beispiel.** Wir gehen zurück zu Beispiel 6.1.7

$$y' = 2\sqrt[3]{y^2}$$

d.h.  $f(x, y) = \sqrt[3]{y^2}$ . Die Ableitung nach  $y$  ist

$$\frac{df}{dy} = \frac{2}{3}y^{-\frac{1}{3}}$$

und in  $(x, y) = (x, 0)$  nicht wohldefiniert. Die Voraussetzung des Satzes ist nicht erfüllt.

Der Existenz- und Eindeutigkeitssatz wird oft als Satz von Picard-Lindelöf zitiert, deren Beweis eine explizite Lösungsmethode liefert.

- **Die Picard-Iteration** Wir setzen  $y_0(x) = y_0$  und bilden iterativ für  $n \in \mathbb{N}$

$$y_n(x) = y_0 + \int_{x_0}^x f(t, y_{n-1}) dt .$$

Dies ist eine Folge von Funktionen. Man beweist mit Hilfe eines Fixpunktsatzes wie 3.3.16, dass die Folge auf einem kleinen Intervall gleichmäßig konvergiert. Sei

$$y = \lim_{n \rightarrow \infty} y_n$$

die Grenzfunktion. Diese erfüllt dann

$$y = y_0 + \int_{x_0}^x f(t, y) dt \Rightarrow y' = f(x, y), y(x_0) = y_0 .$$

- Mit Hilfe von numerischer Integration wird hieraus auch ein numerisches Verfahren zur Integration von Differentialgleichungen.

**6.3.3 Beispiel.** Wir betrachten  $y' = xy, y(0) = 1$ . Picard-Iteration liefert

$$y_0(x) = 1$$

$$y_1(x) = 1 + \int_0^x ty_0(t) dt = 1 + \frac{x^2}{2}$$

$$y_2(x) = y_1(x) = 1 + \int_0^x ty_1(t) dt = 1 + \int_0^x \left( t + \frac{t^3}{2} \right) dt = 1 + \frac{x^2}{2} + \frac{x^4}{8}$$

$$y_3(x) = y_2(x) = 1 + \int_0^x ty_2(t) dt = 1 + \int_0^x \left( t + \frac{t^3}{2} + \frac{t^5}{8} \right) dt = 1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{48}$$

Induktiv erhält man

$$y_n(t) = \sum_{k=0}^n \frac{x^{2k}}{2^k k!} .$$

Diese Folge konvergiert gegen  $\exp^{\frac{1}{2}x^2}$ .

## 6.4 Potenzreihenansatz

Wir betrachten eine DGL der Form

$$F(x, y, \dots, y^{(n)}) = 0 \tag{4.1}$$

wobei  $F$  durch eine konvergente Potenzreihe (z.B. ein Polynom) gegeben sein soll. Wir setzen an

$$y(x) = \sum_{k=0}^{\infty} a_k x^k$$

und erhalten Formeln für die Ableitungen. Dann geht man wie folgt vor:

1. Den Ansatz und die Ableitung in die Gleichung (4.1) einsetzen.
2. Wir rechnen und sortieren nach Potenzen von  $x$ .
3. Durch Koeffizientenvergleich erhalten wir ein unendliches System von Gleichungen für die Koeffizienten  $a_k$ .
4. Mit den Lösungen bestimmen wir den Konvergenzradius.

Im Konvergenzbereich von  $y$  (abhängig auch vom Konvergenzbereich von  $F$ ) erhalten wir eine Lösung der Gleichung.

#### 6.4.1 Beispiel. Die Laguerre-DGL

$$xy'' + (1-x)y' + my = 0 \quad m \in \mathbb{R}$$

Wir machen den Ansatz

$$y(x) = \sum_{k=0}^{\infty} a_k x^k$$

$$y'(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}$$

$$y''(x) = \sum_{k=0}^{\infty} k(k-1) a_k x^{k-2}$$

und erhalten die Gleichung

$$0 = \sum_{k=0}^{\infty} k(k-1) a_k x^{k-1} + \sum_{k=0}^{\infty} k a_k x^{k-1} - \sum_{k=0}^{\infty} k a_k x^k + m \sum_{k=0}^{\infty} a_k x^k$$

$$= \sum_{l=0}^{\infty} [(l+1)(l+1) a_{l+1} + (m-l) a_l] x^l$$

d.h. für jedes  $l \geq 0$  muss gelten

$$(l+1)^2 a_{l+1} + (m-l) a_l \Rightarrow a_{l+1} = \frac{l-m}{(l+1)^2} a_l .$$

Mit frei gewähltem  $a_0$  ergibt sich

$$\begin{aligned} a_1 &= -ma_0 \\ a_2 &= -\frac{m-1}{2^2}a_1 = \frac{m(m-1)}{2^2}a_0 = \binom{m}{2} \frac{a_0}{2} \\ a_3 &= -\frac{m-2}{3^2}a_2 = -\binom{m}{3} \frac{a_0}{2 \cdot 3} \\ a_n &= (-1)^n \binom{m}{n} \frac{a_0}{n!} \\ y(x) &= a_0 \sum_{n=0}^{\infty} (-1)^n \binom{m}{n} \frac{1}{n!} x^n. \end{aligned}$$

Diese Reihe konvergiert für jedes  $x \in \mathbb{R}$  (Vergleich mit der binomischen Reihe 5.2.18).

Speziell für  $m \in \mathbb{N}$  bricht die Reihe bei  $n = m$  ab. Für  $a_0 = 1$  erhält man das **Laguerre-Polynom**  $L_m(x)$  vom Grad  $m$ . Es gilt

$$\begin{aligned} L_m(x) &= \frac{e^x}{m!} \frac{d^m}{dx^m} (x^m e^{-x}) \\ \int_0^{\infty} L_n(x) L_m(x) e^{-x} dx &= \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases} \end{aligned}$$

Gewichtet mit  $e^{-x}$  bilden die Laguerre-Polynome eine orthogonale Schar von Polynomen (vergleiche die Schar  $\sin(nx)$ ,  $\cos(mx)$ ).

**6.4.2 Lemma.** *Für Gleichungen zweiten Grades der Form*

$$p(x)y'' + q(x)y' + r(x)y = 0 \quad (4.2)$$

*führt der Potenzreihenansatz um  $x_0$  zu einer nichttrivialen Lösung, wenn*

$$\frac{q(x)}{p(x)}, \frac{r(x)}{p(x)}$$

*um  $x = x_0$  in Taylor-Reihen entwickelt werden können.*

**6.4.3 Beispiel. Die Legendre-Differentialgleichung**

$$(1-x^2)y'' - 2xy' + m(m+1)y = 0 \quad m \in \mathbb{R} \quad (4.3)$$

Um  $x = 0$  lassen sich die Koeffizienten in eine Reihe mit Konvergenzradius 1 entwickeln, denn

$$\frac{1}{1-x^2} = \sum_{n=0}^{\infty} x^{2n}.$$

Der Ansatz  $y(x) = \sum_{n=0}^{\infty} a_n x^n$  führt auf die Gleichung

$$\begin{aligned} 0 &= (1-x^2) \sum_{n=0}^{\infty} n(n-1)a_n x^{n-2} - 2x \sum_{n=0}^{\infty} na_n x^{n-1} + m(m+1) \sum_{n=0}^{\infty} a_n x^n \\ &= \sum_{n=0}^{\infty} n(n-1)a_n x^{n-2} + \sum_{n=0}^{\infty} (-n(n-1) - 2n + m(m+1))a_n x^n \\ &= \sum_{n=0}^{\infty} [(n+2)(n+1)a_{n+2} + (-n(n-1) - 2n + m(m+1))a_n] x^n. \end{aligned}$$

Dies führt zur Rekursionsgleichung

$$\begin{aligned} (n+2)(n+1)a_{n+2} &= [n(n-1) + 2n - m(m+1)]a_n \\ &= [n(n+1) - m(m+1)]a_n \\ &= [-(m-n)(m+n+1)]a_n \end{aligned}$$

mit frei wählbarem  $a_0, a_1 \in \mathbb{R}$ . Hieraus folgt zunächst ( $n = 2k - 2, 2k - 1$ )

$$\begin{aligned} a_{2k} &= a_{2k-2} \frac{-(m-2k+2)(m+2k-1)}{(2k)(2k-1)} \\ a_{2k+1} &= a_{2k-1} \frac{-(m-2k+1)(m+2k)}{(2k+1)(2k)} \end{aligned}$$

und mit vollständiger Induktion

$$\begin{aligned} a_{2k} &= \frac{(-1)^k}{(2k)!} \prod_{l=1}^k (m-2(l-1))(m+2l-1)a_0 \\ a_{2k+1} &= \frac{(-1)^k}{(2k+1)!} \prod_{l=1}^k (m-(2l-1))(m+2l)a_1 \end{aligned}$$

Die allgemeine Lösung hat also die Form

$$y(x) = a_0 y_1(x) + a_1 y_2(x)$$

mit

$$y_1(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \prod_{l=1}^k (m - 2(l-1))(m + 2l - 1) x^{2k}$$

$$y_2(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \prod_{l=1}^k (m - 2(l-1))(m + 2l) x^{2k+1}$$

Man sieht mit dem Quotientenkriterium, dass wegen

$$\left| \frac{a_{2k+2}}{a_{2k}} \right| = \left| \frac{1}{(2k+1)(2k+2)} (m - 2(k+1-1))(m + 2(k+1) - 1) \right| \xrightarrow{k \rightarrow \infty} 1$$

die Reihe  $y_1$  den Konvergenzradius 1 hat, analog für  $y_2$ .

Sei nun speziell  $m \in \mathbb{N}$ . Dann ist jeweils eine der Basislösungen  $y_1, y_2$  ein Polynom, existiert also auf ganz  $\mathbb{R}$ . Die andere wird jedoch singular in  $x = 1$  oder  $x = -1$ .

Für  $m = 0$  gilt  $a_2 = 0$  und daher  $y_1 = 1$ , jedoch

$$y_2(x) = \frac{1}{2} \ln \left( \frac{1+x}{1-x} \right).$$

Für  $m = 1$  gilt  $a_3 = 0$  und daher  $y_2 = x$ , jedoch

$$y_1(x) = 1 - \frac{x}{2} \ln \left( \frac{1+x}{1-x} \right).$$

**6.4.4 Definition.** Ein Punkt  $x_0$  heißt **singulärer Punkt** der Differentialgleichung (4.2), wenn  $p(x_0) = 0$  gilt. Man nennt  $x_0$  eine **reguläre Singularität**, wenn

$$p(x) = (x - x_0)^2 p_0(x), \quad q(x) = (x - x_0) p_1(x), \quad r(x) = p_2(x)$$

mit Funktionen  $p_i$ , die um  $x_0$  in eine Potenzreihe entwickelt werden können und in  $x_0$  nicht verschwinden.

**6.4.5 Lemma.** Hat (4.2) eine reguläre Singularität in  $x_0$ , so führt der Ansatz

$$y(x) = (x - x_0)^r P(x)$$

zum Ziel, wenn  $r$  eine Nullstelle der **Indexgleichung**

$$r(r-1)p_0(x_0) + r p_1(x_0) + p_2(x_0)$$

und  $P(x)$  eine Potenzreihe ist.

- Im Lemma ist  $r \in \mathbb{R}$  beliebig!

*Beweis.* Wir wollen sehen, wo die Indexgleichung herkommt. Sei  $x_0 = 0$ . Wir machen den Ansatz

$$y(x) = \sum_{k=0}^{\infty} a_k x^{r+k}$$

und setzen in die Differentialgleichung ein:

$$p_0(x) \sum_{k=0}^{\infty} (r+k)(r+k-1)a_k x^{r+k} + p_1(x) \sum_{k=0}^{\infty} (r+k)a_k x^{r+k} + p_2(x) \sum_{k=0}^{\infty} a_k x^{r+k} = 0.$$

Nach Voraussetzung sind  $p_i$  Potenzreihen. Einsetzen und Ausmultiplizieren führt dann auf die gesuchten Rekursionsgleichungen. Die Indexgleichung ist genau die Bedingung, dass der Koeffizient von  $x^r$  von der Form  $0a_0$  ist, also  $a_0$  frei wählbar ist.  $\square$

Wir zeigen das Prinzip in einem wichtigen Spezialfall.

#### 6.4.6 Beispiel. Die Besselsche Differentialgleichung

$$x^2 y'' + x y' + (x^2 - \alpha^2) y = 0 \quad (4.4)$$

mit  $p_0 = p_1 = 1, p_2 = x^2 - \alpha^2$  Die Indexgleichung lautet

$$r(r-1) + r - \alpha^2 = 0 \Rightarrow r = \pm \alpha.$$

Sei zunächst  $\alpha = m \in \mathbb{N}$ . Der Ansatz ist

$$y(x) = x^m \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+m}$$

und damit

$$x y' = \sum_{n=0}^{\infty} (n+m) a_n x^{n+m}, \quad x^2 y'' = \sum_{n=0}^{\infty} (n+m-1)(n+m) a_n x^{n+m}.$$

Einsetzen in die Differentialgleichung ergibt

$$0 = \sum_{n=0}^{\infty} [(n+m) + (n+m-1)(n+m) - \alpha^2] a_n x^{n+m} + \sum_{n=0}^{\infty} a_n x^{n+m+2}$$

und die Rekursionsgleichung

$$((n+m)^2 - m^2)a_n = \begin{cases} 0 & n = 0, 1 \\ a_{n-2} & n \geq 2 \end{cases} .$$

Hieraus folgt

$$a_0 \in \mathbb{R}, a_1 = 0, a_n = \frac{a_{n-2}}{(n+m)^2 - m^2} \Rightarrow \\ a_{2n-1} = 0, a_{2n} = (-1)^n \frac{a_0}{4^n n! (m+1)(m+2) \dots (m+n)} .$$

Die Reihe konvergiert auf ganz  $\mathbb{R}$  und löst die Besselsche Differentialgleichung. Speziell für Die Reihe für  $a_0 = 1/2^m m!$  erhält man die **Bessel-Funktion 1. Art der Ordnung  $m$** :

$$J_m(x) = \frac{x^m}{2^m} \sum_{n=0}^{\infty} \frac{(-1)^n}{4^n n! (m+n)!} x^{2n}$$

Die Vielfachen von  $J_m$  sind noch nicht die allgemeine Lösung, denn wir erwarten bei einer Gleichung der Ordnung zwei auch zwei frei wählbare Parameter. Die zweite Basislösung hat die Form

$$Y_m(x) = cJ_m(x) \ln x + x^{-m} P_2(x) ,$$

wobei  $P_2$  eine Potenzreihe bezeichnet.  $Y_m$  heißt **Bessel-Funktion 2. Art der Ordnung  $m$**  und ist bei 0 singular.

## 6.5 Systeme von Differentialgleichungen

**6.5.1 Definition.** Sei  $n \in \mathbb{N}$ ,  $D \subset \mathbb{R}^{n+1}$  und  $F_1, \dots, F_n : D \rightarrow \mathbb{R}$  Funktionen.

1. Das Gleichungssystem

$$\begin{aligned} y_1' &= F(x, y_1, \dots, y_n) \\ y_2' &= F(x, y_1, \dots, y_n) \\ &\dots \\ y_n' &= F(x, y_1, \dots, y_n) \end{aligned}$$

**System von  $n$  Differentialgleichungen 1. Ordnungen.** Man kürzt oft ab  $F = (F_1, \dots, F_n)$ ,  $y(x) = (y_1(x), \dots, y_n(x))$  und schreibt dann

$$y' = F(x, y) \quad (5.1)$$

für das System.

2. Ein Tupel  $y_1, \dots, y_n : I \rightarrow \mathbb{R}$  heißt **Lösung** des Systems (5.1), wenn für alle  $x \in I$  gilt

$$\begin{aligned} (x, y_1(x), \dots, y_n(x)) &\in D \\ F_1(x, y_1(x), \dots, y_n(x)) &= 0 \\ F_2(x, y_1(x), \dots, y_n(x)) &= 0 \\ &\dots \\ F_n(x, y_1(x), \dots, y_n(x)) &= 0 \end{aligned}$$

**6.5.2 Satz.** Sei  $D \subset \mathbb{R}^{n+1}$  offen (z.B. ein Produkt von offenen Intervallen),  $F_1, \dots, F_n$  in jeder Variablen stetig differenzierbar. Sei  $(a, b_1, \dots, b_n) \in \mathbb{R}^{n+1}$ . Dann gibt es  $\varepsilon > 0$ , so dass das System von Differentialgleichungen 1. Ordnung

$$y' = F(x, y)$$

auf  $(a - \varepsilon, a + \varepsilon)$  eine Lösung mit

$$y(a) = b .$$

(d.h.  $y_i(a) = b_i$ ). Die Lösung ist eindeutig.

*Beweis.* Wie im Falle einer Gleichung erster Ordnung (Satz 6.3.1) wird mit Picard-Iteration

$$y_k = b + \int_a^x F(t, y_{k-1}) dt$$

eine Folge von Tupeln von Funktion definiert, die gleichmäßig konvergiert. Die Grenzfunktionen bilden dann die Lösung.  $\square$

**6.5.3 Folgerung.** Sei  $D \subset \mathbb{R}^{n+1}$  und  $f : D \rightarrow \mathbb{R}$  in jeder Variablen stetig differenzierbar. Sei  $(a, b_0, \dots, b_{n-1}) \in D$ . Dann hat das Anfangswertproblem

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \quad f^{(k)}(a) = b_k$$

eine lokale Lösung  $y : I \rightarrow \mathbb{R}$  um  $a$ . Diese ist eindeutig.

*Beweis.* Wir gehen über zu einem System von lineare Differentialgleichungen 1. Ordnung. Wir setzen

$$\begin{aligned} y_1 &:= y \\ y_2 &:= y' \\ &\dots \\ y_n &:= y^{(n-1)} \end{aligned}$$

und betrachten

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\dots \\ y_{n-1}' &= y_n \\ y_n' &= f(x, y_1, \dots, y_{n-1}) \end{aligned} \qquad = y_{n-2}$$

Offensichtlich liefert jede Lösung des Systems eine Lösung des Gleichung  $n$ -ter Ordnung. Die Folgerung folgt aus dem Satz.  $\square$

**6.5.4 Beispiel.** Die Differentialgleichung 2. Ordnung

$$y'' = -y$$

ist äquivalent zu dem System

$$y_1' = y_2, y_2' = -y_1$$

Die Lösung ist offensichtlich  $y = y_1 = c_1 \sin x + c_2 \cos x, y_2 = c_1 \cos x - c_2 \sin x$ .

# Kapitel 7

## Lineare Algebra

### 7.1 Lineare Gleichungssysteme und Matrizen

**7.1.1 Definition.** Ein rechteckiges Zahlenschema der Form

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (1.1)$$

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

$$\mathbf{A} = (a_{ij}) \quad (1.2)$$

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  $\mathbb{R}^{m \times n}$ . Wir schreiben  $\mathbb{R}^n := \mathbb{R}^{n \times 1}$  (Spaltenvektoren) und  $\mathbb{R}_m := \mathbb{R}^{1 \times m}$  (Zeilenvektoren).

**7.1.2 Definition.** 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  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ .

- Genauso werden addiert man Spaltenvektoren und multipliziert sie mit Skalaren. Zeilenvektoren werden analog behandelt.
- Matrizen oder Vektoren unterschiedlicher Größe lassen sich nicht addieren.
- 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

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

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

$$\mathbf{a} = a_1 \mathbf{e}_1 + \cdots + a_n \mathbf{e}_n \quad (1.3)$$

von Vielfachen der **Standard Basisspaltenvektoren**

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n := \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (1.4)$$

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  $\mathbb{R}_n$ .

- 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**.



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:

$$3x_1 + 2x_2 = 1$$

$$3x_1 + 2x_2 = 2$$

- b) genau eine Lösung:

$$3x_1 + 2x_2 = 1$$

$$3x_1 + x_2 = 5$$

Wir wollen das Verfahren am konkreten Beispiel vorstellen:

- Koeffizienten von  $x_1$  in 1. Gleichung = 1, d.h. 1. Gleichung  $\times \frac{1}{3}$ ; Eliminieren von  $x_1$  aus 2. Gleichung, d.h. 2. Gleichung -  $3 \times$  neue 1. Gleichung

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

- Koeffizient von  $x_2$  in 2. Gleichung = 1, d.h. 2. Gleichung  $\times (-1)$ ; Eliminieren von  $x_2$  aus 1. Gleichung, d.h. 1. Gleichung -  $\frac{2}{3} \times$  neue 2. Gleichung

$$x_1 = 3$$

$$x_2 = -4$$

Also ist  $x_1 = 3, x_2 = -4$  die Lösung. *Probe:*

$$3 \cdot 3 + 2 \cdot (-4) = 1$$

$$3 \cdot 3 + (-4) = 5$$

c) unendlich viele Lösungen

$$3x_1 + 2x_2 = 1$$

Das ist eine Geradengleichung, mit einem Parameter  $\lambda$ . Also ist

$$\begin{aligned} x_2 &= \lambda \\ x_1 &= \frac{1}{3}(1 - 2\lambda) \end{aligned}$$

eine Lösung für alle  $\lambda \in \mathbb{R}$ .

**7.1.4 Definition.** Sei  $\sum_{j=1}^n a_{ij}x_j = b_i$ ,  $i = 1, \dots, m$  ein lineares Gleichungssystem. Die Matrix

$$\mathbf{A} := (a_{ij}) \in \mathbb{R}^{m \times n}$$

heißt **Koeffizientenmatrix**, der Spaltenvektor

$$\mathbf{b} := (b_1, \dots, b_m)^T \in \mathbb{R}^m$$

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**.

**7.1.5 Satz.** 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.

• Die Umformungen aus Satz 7.1.5 haben Entsprechungen für die erweiterte Koeffizientenmatrix.

**7.1.6 Definition.** Sei  $(\mathbf{A}|\mathbf{b})$  die 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  $\alpha$ -fachen einer Zeile zu einer anderen.

**7.1.7 Satz.** *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.*

**7.1.8** 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:*

0. Initialisierung. Wir beginnen mit der *Zeilennummer*  $i = 1$ . und der leeren *Buchführungsmenge*  $B = \emptyset$ .
1. Nächste Spalte. Es sei  $j_i$  die Nummer der ersten Spalte von  $\mathbf{A}$ , die vom  $i$ -ten Element an nach unten nicht nur aus Nullen besteht, das heißt, es gibt ein  $k \in \{i, \dots, m\}$ , so dass  $a_{k,j_i} \neq 0$ . Falls es keine solche Spalte mehr gibt: **Ende.** Ansonsten ersetze die Buchführungsmenge  $B$  durch  $B \cup \{j_i\} = \{j_1, \dots, j_i\}$ .
2. Vertauschen zweier Zeilen (Operation a). Wenn  $a_{i,j_i} = 0$  ist, vertausche die  $i$ -te Zeile  $\mathbf{z}_i$  der erweiterten Koeffizientenmatrix mit der  $k$ -Zeile  $\mathbf{z}_k$  ( $k$  wie in Schritt 1). **Alternativ:** vertausche die  $i$ -te Zeile mit der Zeile  $k \geq 1$ , für die  $a_{k,j_i}$  betragsmäßig am größten ist.
3. Normierung der  $i$ -ten Zeile (Operation b). Multipliziere die gesamte  $i$ -te Zeile  $\mathbf{z}_i$  elementweise mit  $a_{i,j_i}^{-1}$ . In der neuen erweiterten Koeffizientenmatrix ist also  $a_{i,j_i} = 1$ , alle Glieder links davon sind 0.
4. Ausräumen der  $j_i$ -ten Spalte (Operation c). Zu allen anderen Zeilen  $\mathbf{z}_k$  mit  $k \neq i$  addiere das  $-a_{k,j_i}$ -fache der  $i$ -ten Zeile  $\mathbf{z}_i$ . Hinterher ist  $a_{k,j_i} = 0$  für alle  $k \neq i$ , das heißt, in der  $j_i$ -ten Spalte steht jetzt der  $i$ -te Standardbasisvektor  $\mathbf{e}_i$ .
5. Schleife. Falls  $i = m$ , haben wir alle Zeilen bearbeitet: **Ende.** Ansonsten ersetze  $i$  durch  $i + 1$  und mache weiter mit Schritt 1.

Es reicht, pro Schleifendurchlauf eine neue Matrix aufzuschreiben.

Wir erhalten dadurch eine Buchführungsmenge  $B = \{j_1, \dots, j_r\}$  mit  $j_1 < j_2 < \dots < j_r \leq n$  und eine erweiterte Koeffizientenmatrix in **Stufenform**.  
Etwa bei 6 Gleichungen mit 10 Unbekannten und  $B = \{2, 5, 8, 9\}$ .

$$\left( \begin{array}{cccccccccc|c} 0 & 1 & * & * & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & 0 & * & * \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & 1 & 0 & * & * \\ \vdots & 0 & 1 & * & * \\ \vdots & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{array} \right)$$

- Die Buchführungsmenge  $B = \{j_1, j_2, \dots, j_r\}$  ist dadurch gekennzeichnet, daß in der  $j_k$ -ten Spalte der  $k$ -te Standardbasisvektor  $\mathbf{e}_k$  steht, also steht in der  $k$ -ten Zeile eine 1 und sonst nur Nullen.

**7.1.9 Satz Lösbarkeit.** *Sei eine erweiterte Koeffizientenmatrix in Stufenform mit Buchführungsmenge  $B = \{j_1, \dots, j_r\}$ . Dann besitzt das zugehörige lineare Gleichungssystem genau dann Lösungen, wenn alle Absolutglieder  $b_i$  mit Index  $i > r$  Null sind.*

**7.1.10 Satz 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.*

**7.1.11 Definition.** *Die im Beweis von Satz 7.1.10 in Abhängigkeit von den Parametern  $c_j, j \notin B$ , konstruierte Lösung heißt **allgemeine Lösung**. Jede spezielle Wahl der Parameter liefert eine **spezielle Lösung**.*

**7.1.12 Satz Eindeutigkeit.** *Das System ist genau dann eindeutig lösbar wenn es lösbar ist und  $n = r$ , d.h.  $B = \{1, \dots, n\}$ .*

**7.1.13 Folgerung.** *Das homogene System, d.h.  $b_i = 0, i = 1, \dots, m$ , besitzt genau dann nur die triviale Lösung, wenn  $r = n$ .*

**7.1.14 Satz 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.

b) 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.

## 7.2 Matrizenmultiplikation

**7.2.1 Definition.** 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  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b} \in \mathbb{R}^n$

$$\begin{aligned} (\mathbf{a}_1 + \mathbf{a}_2) \cdot \mathbf{b} &= \mathbf{a}_1 \cdot \mathbf{b} + \mathbf{a}_2 \cdot \mathbf{b}, \\ \mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) &= \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2, \\ \alpha (\mathbf{a} \cdot \mathbf{b}) &= (\alpha \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha \mathbf{b}). \end{aligned} \tag{2.1}$$

**7.2.2 Definition.** 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

$$c_{ij} := \sum_{k=1}^n a_{ik} b_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, r, \tag{2.2}$$

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

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad (2.3)$$

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:

$$\mathbf{A} \cdot \mathbf{s} = \begin{pmatrix} \mathbf{z}_1 \cdot \mathbf{s} \\ \vdots \\ \mathbf{z}_m \cdot \mathbf{s} \end{pmatrix} \quad (2.4)$$

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:

$$\mathbf{A} \cdot \mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i = x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n. \quad (2.5)$$

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  $\mathbf{A}$ .

- Im Allgemeinen gilt:

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A},$$

auch wenn beide Produkte definiert sind. **Beispiel:**

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & -3 \\ 14 & 19 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 14 & 11 \end{pmatrix}$$

- Für alle  $n$  heißt

$$\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \quad (2.6)$$

die  $n \times n$  **Einheitsmatrix**, sie hat die Spaltendarstellung  $\mathbf{E}_n = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  mit den Basisspaltenvektoren  $\mathbf{e}_i \in \mathbb{R}^n$  aus (1.4).

- Die **Nullmatrizen** sind die  $m \times n$ -Matrizen

$$\mathbf{0}_{m,n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}. \quad (2.7)$$

**7.2.3 Satz.** Für alle Matrizen  $\mathbf{A}, \mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B}, \mathbf{B}_1, \mathbf{B}_2 \in \mathbb{R}^{n \times r}$ ,  $\mathbf{C} \in \mathbb{R}^{r \times s}$  und  $\alpha \in \mathbb{R}$  gilt:

- $(\mathbf{A}_1 + \mathbf{A}_2) \cdot \mathbf{B} = \mathbf{A}_1 \cdot \mathbf{B} + \mathbf{A}_2 \cdot \mathbf{B}$ ,
- $\mathbf{A} \cdot (\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A} \cdot \mathbf{B}_1 + \mathbf{A} \cdot \mathbf{B}_2$ ,
- $\alpha(\mathbf{A} \cdot \mathbf{B}) = (\alpha\mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\alpha\mathbf{B})$ ,
- $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$ ,
- $\mathbf{E}_m \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{E}_n = \mathbf{A}$ .
- $\mathbf{0}_{\ell,m} \cdot \mathbf{A} = \mathbf{0}_{\ell,n}$  und  $\mathbf{A} \cdot \mathbf{0}_{n,r} = \mathbf{0}_{m,r}$ .

- Es gibt Matrizen  $\mathbf{A}, \mathbf{B} \neq \mathbf{0}$  so daß  $\mathbf{A} \cdot \mathbf{B} = \mathbf{0}$ . Dies ist bei reellen Zahlen nicht möglich, denn für  $a, b \in \mathbb{R}$  gilt:  $a \cdot b = 0 \Rightarrow (a = 0) \vee (b = 0)$ .
- **Potenzen** von quadratischen Matrizen  $\mathbf{A} \in \mathbb{R}^{n \times n}$  sind rekursiv definiert:

$$\mathbf{A}^0 = \mathbf{E}_n, \quad \mathbf{A}^k := \mathbf{A}^{k-1} \cdot \mathbf{A} = \underbrace{\mathbf{A} \cdot \dots \cdot \mathbf{A}}_{k\text{-mal}} \quad (2.8)$$

**7.2.4 Definition.** Zu jeder  $m \times n$  Matrix  $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$  gibt es die **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

$$\mathbf{A} = \left( \begin{array}{ccc} \boxed{a_{11}} & \cdots & \boxed{a_{1n}} \\ \vdots & & \vdots \\ \boxed{a_{m1}} & \cdots & \boxed{a_{mn}} \end{array} \right) \in \mathbb{R}^{m \times n}$$

ist die transponierte Matrix gegeben durch

$$\mathbf{A}^T = \left( \begin{array}{ccc} \boxed{a_{11}} & \cdots & \boxed{a_{m1}} \\ \vdots & & \vdots \\ \boxed{a_{1n}} & \cdots & \boxed{a_{mn}} \end{array} \right) \in \mathbb{R}^{n \times m}.$$

- Die Bezeichnung ist konsistent mit unserer Schreibweise  $(a_1, \dots, a_n)^T$  für Spaltenvektoren
- Es gelten folgende Rechenregeln für alle  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R}$ :

$$\begin{aligned} (\mathbf{A} + \mathbf{B})^T &= \mathbf{A}^T + \mathbf{B}^T, \\ (\alpha \mathbf{A})^T &= \alpha \mathbf{A}^T, \\ (\mathbf{A}^T)^T &= \mathbf{A}, \\ (\mathbf{A} \cdot \mathbf{B})^T &= \mathbf{B}^T \cdot \mathbf{A}^T. \end{aligned} \tag{2.9}$$

**7.2.5 Definition.** Eine  $n \times n$  Matrix  $\mathbf{A}$  heißt **symmetrisch**, wenn  $\mathbf{A}^T = \mathbf{A}$  gilt; sie heißt **schiefssymmetrisch**, 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:

$$\mathbf{A} = \mathbf{A}^T \quad \Leftrightarrow \quad a_{ij} = a_{ji} \quad \forall i, j = 1, \dots, n, \tag{2.10}$$

und analog gilt für schiefsymmetrische Matrizen:

$$\mathbf{A} = -\mathbf{A}^T \Leftrightarrow a_{ij} = -a_{ji} \quad \forall i, j = 1, \dots, n. \quad (2.11)$$

Insbesondere gilt für die Diagonalelemente einer schiefsymmetrischen Matrix  $a_{ii} = 0$ .

**7.2.6 Definition.** 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  $\mathbf{A}$ .

- Die Eindeutigkeit von  $\mathbf{B}$  muß bewiesen werden.

**7.2.7 Satz.** Wenn es zu einer Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  zwei Matrizen  $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$  gibt mit  $\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{C} = \mathbf{E}_n$ , dann ist  $\mathbf{A}$  invertierbar und es gilt:

$$\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}.$$

- Beispiel: Für  $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  mit  $ad - bc \neq 0$  gilt:

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

**7.2.8 Satz.**

- a) Die inverse Matrix einer invertierbaren Matrix  $\mathbf{A}$  ist invertierbar und es gilt:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}.$$

- b) Das Produkt zweier invertierbarer Matrizen ist invertierbar und es gilt:

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}.$$

- c) Die Transponierte  $\mathbf{A}^T$  einer Matrix ist genau dann invertierbar, wenn  $\mathbf{A}$  invertierbar ist. In diesem Fall gilt:

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

- Für das Produkt mehrerer invertierbarer Matrizen  $\mathbf{A}_i, i = 1, \dots, n$ , gilt:

$$(\mathbf{A}_1 \cdot \dots \cdot \mathbf{A}_n)^{-1} = \mathbf{A}_n^{-1} \cdot \mathbf{A}_{n-1}^{-1} \cdot \dots \cdot \mathbf{A}_1^{-1}. \quad (2.12)$$

- 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.

- Wenn  $\mathbf{A}$  selbst nicht invertierbar ist, kann man trotzdem mit einer invertierbaren Matrix  $\mathbf{P}$  multiplizieren und erhält das äquivalente Gleichungssystem

$$(\mathbf{P} \cdot \mathbf{A}) \cdot \mathbf{x} = \mathbf{P} \cdot \mathbf{b}.$$

Genau das passiert beim Gauß-Verfahren, wie wir gleich sehen werden.

**7.2.9 Definition.** Eine  $m \times m$  Matrix  $\tilde{\mathbf{E}}$  heißt **Elementarmatrix vom Typ  $i$** ,  $i=(a), (b), (c)$ , wenn sie aus der  $m \times m$  Einheitsmatrix  $\mathbf{E}$  durch **eine** elementare Zeilenumformung vom Typ  $i$  aus Definition 7.1.6 hervorgeht.

Beispiele:

$$\begin{pmatrix} \mathbf{e}'_2 \\ \mathbf{e}'_1 \\ \mathbf{e}'_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Vertauschen von } \mathbf{e}'_1 \text{ und } \mathbf{e}'_2: \text{ Typ (a)}$$

$$\begin{pmatrix} \mathbf{e}'_1 \\ \alpha \mathbf{e}'_2 \\ \mathbf{e}'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Multiplikation von } \mathbf{e}'_2 \text{ mit } \alpha \neq 0: \text{ Typ (b)}$$

$$\begin{pmatrix} \mathbf{e}'_1 + \alpha \mathbf{e}'_3 \\ \mathbf{e}'_2 \\ \mathbf{e}'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Addition von } \alpha \cdot \mathbf{e}'_3 \text{ zu } \mathbf{e}'_1 \Rightarrow \text{ Typ (c)}$$

**7.2.10 Satz.**

- a) Entsteht  $\tilde{\mathbf{A}}$  aus  $\mathbf{A} \in \mathbb{R}^{m \times n}$  durch eine elementare Zeilenumformung, dann gilt:

$$\tilde{\mathbf{A}} = \tilde{\mathbf{E}}\mathbf{A}$$

mit der zugehörigen Elementarmatrix  $\tilde{\mathbf{E}}$ .

- b) Die Elementarmatrizen sind invertierbar, die Inversen sind ebenfalls Elementarmatrizen.
- c) Entsteht  $\mathbf{M}$  bzw.  $\mathbf{N}$  aus  $\mathbf{A} \in \mathbb{R}^{m \times n}$  durch endlich viele elementare Zeilen- bzw. Spaltenumformungen, dann gibt es invertierbare Matrizen  $\mathbf{P} \in \mathbb{R}^{m \times m}$ ,  $\mathbf{Q} \in \mathbb{R}^{n \times n}$  mit

$$\mathbf{M} = \mathbf{P}\mathbf{A}, \quad \mathbf{N} = \mathbf{A}\mathbf{Q}. \quad (2.13)$$

$\mathbf{P}$  und  $\mathbf{Q}$  sind Produkte von Elementarmatrizen.

- Insbesondere wählt der Gauß-Algorithmus 7.1.8 eine invertierbare Matrix  $\mathbf{P}$  so aus, dass  $\mathbf{P} \cdot \mathbf{A}$  in strenger Zeilenstufenform ist.
- Es seien  $\mathbf{A}$  und  $\mathbf{B} \in \mathbb{R}^{m \times m}$ , und  $\mathbf{B} = (\mathbf{s}_1, \dots, \mathbf{s}_m)$  sei die Spalten-Darstellung. Dann hat  $\mathbf{A} \cdot \mathbf{B}$  die Spaltendarstellung

$$\mathbf{A} \cdot \mathbf{B} = (\mathbf{A} \cdot \mathbf{s}_1, \dots, \mathbf{A} \cdot \mathbf{s}_m).$$

- Also erfüllen die einzelnen Spalten  $\mathbf{s}_j$  der Inversen von  $\mathbf{A}$  jeweils ein lineares Gleichungssystem

$$\mathbf{A} \cdot \mathbf{s}_j = \mathbf{e}_j.$$

Um die inverse Matrix zu berechnen, lösen wir all diese Gleichungssysteme simultan, indem wir das Gauß-Verfahren 7.1.8 auf die erweiterte Koeffizientenmatrix  $(\mathbf{A}|\mathbf{E}_m)$  anwenden. Das ist möglich, da die einzelnen Zeilenoperationen, die wir im Laufe des Verfahrens anzuwenden haben, einzig durch die linke Seite bestimmt werden.

**7.2.11 Satz.** *Es sei  $\mathbf{A}$  eine  $m \times m$ -Matrix. Wenn das Gauß-Verfahren für  $(\mathbf{A}|\mathbf{E}_m)$  die Buchführungsmenge  $B = \{1, \dots, m\}$  liefert, dann hat das Ergebnis die Gestalt  $(\mathbf{E}_m|\mathbf{B})$ , und  $\mathbf{A}$  ist invertierbar mit der Inversen  $\mathbf{A}^{-1} = \mathbf{B}$ . Andernfalls ist  $\mathbf{A}$  nicht invertierbar.*

Wir dürfen das Gauß-Verfahren frühzeitig abbrechen, wenn wir in Schritt 1 eine Spalte überspringen müssten, da wir dann nicht mehr die volle Buchführungsmenge  $B = \{1, \dots, m\}$  erhalten können.

**7.2.12 Folgerung.** *Es sei  $\mathbf{A} \in \mathbb{R}^{m \times m}$  eine quadratische Matrix. Dann sind äquivalent:*

1.  $\mathbf{A}$  is invertierbar,
2. es gibt  $\mathbf{B} \in \mathbb{R}^{m \times m}$  mit  $\mathbf{A} \cdot \mathbf{B} = \mathbf{E}_m$ ,
3. es gibt  $\mathbf{B} \in \mathbb{R}^{m \times m}$  mit  $\mathbf{B} \cdot \mathbf{A} = \mathbf{E}_m$ ,
4.  $\mathbf{A}$  lässt sich als Produkt von Elementarmatrizen darstellen,
5. aus  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  folgt  $\mathbf{x} = \mathbf{0}$ .

Wenn  $\mathbf{A}$  nicht quadratisch ist, ist die obige Aussage im Allgemeinen nicht richtig.

**7.2.13** 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*

$$\text{Diag}(a_1, \dots, a_n) := \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_n \end{pmatrix}.$$

*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.*

**7.2.14 Satz.** *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:

$$\mathbf{A}^{-1} = \text{Diag}(a_1^{-1}, \dots, a_n^{-1})$$

## 7.3 Vektorräume und Basen

**7.3.1 Definition.** 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  $\lambda$ -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)  $(\mu + \lambda)\mathbf{a} = \mu\mathbf{a} + \lambda\mathbf{a}$  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 (1.5)).

2) Insbesondere sind

$$\mathbb{R}_n = \{(a_1, \dots, a_n); a_i \in \mathbb{R}\}$$

$$\mathbb{R}^n = \{(a_1, \dots, a_n)^T; a_i \in \mathbb{R}\}$$

Vektorräume.

3) Die Menge

$$P_n(\mathbb{R}) := \{a_0 + a_1x + \dots + a_nx^n; a_i \in \mathbb{R}, i = 0, \dots, n\}$$

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

$$(f + g)(x) := f(x) + g(x)$$

$$(\lambda f)(x) := \lambda f(x)$$

bildet den Vektorraum der stetigen Funktionen.

- Im folgenden sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Für all  $\mathbf{a} \in V$  gilt

$$0 \cdot \mathbf{a} = \mathbf{0} ,$$

$$(-1) \cdot \mathbf{a} = -\mathbf{a} .$$

**7.3.2 Definition.** 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 sofort, dass auch  $U$  ein Vektorraum ist.

**Beispiele:**

- 1)  $V$  besitzt die **trivialen Unterräume**  $U = \{\mathbf{0}\}, U = V$ .
- 2) Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Dann ist die Lösungsmenge

$$\text{Ker } \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^n; \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

des homogenen Gleichungssystems  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  ein Unterraum von  $\mathbb{R}^n$ .

**Achtung:** Sei  $\mathbf{b} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ , dann ist die Lösungsmenge des inhomogenen Gleichungssystems  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  **kein** Unterraum von  $\mathbb{R}^n$ , denn  $\mathbf{0}$  ist keine Lösung.

- 3) Die Lösungsmenge eines **homogenen Systems linearer Differentialgleichungen** ist ein Unterraum von  $C^0(I)$  (sogar von  $C^\infty(I)$ , wenn die Koeffizientenfunktionen glatt sind). Etwa Beispiel 6.5.4: die Lösungsmenge von  $f'' + f = 0$  ist

$$\{\lambda \cos + \mu \sin \mid \lambda, \mu \in \mathbb{R}\},$$

ein Unterraum von  $C^\infty(\mathbb{R})$ .

**7.3.3 Definition.** Jede aus endlich vielen Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$  gebildete Summe der Form

$$\sum_{i=1}^k \alpha_i \mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k$$

mit Koeffizienten  $\alpha_i \in \mathbb{R}$  heißt **Linearkombination** der  $\mathbf{v}_i$ . Eine solche Linearkombination heißt *trivial*, wenn  $\alpha_i = 0$  für  $i = 1, \dots, k$  gilt. Die Menge aller Linearkombinationen der  $\mathbf{v}_i$  heißt **lineare Hülle der  $\mathbf{v}_i$**  und wird mit

$$\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) := \left\{ \sum_{i=1}^k \alpha_i \mathbf{v}_i; \alpha_i \in \mathbb{R}, i = 1, \dots, k \right\}$$

bezeichnet.

**7.3.4 Lemma.** Die lineare Hülle der Vektoren  $\mathbf{v}_i \in V, i = 1, \dots, k$ , ist ein Unterraum von  $V$ .

**7.3.5 Definition.** Man sagt, ein Unterraum  $U$  von  $V$  wird von den Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k$  **erzeugt** oder auch,  $(\mathbf{v}_1, \dots, \mathbf{v}_k)$  ist ein **Erzeugendensystem** von  $U$ , wenn

$$U = \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

Beispiele:

1)  $\mathbb{R}^n$  wird wegen (1.3) 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 einem Wechsel des Koordinatensystems.

**7.3.6 Definition.**

a) *Endlich viele Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$  heißen **linear abhängig**, wenn  $\mathbf{0}$  eine nicht-triviale Linearkombination der  $\mathbf{v}_i$  ist, das heißt, 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.*

$$\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k = \mathbf{0} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \dots = \alpha_k = 0.$$

- Zum Beispiel sind  $(\cos, \sin)$  in  $C^0(\mathbb{R})$  linear unabhängig. Denn sei  $\lambda \cos + \mu \sin = \mathbf{0}$ , also  $\lambda \cos(x) + \mu \sin(x) = 0$  für alle  $x \in \mathbb{R}$ . Einsetzen von  $x = 0$  und  $x = \frac{\pi}{2} = 90^\circ$  liefert

$$\begin{aligned} \lambda &= \lambda \underbrace{\cos(0)}_{=1} + \mu \underbrace{\sin(0)}_{=0} = 0 \\ \mu &= \lambda \underbrace{\cos\left(\frac{\pi}{2}\right)}_{=0} + \mu \underbrace{\sin\left(\frac{\pi}{2}\right)}_{=1} = 0. \end{aligned}$$

- Man kann eine Linearkombination von Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$  auch kurz als Matrixprodukt schreiben:

$$\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k = (\mathbf{v}_1, \dots, \mathbf{v}_k) \cdot (\alpha_1, \dots, \alpha_k)^T \in \mathbb{R}^n.$$

- Um zu überprüfen, ob  $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$  linear unabhängig sind, muss man das lineare Gleichungssystem in  $x_1, \dots, x_k$

$$x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k = \mathbf{0} \quad (3.1)$$

betrachten (vergleiche (2.5)) und überprüfen, ob es nichttriviale Lösungen gibt. Es gilt:

- i)  $\mathbf{v}_1, \dots, \mathbf{v}_k$  linear abhängig  $\iff$  (3.1) besitzt eine nichttriviale Lösung.
  - ii)  $\mathbf{v}_1, \dots, \mathbf{v}_k$  linear unabhängig  $\iff$  (3.1) besitzt nur die triviale Lösung.
- Sonderfälle von Definition 7.3.6
    - 1)  $k = 1$ :  $\mathbf{v} \in V$  ist linear unabhängig  $\iff \mathbf{v} \neq \mathbf{0}$ .
    - 2)  $k = 2$ :  $(\mathbf{u}, \mathbf{v})$  sind linear abhängig  $\iff \exists \alpha, \beta \in \mathbb{R}$ , nicht beide gleich Null, mit

$$\alpha \mathbf{u} + \beta \mathbf{v} = \mathbf{0} \quad \iff \quad \begin{cases} \mathbf{u} = -\frac{\beta}{\alpha} \mathbf{v}, & \text{falls } \alpha \neq 0, \\ \mathbf{v} = -\frac{\alpha}{\beta} \mathbf{u}, & \text{falls } \beta \neq 0. \end{cases}$$

**7.3.7 Lemma.** *Die Vektoren sind genau dann linear abhängig, wenn sich einer von ihnen als Linearkombination der anderen darstellen läßt.*

**7.3.8 Lemma.**

- 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.*

**7.3.9 Satz.** *Für Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w} \in V$  gilt:*

- a)  $\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}) = \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k) \iff \mathbf{w} \in \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k)$
- b) *Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k$  sind linear unabhängig  $\iff$  Zur Erzeugung von  $\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k)$  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**.*

**7.3.10 Definition.** 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.  $\text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_n) = V$ .

**7.3.11 Satz.** 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

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n. \quad (3.2)$$

Ferner sind je  $m$  Vektoren aus  $V$  linear abhängig, falls  $m > n$ .

- Die Menge der Basisspaltenvektoren  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  ist eine Basis des  $\mathbb{R}^n$  im Sinne obiger Definition.
- Im Raum  $P_k(\mathbb{R}) = \{\text{Polynome vom Grad } \leq k\}$  bilden die Polynome  $\{1, x, x^2, \dots, x^k\}$  eine Basis.
- Die Funktionen  $(\cos, \sin)$  bilden eine Basis des Lösungsraums der Differentialgleichung  $f'' + f = 0$  aus Beispiel 6.5.4. Folglich gibt es zu jeder Lösung  $f$  genau ein Paar  $(\lambda, \mu)^T \in \mathbb{R}^2$ , so dass

$$f(x) = \lambda \cos(x) + \mu \sin(x) \quad \text{für alle } x.$$

**7.3.12 Definition.** Es sei  $V$  ein Vektorraum mit Basis  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ . Mit Satz 7.3.11 können wir eine Abbildung  $k_B: V \rightarrow \mathbb{R}^n$  definieren, die jedem Vektor  $\mathbf{a} \in V$  wie in (3.2) seine **Koordinaten**

$$k_B(\mathbf{a}) = (\alpha_1, \dots, \alpha_n)^T \in \mathbb{R}^n$$

zuordnet. Diese Abbildung heißt **Koordinatenabbildung bezüglich der Basis  $\mathbf{B}$** .

Mit diesen Koordinaten können wir genauso rechnen wie mit den Vektoren aus  $V$ . Die Umkehrabbildung  $k_B^{-1}$  bildet  $(\alpha_1, \dots, \alpha_n)^T$  auf die entsprechende Linearkombination der Basisvektoren ab.

**7.3.13 Definition.** Ein Vektorraum  $V$  heißt **endlichdimensional**, wenn es endlich viele Vektoren  $\mathbf{w}_1, \dots, \mathbf{w}_r$  mit  $V = \text{Lin}(\mathbf{w}_1, \dots, \mathbf{w}_r)$  gibt.

- Die Räume  $\mathbb{R}^n, \mathbb{R}_n, P_k(\mathbb{R})$  sind endlichdimensional.
- Der Raum  $P(\mathbb{R}) = \{\text{Polynom mit beliebigem Grad}\}$  ist nicht endlichdimensional.

**7.3.14 Satz** 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.*

**7.3.15 Satz.** *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:  $m = n$ .*

**7.3.16 Definition.** *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:

$$\dim \mathbb{R}^n = \dim \mathbb{R}_n = n,$$

$$\dim P_k(\mathbb{R}) = k + 1,$$

$$\dim\{x \in \mathbb{R}^3, 3x_1 + x_2 + x_3 = 0\} = 2. \quad (\text{zwei Parameter frei wählbar})$$

**7.3.17 Lemma.** *Ist  $r$  die Maximalzahl linear unabhängiger Vektoren aus  $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ , dann gilt:*

$$r = \dim \text{Lin}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

**7.3.18 Satz.** *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.*

**7.3.19 Satz.** *Jeder Unterraum  $U$  eines endlichdimensionalen Vektorraumes  $V$  ist endlichdimensional. Im Falle  $U \neq V$  gilt:  $\dim U < \dim V$ .*

- Nach Satz 7.3.19 ist ein Unterraum  $U$  eines endlichdimensionalen Vektorraumes wieder ein endlichdimensionaler Vektorraum. Also gelten die Sätze 7.3.14, 7.3.15, 7.3.18 analog für  $U$ .

- Der Raum  $C^0(\mathbb{R})$  ist unendlich-dimensional, denn  $P(\mathbb{R}) \subset C^0(\mathbb{R})$  ist ein unendlich-dimensionaler Unterraum.

**7.3.20 Satz.** Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$  eine quadratische Matrix. Dann sind die folgenden Aussagen äquivalent:

- $\mathbf{A}$  ist invertierbar.
- Die Spalten von  $\mathbf{A}$  sind linear unabhängig.
- Die Zeilen von  $\mathbf{A}$  sind linear unabhängig.
- Die Spalten von  $\mathbf{A}$  erzeugen  $\mathbb{R}^n$ .
- Die Zeilen von  $\mathbf{A}$  erzeugen  $\mathbb{R}^n$ .
- Die Spalten von  $\mathbf{A}$  bilden eine Basis von  $\mathbb{R}^n$ .
- Die Zeilen von  $\mathbf{A}$  bilden eine Basis von  $\mathbb{R}^n$ .

**7.3.21 Definition.** Sei  $V$  ein Vektorraum, und seien  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$  und  $\mathbf{C} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$  Basen von  $V$ . Die **Basiswechselmatrix**  $\mathbf{M}_C^B(\text{id})$  von der Basis  $\mathbf{B}$  zur Basis  $\mathbf{C}$  ist diejenige Matrix, deren  $j$ -te Spalte aus den  $\mathbf{C}$ -Koordinaten des  $j$ -ten Basisvektors von  $\mathbf{B}$  besteht, also

$$\mathbf{M}_C^B(\text{id}) = (k_C(\mathbf{v}_1), \dots, k_C(\mathbf{v}_n)). \quad (3.3)$$

Merkvers: „Bilder der alten Basis in die Spalten“.

**7.3.22 Satz.** Sei  $\mathbf{M}_S^B(\text{id})$  die Basiswechselmatrix für den Übergang von der Basis  $\mathbf{B}$  zur Basis  $\mathbf{C}$ . Dann gilt für alle  $\mathbf{v} \in V$ :

$$k_C(\mathbf{v}) = \mathbf{M}_C^B(\text{id}) \cdot k_B(\mathbf{v}), \quad (3.4)$$

d.h. man erhält die  $\mathbf{C}$ -Koordinaten von  $\mathbf{v}$  indem man  $\mathbf{M}_C^B(\text{id})$  mit den  $\mathbf{B}$ -Koordinaten von  $\mathbf{v}$  multipliziert.

- Im Spezialfall, dass  $\mathbf{S} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  die Standardbasis von  $V = \mathbb{R}^n$  ist, bedeutet (3.3) gerade

$$\mathbf{M}_S^B(\text{id}) = (\mathbf{v}_1, \dots, \mathbf{v}_n) =: \mathbf{B},$$

- $\mathbf{M}_C^B(\text{id})$  ist invertierbar und es gilt

$$k_B(\mathbf{v}) = (\mathbf{M}_C^B(\text{id}))^{-1} \cdot k_C(\mathbf{v}), \quad (3.5)$$

d.h.  $(\mathbf{M}_C^B(\text{id}))^{-1} = \mathbf{M}_B^C(\text{id})$  ist die Basiswechsellmatrix von der Basis  $\mathbf{C}$  zur Basis  $\mathbf{B}$ .

- Umgekehrt sei  $\mathbf{A} = (a_{ij})$  eine invertierbare  $n \times n$ -Matrix und  $\mathbf{C}$  eine Basis von  $V$ . Es sei  $\mathbf{v}_j \in V$  der Vektor mit den  $\mathbf{C}$ -Koordinaten  $a_{ij}$  ( $j$ -te Spalte von  $\mathbf{A}$ ). Dann ist  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$  eine Basis von  $V$ , und wir erhalten die Basiswechsellmatrix  $\mathbf{M}_C^B(\text{id}) = \mathbf{A}$ .

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

$$\mathbf{M}_S^B(\text{id}) = \begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix}.$$

Sei  $k_B(\mathbf{v}) = (1, 3)^T$  die  $\mathbf{B}$ -Koordinaten von  $\mathbf{v}$ . Welcher Vektor ist dies?

$$k_S(\mathbf{v}) = \mathbf{M}_S^B(\text{id}) \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 10 \end{pmatrix}.$$

- b) Übergang  $\mathbf{S} \rightarrow \mathbf{B}$ . Aus (3.5) folgt:

$$\mathbf{M}_B^S(\text{id}) = (\mathbf{M}_S^B(\text{id}))^{-1} = \frac{1}{10} \begin{pmatrix} 2 & 1 \\ -4 & 3 \end{pmatrix}.$$

Also sind die  $\mathbf{B}$ -Koordinaten von  $\mathbf{v} = (1, 1)^T$

$$k_B(\mathbf{v}) = \frac{1}{10} \begin{pmatrix} 2 & 1 \\ -4 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 3 \\ -3 \end{pmatrix}.$$

## 7.4 Matrizen und lineare Abbildungen

**7.4.1 Definition.** 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 7.3.17 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  $\mathbf{A}$  wieder 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

$$\begin{aligned} \text{Spaltenraum von } \mathbf{A} &= \{\mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n\}, \\ \text{Zeilenraum von } \mathbf{A} &= \{\mathbf{y}\mathbf{A}, \mathbf{y} \in \mathbb{R}_m\} = \{\mathbf{y}^T \mathbf{A}, \mathbf{y} \in \mathbb{R}^m\}. \end{aligned} \quad (4.1)$$

**7.4.2 Satz.** 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.

- Man kann statt Zeilenumformungen auch Spaltenumformungen durchführen und erhält so die sogenannte „Spaltenstufenform“, z.B.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ * & * & * & 0 & 0 \end{pmatrix}$$

**7.4.3 Folgerung** aus Satz 7.2.10. *Der Zeilenraum einer Matrix ändert sich nicht bei elementaren Zeilenumformungen, der Spaltenraum nicht bei elementaren Spaltenumformungen.*

**7.4.4 Folgerung.** *In einer Matrix in Zeilenstufenform bilden die nicht-trivialen Zeilenvektoren eine Basis des Zeilenraumes. In einer Matrix in „Spaltenstufenform“ bilden die nichttrivialen Spaltenvektoren eine Basis des Spaltenraumes. Die Dimension ist jeweils 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.

#### 7.4.5 Satz.

- Rangatz. Die Dimension des Spaltenraumes von  $\mathbf{A}$  ist gleich der Dimension des Zeilenraumes von  $\mathbf{A}$ .*
- P-Q-Normalform. Es gibt invertierbare Matrizen  $\mathbf{P} \in \mathbb{R}^{m \times m}$ ,  $\mathbf{Q} \in \mathbb{R}^{n \times n}$ , so dass*

$$\mathbf{PAQ} = \begin{pmatrix} \mathbf{E}_r & 0 \\ 0 & 0 \end{pmatrix}, \quad (4.2)$$

wobei  $r = \dim$  (Zeilenraumes von  $\mathbf{A}$ )

- 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  $\mathbf{A}$ .*

- Eine andere Methode zur Bestimmung von  $\text{Ker } \mathbf{A}$  haben wir in Satz 7.1.10 kennengelernt.

**7.4.6 Definition.** *Der **Rang** einer Matrix ist die Dimension ihres Zeilenraumes und wird mit  $\text{Rang } \mathbf{A}$  bezeichnet.*

**7.4.7 Folgerung** Dimensionsformel. *Für jede  $m \times n$  Matrix  $\mathbf{A}$  gilt:*

$$\text{Rang } \mathbf{A} + \dim(\text{Ker } \mathbf{A}) = n.$$

**7.4.8 Folgerung.** 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  $\mathbf{P} \in \mathbb{R}^{m \times m}$ ,  $\mathbf{Q} \in \mathbb{R}^{n \times n}$

$$\text{Rang}(\mathbf{PAQ}) = \text{Rang}(\mathbf{A}). \quad (\text{ii})$$

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.

**7.4.9 Definition.** 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

$$f(\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_n f(\mathbf{v}_n) \quad \forall \alpha_i \in \mathbb{R}, \mathbf{v}_i \in V. \quad (4.3)$$

### Beispiele:

- Die Multiplikation mit einer Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  liefert die lineare Abbildung  $\ell_{\mathbf{A}} : \mathbb{R}^n \rightarrow \mathbb{R}^m : \mathbf{x} \rightarrow \mathbf{A} \cdot \mathbf{x}$  (siehe Satz 7.2.3 b), c))
- Der Differentialoperator  $\frac{d}{dx} : C^1(\mathbb{R}) \rightarrow C^0(\mathbb{R}) : f \mapsto f'$  ist linear (siehe Kapitel 3, Satz 3.2.5).
- 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 4.1.6)
- Sei  $V$  ein Vektorraum mit Basis  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ , dann ist die Koordinatenabbildung  $k_{\mathbf{B}} : V \rightarrow \mathbb{R}^n$  aus Definition 7.3.12 linear.

**7.4.10 Definition.** Seien  $U, V, W$  Vektorräume,  $f, g: V \rightarrow W$  und  $h: U \rightarrow V$  lineare Abbildungen und  $\alpha \in \mathbb{R}$ . Wir definieren die Abbildungen

$$f + g: V \rightarrow W, \quad \alpha f: V \rightarrow W,$$

punktweise, das heißt,

$$\begin{aligned} (f + g)(\mathbf{v}) &:= f(\mathbf{v}) + g(\mathbf{v}), \\ (\alpha f)(\mathbf{v}) &:= \alpha f(\mathbf{v}). \end{aligned}$$

Die **Komposition** zweier linearer Abbildungen definieren wir durch

$$f \circ h: U \rightarrow W \quad (f \circ h)(\mathbf{u}) := f(h(\mathbf{u})) .$$

- **Achtung:** Man kann  $f \circ h$  nur bilden, wenn der Bildraum von  $h$  der Definitionsbereich von  $f$  ist.
- Man überprüft leicht, dass die Abbildungen  $f + g$ ,  $\alpha f$  und  $f \circ h$  wieder linear sind.
- Mit den Operationen  $+$ ,  $\cdot$  bildet die Menge

$$\text{Hom}(V, W) := \{f: V \rightarrow W \text{ linear}\}$$

der **Homomorphismen** von  $V$  nach  $W$  wieder einen Vektorraum.

**7.4.11 Satz.** Sei  $f: V \rightarrow W$  eine lineare Abbildung, sei  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$  eine Basis von  $V$  und  $\mathbf{C} = (\mathbf{w}_1, \dots, \mathbf{w}_m)$  eine Basis von  $W$ . Für die Matrix  $\mathbf{M}_C^B(f) := (k_C(f(\mathbf{v}_1)), \dots, k_C(f(\mathbf{v}_n))) \in \mathbb{R}^{m \times n}$ , deren Spalten die  $\mathbf{C}$ -Koordinaten der Bilder der Basisvektoren  $\mathbf{v}_i$  sind, gilt

$$\mathbf{M}_C^B(f) \cdot k_B(\mathbf{v}) = k_C(f(\mathbf{v})) \quad (4.4)$$

für alle  $\mathbf{v} \in V$ . Man nennt  $\mathbf{M}_C^B(f)$  die **Abbildungsmatrix** von  $f$  bezüglich der Basen  $B$  und  $C$ .

- Im Spezialfall  $W = V$  und  $f = \text{id}$  erhalten wir die Basiswechsellmatrizen  $\mathbf{M}_C^B(\text{id})$ ; vergleiche dazu (3.4) mit (4.4).
- Sei umgekehrt  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Mit (4.4) erhalten wir eine lineare Abbildung  $f: V \rightarrow W$  mit  $\mathbf{M}_C^B(f) = \mathbf{A}$ . Also können wir lineare Abbildungen genauso durch Matrizen beschreiben, wie wir Vektoren in  $V$  mit Hilfe der Koordinatenabbildungen aus Definition 7.3.12 durch Spaltenvektoren im  $\mathbb{R}^n$  beschreiben können.

- Wir betrachten den Spezialfall  $V = \mathbb{R}^n$ ,  $W = \mathbb{R}^m$  mit den Standardbasen  $\mathbf{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  beziehungsweise  $\mathbf{C} = (\mathbf{e}_1, \dots, \mathbf{e}_m)$ . In diesem Fall erhalten wir die Abbildungsmatrix

$$\mathbf{M}(f) = \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f) = (f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)) \in \mathbb{R}^{m \times n},$$

und es gilt

$$f(\mathbf{x}) = \mathbf{M}(f) \cdot \mathbf{x}.$$

**7.4.12 Lemma.** *Seien  $f, g: V \rightarrow W$  und  $h: U \rightarrow V$  lineare Abbildungen, sei  $\alpha \in \mathbb{R}$ , und seien  $\mathbf{A}, \mathbf{B}$  und  $\mathbf{C}$  Basen von  $U, V$  beziehungsweise  $W$ . Dann gilt für die Abbildungsmatrizen*

$$(i) \quad \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(\alpha f) = \alpha \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f),$$

$$(ii) \quad \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f + g) = \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f) + \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(g),$$

$$(iii) \quad \mathbf{M}_{\mathbf{C}}^{\mathbf{A}}(f \circ h) = \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f) \cdot \mathbf{M}_{\mathbf{B}}^{\mathbf{A}}(h),$$

- (iv) *Ist  $f$  **invertierbar**, d.h. zu jedem  $\mathbf{y} \in \mathbb{R}^m$  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  $m = n$  und*

$$\mathbf{M}_{\mathbf{B}}^{\mathbf{C}}(f^{-1}) = (\mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f))^{-1}.$$

- Es sei  $\mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f)$  die Abbildungsmatrix von  $f$  zu den Basen  $\mathbf{B}, \mathbf{C}$  wie oben. Sei  $\mathbf{S}$  eine weitere Basis von  $V$ , und sei  $\mathbf{T}$  eine weitere Basis von  $W$ . Wir erhalten die Basiswechselmatrizen  $\mathbf{M}_{\mathbf{B}}^{\mathbf{S}}(\text{id}_V)$  und  $\mathbf{M}_{\mathbf{T}}^{\mathbf{C}}(\text{id}_W)$ . Aus Lemma 7.4.12 (iii) ergibt sich die Abbildungsmatrix von  $f$  zu den Basen  $\mathbf{S}$  und  $\mathbf{T}$  als

$$\mathbf{M}_{\mathbf{T}}^{\mathbf{S}}(f) = \mathbf{M}_{\mathbf{T}}^{\mathbf{C}}(\text{id}_W) \cdot \mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f) \cdot \mathbf{M}_{\mathbf{B}}^{\mathbf{S}}(\text{id}_V).$$

**7.4.13 Definition.** *Es sei  $f: V \rightarrow W$  eine lineare Abbildung. Dann sind **Kern**, **Bild** und **Rang** von  $f$  gegeben durch*

$$\begin{aligned} \text{Ker}(f) &= \{ \mathbf{v} \in V \mid f(\mathbf{v}) = 0 \} && \subset V, \\ f(V) &= \{ f(\mathbf{v}) \mid \mathbf{v} \in V \} && \subset W, \\ \text{Rang}(f) &= \dim(f(V)). \end{aligned}$$

Kern und Bild von  $f$  sind jeweils Untervektorräume.

**7.4.14 Folgerung** aus Satz 7.4.5. *Es seien  $V$  und  $W$  endlich dimensionale Vektorräume und  $f: V \rightarrow W$  linear. Dann gilt*

- a)  $\text{Rang}(f) = \text{Rang}(\mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f))$  für alle Basen  $\mathbf{B}$  von  $V$  und  $\mathbf{C}$  von  $W$ .
- b)  $\dim(\text{Ker}(f)) = \dim V - \text{Rang}(f)$ .
- c) Sei  $s = \text{Rang}(f)$ . Es gibt Basen  $\mathbf{B}$  von  $V$  und  $\mathbf{C}$  von  $W$ , so dass

$$\mathbf{M}_{\mathbf{C}}^{\mathbf{B}}(f) = \begin{pmatrix} \mathbf{E}_s & 0 \\ 0 & 0 \end{pmatrix}.$$

## 7.5 Determinanten und Eigenwerte

**7.5.1** Zweireihige Determinanten. *Das von zwei Vektoren  $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ ,  $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$  in der Ebene  $\mathbb{R}^2$  erzeugte Parallelogramm hat den Flächeninhalt (siehe Übungen)*

$$F = |a_1 b_2 - a_2 b_1|.$$

Die Zahl

$$\det \mathbf{A} := a_1 b_2 - a_2 b_1 \tag{5.1}$$

heißt **Determinante** der Matrix  $\mathbf{A} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ . Man berechnet sie nach dem Schema

$$\det \begin{pmatrix} \overset{+}{a_1} & b_1 \\ a_2 & \underset{-}{b_2} \end{pmatrix} \begin{matrix} \nearrow \\ \searrow \end{matrix} = a_1 b_2 - a_2 b_1,$$

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

$$\begin{aligned} \det \mathbf{A} = 0 &\Leftrightarrow \mathbf{a}, \mathbf{b} \text{ sind linear abhängig} \\ &\Leftrightarrow \mathbf{A} \text{ ist nicht invertierbar.} \end{aligned}$$

**7.5.2** Dreireihige Determinanten. *Der Spat dreier Vektoren  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$  hat das Volumen (siehe Übungen)*

$$\begin{aligned} V &= |[\mathbf{a}, \mathbf{b}, \mathbf{c}]| := |\mathbf{a}^T \cdot (\mathbf{b} \times \mathbf{c})| \\ &= |a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)| \end{aligned}$$

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 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  $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{3 \times 3}$

$$\det \mathbf{A} = a_{11} \det \begin{pmatrix} a_{12} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}, \quad (5.2)$$

d.h. die Berechnung einer Determinante einer  $3 \times 3$  Matrix kann auf die Berechnung von Determinanten von  $2 \times 2$  Matrizen zurückgeführt werden.

**7.5.3 Definition.** Sei  $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times n}$ . Wir definieren die **Determinante**  $\det \mathbf{A}$  rekursiv durch.

$$n = 1 : \det \mathbf{A} := a_{11}$$

$$\begin{aligned} n \geq 2 : \det \mathbf{A} &:= a_{11} \det \mathbf{A}_{11} - a_{21} \det \mathbf{A}_{21} + \cdots + (-1)^{n+1} a_{n1} \det \mathbf{A}_{n1}, \\ &\text{wobei } \mathbf{A}_{i1} \text{ die } (n-1) \times (n-1) \text{ Matrix ist, die aus } \mathbf{A} \text{ durch Entfernen der 1. Spalte und der } i\text{-ten Zeile entsteht.} \end{aligned}$$

**Beispiel:**

$$\begin{aligned}
\det \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 6 & 3 & 2 \\ 2 & 4 & 3 & 1 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} -2 & 1 & 0 \\ 6 & 3 & 2 \\ 4 & 3 & 1 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ 6 & 3 & 2 \\ 4 & 3 & 1 \end{pmatrix} \\
&+ 0 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 0 \\ 6 & 3 & 2 \end{pmatrix} \\
&= -2(3 \cdot 1 - 2 \cdot 3) - 6(1 \cdot 1 - 0 \cdot 3) + 4(1 \cdot 2 - 0 \cdot 3) \\
&\quad - 3[2(3 \cdot 1 - 2 \cdot 3) - 6(0 \cdot 1 - 1 \cdot 3) + 4(0 \cdot 2 - 1 \cdot 3)] \\
&\quad - 2[2(1 \cdot 2 - 0 \cdot 3) + 2(0 \cdot 2 - 1 \cdot 3) + 6(0 \cdot 0 - 1 \cdot 1)] \\
&= -46
\end{aligned}$$

- Wenn man die rekursive Definition 7.5.3 vollständig entwickelt erhält man:

$$\det \mathbf{A} = \sum_{i=i_1, \dots, i_n} (-1)^{\varepsilon(i)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}, \quad (5.3)$$

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 verwandeln. Man benötigt also  $n!$  Summanden und  $(n-1) \cdot n!$  Produkte. Wenn man Determinanten von Untermatrizen zwischenspeichert, kann man den Rechenaufwand deutlich reduzieren auf  $n(2^{n-1} - 1)$  Produkte, was aber immer noch sehr viel ist.

**7.5.4 Satz.** 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:  $\det \mathbf{E} = 1$ ,  $\det(\alpha \mathbf{E}) = \alpha^n$ .

**7.5.5 Satz.** 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.

**7.5.6 Satz.** 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:

$$\begin{aligned} \text{Typ 1:} & \quad \det \tilde{\mathbf{A}} = - \det \mathbf{A}, \\ \text{Typ 2:} & \quad \det \tilde{\mathbf{A}} = \alpha \det \mathbf{A} \quad \alpha \neq 0, \\ \text{Typ 3:} & \quad \det \tilde{\mathbf{A}} = \det \mathbf{A}. \end{aligned} \tag{5.4}$$

b) Es gilt:

$$\det \mathbf{A} = \det \mathbf{A}^T,$$

insbesondere gilt Satz 7.5.5 auch für Spalten.

c) Es gilt:

$$\det(\mathbf{A} \cdot \mathbf{B}) = \det \mathbf{A} \cdot \det \mathbf{B}. \tag{5.5}$$

d) Es gilt:

$$\begin{aligned} \mathbf{A} \text{ invertierbar} & \Leftrightarrow \det \mathbf{A} \neq 0, \\ \text{Rang } \mathbf{A} < n & \Leftrightarrow \det \mathbf{A} = 0. \end{aligned}$$

Wegen (5.4) können wir Determinanten auch nach dem Gauß-Verfahren aus 7.1.8 bestimmen, mit folgenden Variationen.

1'. Nächste Spalte. Falls  $a_{i,i} = \dots = a_{n,i} = 0$ , ist  $\det \mathbf{A} = 0$ , also **Ende**.

- 2'. Vertauschen. Wenn zwei Zeilen vertauscht werden, ändere das Vorzeichen vor der Determinante.
- 3'. Normieren. Schreibe zusätzlich den Faktor  $a_{i,i}$  vor die Determinante.
- 4'. Ausräumen. Wir müssen wegen Satz 7.5.4 nur unterhalb der Diagonalen ausräumen.
- 5'. Schleife. Es bietet sich an, die letzte  $2 \times 2$ -Determinante direkt wie in 7.5.1 auszurechnen.

Dabei brauchen wir nach dem  $i$ -ten Durchlauf die Zeilen und Spalten 1 bis  $i$  nicht weiter aufzuschreiben. Der Rechenaufwand für eine  $n \times n$ -Matrix beträgt jetzt nur noch  $\frac{n^3+2n}{3} - 2$  Multiplikationen und Divisionen. Ab  $n \geq 4$  lohnt sich dieses Verfahren.

**Beispiel:** Wir berechnen mit dem Gauß-Verfahren

$$\begin{aligned}
 & \det \begin{pmatrix} 0 & 2 & 4 & 6 \\ 3 & 6 & 3 & 0 \\ 2 & 2 & 3 & 3 \\ 1 & 0 & 2 & 1 \end{pmatrix} \\
 &= -3 \cdot \det \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & 4 & 6 \\ 0 & -2 & 1 & 3 \\ 0 & -2 & 1 & 1 \end{pmatrix} \\
 &= -6 \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 9 \\ 0 & 5 & 7 \end{pmatrix} \\
 &= -6 \cdot (5 \cdot 7 - 5 \cdot 9) = 60.
 \end{aligned}$$

**7.5.7 Folgerung.** *Es gelten für  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ :*

- a)  $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{B} \cdot \mathbf{A})$ ,
- b)  $\det(\mathbf{A}^k) = (\det \mathbf{A})^k$ ,  $k \in \mathbb{N}$ ,
- 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 Blockmatrizen  $\mathbf{B} \in \mathbb{R}^{k \times k}, \mathbf{D} \in \mathbb{R}^{(n-k) \times (n-k)}$ , so gilt:

$$\det \mathbf{A} = \det \mathbf{B} \cdot \det \mathbf{D}.$$

**7.5.8.** 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 7.5.3 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  $\det \mathbf{A}$  für  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mit  $n \geq 3$  sollte man die Struktur der Matrix beachten (evtl. Blockmatrix / Zeilen oder Spalten mit vielen Nullen) 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  $\mathbf{A}\mathbf{x} = \mathbf{b}$  die Lösung  $\mathbf{x} = (x_1, \dots, x_n)^T$  mit

$$x_i = \frac{1}{\det \mathbf{A}} \det (\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n).$$

- Für invertierbare Matrizen gilt

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \left( (-1)^{i+j} \det \mathbf{A}_{ji} \right).$$

Beide Formeln sind für praktische Rechnungen zu aufwändig. Sie sind dennoch für manche theoretische Überlegung hilfreich.

**7.5.9 Definition.** *Es sei  $V$  ein Vektorraum. Eine **Endomorphismus** von  $V$  ist eine lineare Abbildung  $f: V \rightarrow V$ , und wir schreiben*

$$\text{End}(V) = \text{Hom}(V, V).$$

- Seien  $f, g \in \text{End}(V)$  und  $\alpha \in \mathbb{R}$ . Wie in Definition 7.4.10 erhalten wir

$$f + g, \quad \alpha f, \quad f \circ g \in \text{End}(V).$$

Mit den Operationen  $+$ ,  $\cdot$  bildet  $\text{End } V$  einen Vektorraum.

- Sei  $\mathbf{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$  eine Basis von  $V$ . Wie in Satz 7.4.11 erhalten wir die Abbildungsmatrix  $\mathbf{M}_B^B(f) \in \mathbb{R}^{n \times n}$ . Mit den Koordinatenabbildungen aus Definition 7.3.12 erhalten wir

$$k_B(f(\mathbf{v})) = \mathbf{M}_B^B(f) \cdot k_B(\mathbf{v}).$$

- Mit Abbildungsmatrizen kann man rechnen:

$$\mathbf{M}_B^B(f + g) = \mathbf{M}_B^B(f) + \mathbf{M}_B^B(g),$$

$$\mathbf{M}_B^B(\lambda f) = \lambda \mathbf{M}_B^B(f),$$

$$\mathbf{M}_B^B(f \circ g) = \mathbf{M}_B^B(f) \cdot \mathbf{M}_B^B(g).$$

**7.5.10 Folgerung** aus Lemma 7.4.12. *Es seien  $\mathbf{B}$  und  $\mathbf{S}$  Basen von  $V$  mit Basiswechselmatrix  $\mathbf{A} = \mathbf{M}_S^B(\text{id})$  und  $f \in \text{End } V$ . Dann gilt*

$$\mathbf{M}_S^S(f) = \mathbf{M}_S^B(\text{id}) \cdot \mathbf{M}_B^B(f) \cdot \mathbf{M}_B^S(\text{id}) = \mathbf{A} \cdot \mathbf{M}_B^B(f) \cdot \mathbf{A}^{-1}.$$

**7.5.11 Definition.** Es sei  $V$  ein endlich-dimensionaler Vektorraum und  $f \in \text{End } V$ . Die **Determinante** von  $f$  ist definiert durch

$$\det f = \det(\mathbf{M}_B^B(f)).$$

**7.5.12 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  $|\det f| = |\det(\mathbf{M}(f))|$  die Volumenverzerrung der linearen Abbildung an.

- Genauso gibt  $\det f$  für  $f \in \text{End } V$  die Volumenverzerrung der Abbildung  $f$  an — diese hängt nicht davon, mit welcher Koordinatenabbildung  $k_B: V \rightarrow \mathbb{R}^n$  wir  $V$  mit  $\mathbb{R}^n$  identifizieren, um einen Volumenbegriff auf  $V$  zu erhalten.
- Eine invertierbare Matrix  $\mathbf{A}$  heißt *orientierungserhaltend*, wenn  $\det \mathbf{A} \geq 0$ , und *orientierungsumkehrend*, wenn  $\det \mathbf{A} < 0$ . Genauso verhält es sich mit invertierbaren Endomorphismen.

**7.5.13 Definition.** 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

$$\mathbf{C} = \mathbf{B}^{-1} \mathbf{A} \mathbf{B}.$$

Eine zu einer Diagonalmatrix ähnliche Matrix heißt **diagonalisierbar**. Ein Endomorphismus  $f \in \text{End } V$  heißt **diagonalisierbar**, wenn es eine Basis  $\mathbf{B}$  gibt, so dass  $\mathbf{M}_B^B(f)$  eine Diagonalmatrix ist.

- Nach Folgerung 7.5.10 sind Abbildungsmatrizen eines festen Endomorphismus  $f \in \text{End } V$  zu verschiedenen Basen zueinander ähnlich.
- Ein Endomorphismus  $f \in \text{End } V$  ist genau dann diagonalisierbar, wenn seine Abbildungsmatrix  $\mathbf{M}_S^S(f)$  zu einer beliebigen Basis  $\mathbf{S}$  diagonalisierbar ist. Verstehe dazu die Matrix  $\mathbf{B}$  als Basiswechselmatrix.

**Beispiel:** Sei  $s$  eine Spiegelung an der zu  $\mathbf{a}$ ,  $\|\mathbf{a}\| = 1$ , orthogonalen Ebene. Dann ist

$$s(\mathbf{x}) = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{a})\mathbf{a},$$

$$\mathbf{M}_S^S(s) = \begin{pmatrix} 1 - 2a_1^2 & -2a_2a_1 & -2a_1a_3 \\ -2a_1a_2 & 1 - 2a_2^2 & -2a_3a_2 \\ -2a_1a_3 & -2a_2a_3 & 1 - 2a_3^2 \end{pmatrix}.$$

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  $\mathbf{M}_B^B(s) = \mathbf{B}^{-1}\mathbf{M}_S^S(s)\mathbf{B}$ , das heißt, die Spiegelung  $s$  ist diagonalisierbar.

- Um eine Matrix oder Abbildung zu diagonalisieren, muss man eine Basis  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  geschickt wählen. Beispielsweise gilt

$$\mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & * & \cdots \\ 0 & \ddots & * \\ 0 & * & \ddots \end{pmatrix} \iff f(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1,$$

$$\mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \ddots \end{pmatrix} \iff f(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1, f(\mathbf{b}_2) = \lambda_2 \mathbf{b}_2.$$

Analog für Matrizen:

$$\mathbf{B}^{-1} \cdot da \cdot \mathbf{B} = \begin{pmatrix} \lambda_1 & * & \cdots \\ 0 & \ddots & * \\ 0 & * & \ddots \end{pmatrix} \iff \mathbf{A} \cdot \mathbf{b}_1 = \lambda_1 \mathbf{b}_1,$$

und so weiter. **Frage:** 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 (mit Ausnahme der Bedeutung des Vorzeichens der Determinante) bleiben wahr.

**7.5.14 Definition.** Eine Zahl  $\lambda \in \mathbb{C}$  heißt **Eigenwert** einer Matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , wenn es einen Spaltenvektor  $\mathbf{b} \in \mathbb{C}^n$ ,  $\mathbf{b} \neq \mathbf{0}$ , gibt mit

$$\mathbf{A} \cdot \mathbf{b} = \lambda \mathbf{b}. \quad (5.6)$$

Jeder Vektor  $\mathbf{b} \neq \mathbf{0}$ , der (5.6) erfüllt, heißt **Eigenvektor** von  $\mathbf{A}$  zum Eigenwert  $\lambda$ .

Genauso kann man Eigenwerte für Endomorphismen  $f \in \text{End}(V)$  definieren, aber nicht für  $f \in \text{Hom}(V, W)$  wenn  $V \neq W$ .

**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  $\mathbf{M}_S^S(s)$ .

**7.5.15 Definition.** Das charakteristische Polynom einer Matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$  ist definiert durch

$$\chi_{\mathbf{A}}(\lambda) := \det(\lambda \mathbf{E} - \mathbf{A}) \quad (5.7)$$

- Die Determinante in (5.7) ist explizit zu berechnen, wobei  $\lambda$  eine Variable ist. Hier eignet sich das Gauß-Verfahren nicht, also sollte man lieber nach geeigneten Zeilen oder Spalten entwickeln. 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 die **Spur** von  $\mathbf{A}$  definiert ist durch

$$\text{Sp}(\mathbf{A}) := \sum_{i=1}^n a_{ii}. \quad (5.8)$$

**7.5.16 Satz.** Es ist  $\lambda \in \mathbb{C}$  genau dann Eigenwert von  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , wenn  $\lambda$  eine Nullstelle des charakteristischen Polynoms  $\chi_{\mathbf{A}}$  ist.

**7.5.17 Definition.** 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  $\lambda$  ist die Ordnung von  $\lambda$  als Nullstelle des charakteristischen Polynoms. Der **Eigenraum**  $E(\lambda)$  zum Eigenwert  $\lambda$  ist der Lösungsraum des homogenen Gleichungssystems  $(\lambda \mathbf{E} - \mathbf{A}) \mathbf{x} = 0$ , d.h.

$$E(\lambda) = \text{Ker}(\lambda \mathbf{E} - \mathbf{A}). \quad (5.9)$$

Die Dimension des Eigenraums heißt **geometrische Vielfachheit**  $m(\lambda)$  des Eigenwerts  $\lambda$ .

**7.5.18 Satz.** 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.

**7.5.19 Satz.** 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  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_r) \in \mathbb{C}^{n \times n}$

$$\mathbf{B}^{-1} \mathbf{A} \mathbf{B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}. \quad (5.10)$$

**7.5.20 Satz.** Eine Matrix besitzt genau dann  $n$  linear unabhängige Eigenvektoren, wenn die Summe der geometrischen Vielfachheiten der Eigenwerte gleich  $n$  ist, d.h.

$$\sum_{i=1}^k m(\lambda_i) = n. \quad (5.11)$$

**7.5.21 Folgerung.** Falls (5.11) gilt, ist  $\mathbf{A}$  ähnlich einer Diagonalmatrix, d.h. (5.10) gilt.

**7.5.22 Folgerung.** 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 (5.10) gilt.

- Durch die Identifizierung von linearen Abbildungen und Matrizen erhalten wir folgenden Satz.

**7.5.23 Satz.** 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:

$$\mathbf{M}_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

**7.5.24. 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  $\lambda_i$  die Eigenwerte von  $\mathbf{A}$  mit der algebraischen Vielfachheit  $k(\lambda_i) = k_i$  sind. **Achtung:** Für  $n \geq 5$  kann man die Nullstellen nur noch raten oder versuchen, durch Approximation zu finden.

b) Bestimme die Eigenräume  $\mathbf{E}(\lambda_i) = \text{Ker}(\lambda_i \mathbf{E} - \mathbf{A})$ ,  $i = 1, \dots, r$ , und berechne für jeden eine Basis. Die Dimension von  $\mathbf{E}(\lambda_i)$ ,  $i = 1, \dots, r$  ist die geometrische Vielfachheit  $m(\lambda_i)$  des Eigenwertes  $\lambda_i$ .

c) Überprüfe, ob  $k(\lambda_i) = m(\lambda_i)$  für alle  $i = 1, \dots, r$  gilt. Falls ja, so ist  $\mathbf{A}$  bezüglich der aus den Basen der Eigenräume  $\mathbf{E}(\lambda_i)$  zusammengesetzten Basis ähnlich einer Diagonalmatrix.

- Falls  $\mathbf{A} \in \mathbb{R}^{n \times n}$  und alle Eigenwerte reell sind, d.h.  $\lambda_i \in \mathbb{R}$ , dann gelten Satz 7.5.19, Folgerung 7.5.22 und Satz 7.5.23 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 7.5.24 berechnet werden.

**Beispiele:** Für reelle  $2 \times 2$ -Matrizen gibt es vier mögliche Fälle.

1)  $\lambda_1, \lambda_2 \in \mathbb{R}$ ,  $\lambda_1 \neq \lambda_2$ : Sei  $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ , dann gilt:

$$\chi_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 1 \end{pmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3).$$

Also sind  $\lambda_1 = -1$ , und  $\lambda_2 = 3$  die Eigenwerte von  $\mathbf{A}$ . Das System  $(-\mathbf{E} - \mathbf{A})\mathbf{v}_1 = \mathbf{0}$  hat die Lösung  $\mathbf{v}_1 = (1, -1)^T$ . Die Lösung von  $(3\mathbf{E} - \mathbf{A})\mathbf{v}_2 = \mathbf{0}$  ist  $\mathbf{v}_2 = (1, 1)^T$ . Also gilt:  $\dim \mathbf{E}(\lambda_i) = 1, i = 1, 2$ , und  $\mathbf{A}$  ist ähnlich einer Diagonalmatrix, d.h. mit  $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$  gilt:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{B} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}.$$

2)  $\lambda_1 = \lambda_2 =: \lambda \in \mathbb{R}, m(\lambda) = k(\lambda) = 2$ : Für die Matrix  $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$  gilt

$\chi_{\mathbf{A}}(\lambda) = (\lambda - 3)^2$ . Also ist  $\lambda = 3$  der einzige Eigenwert. Aus  $\lambda\mathbf{E} - \mathbf{A} = \mathbf{0}$  folgt  $k(\lambda) = 2$  und  $\mathbf{E}(\lambda) = \mathbb{R}^2$ , also  $m(\lambda) = k(\lambda)$ . Und in der Tat ist  $\mathbf{A}$  bereits in Diagonalgestalt.

3)  $\lambda_1 = \lambda_2 =: \lambda \in \mathbb{R}, k(\lambda) = 2, m(\lambda) = 1$ : 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  $\lambda = 3$  der einzige Eigenwert von  $\mathbf{A}$  und  $k(\lambda) = 2$ . Da  $\lambda\mathbf{E} - \mathbf{A} \neq \mathbf{0}$ , folgt  $m(\lambda) < k(\lambda)$ , und  $\mathbf{A}$  ist nicht ähnlich einer Diagonalmatrix.

Aber man kann  $\mathbf{A}$  mit folgendem Verfahren in eine einfachere Form bringen. Da

$$\lambda\mathbf{E} - \mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix},$$

ist  $\mathbf{v}_1 = (1, -1)^T \in \text{Ker}(\lambda\mathbf{E} - \mathbf{A})$  ein Eigenvektor. Sei  $\mathbf{v}_2$  eine Lösung von  $(\lambda\mathbf{E} - \mathbf{A})\mathbf{v}_2 = -\mathbf{v}_1$ , zum Beispiel  $\mathbf{v}_2 = (-1, 0)^T$ . Setze man  $\mathbf{B} = (\mathbf{v}_1, \mathbf{v}_2)$ , dann gilt:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{B} = - \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

Man sagt, dass  $\mathbf{B}^{-1}\mathbf{A}\mathbf{B}$  ein  $\lambda$ -Block der Länge zwei ist. Allgemein ist ein  $\lambda$ -Block der Länge  $k$  gegeben durch

$$\begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

- 4)  $\lambda_1 \in \mathbb{C} \setminus \mathbb{R}$ : da das charakteristische Polynom reell ist, ist auch  $\lambda_2 = \bar{\lambda}_1$  ein Eigenwert. Für  $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ -1 & 2 \end{pmatrix}$  haben wir

$$\chi_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} \lambda & -2 \\ 1 & \lambda - 2 \end{pmatrix} = \lambda^2 - 2\lambda + 2 = (\lambda - i - 1)(\lambda + i - 1).$$

Also sind  $\lambda_1 = 1 + i$  und  $\lambda_2 = 1 - i$  die Eigenwerte von  $\mathbf{A}$ . Eine Basis von  $\mathbf{E}(\lambda_1)$  berechnet sich durch

$$\begin{pmatrix} 1+i & -2 \\ 1 & i-1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i-1 \\ 1+i & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i-1 \\ 0 & 0 \end{pmatrix},$$

d.h.  $\mathbf{v}_1 = (1 - i, 1)^T$ . Für  $\mathbf{E}(\lambda_2)$  erhalten wir den komplex konjugierten Eigenvektor, denn

$$\mathbf{A} \cdot \bar{\mathbf{v}}_1 = \overline{\mathbf{A} \cdot \mathbf{v}_1} = \overline{\lambda_1 \mathbf{v}_1} = \lambda_2 \bar{\mathbf{v}}_1.$$

Also ist  $\mathbf{v}_2 = \bar{\mathbf{v}}_1 = (1 + i, 1)^T$ . Somit

$$\begin{aligned} \mathbf{B}^{-1} \mathbf{A} \mathbf{B} &= \frac{i}{2} \begin{pmatrix} 1 & -1-i \\ -1 & 1-i \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \in \mathbb{C}^{2 \times 2}. \end{aligned}$$

**7.5.25 Satz** 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  $\lambda_j$ -Blöcken.

- Zu einem Eigenwert  $\lambda$  können verschiedene  $\lambda$ -Blöcke auftreten.

$$\left( \begin{array}{ccc|ccc} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{array} \right) \begin{array}{l} \text{Länge 3} \\ \text{Länge 2} \\ \text{Länge 1} \end{array}$$

- Bis auf die Reihenfolge der Blöcke ist die diagonale Blockmatrix eindeutig bestimmt.

## 7.6 Skalarprodukt und orthogonale Abbildungen

**7.6.1 Definition.** 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  $\|\mathbf{x}\| = 1$ .

**7.6.2 Lemma.** 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  $\mathbf{x} \neq \mathbf{0}$
- 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)

**7.6.3 Definition.** 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  $\mathbb{R}^n$ .

**7.6.4 Definition.**

- 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.

$$\mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij} \quad \forall i, j = 1, \dots, n.$$

**7.6.5 Satz.** Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Dann sind äquivalent:

a)  $\mathbf{A}$  ist orthogonal,

b)  $\|\mathbf{Ax}\| = \|\mathbf{x}\|$  für alle  $\mathbf{x} \in \mathbb{R}^n$ ,

c)  $(\mathbf{Ax}) \cdot (\mathbf{Ay}) = \mathbf{x} \cdot \mathbf{y}$  für alle  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,

d) 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.  $\angle(f(\mathbf{x}), f(\mathbf{y})) = \angle(\mathbf{x}, \mathbf{y})$ .

**7.6.6** 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 sofort aus, dass

$$\begin{aligned} \|s(\mathbf{x})\|^2 &= \|\mathbf{x}\|^2 \\ s(s(\mathbf{x})) &= \mathbf{x} \end{aligned}$$

und daher gilt (siehe Satz 7.6.5 und Lemma 7.4.12)

$$\mathbf{M}(s) = \mathbf{M}(s)^{-1} = \mathbf{M}(s)^T,$$

d.h. die Abbildungsmatrix  $\mathbf{M}(s)$  ist eine symmetrische, orthogonale Matrix mit  $\det(\mathbf{M}(s)) = -1$ , für die gilt:

$$\mathbf{M}(s) = \mathbf{E} - 2\mathbf{a}\mathbf{a}^T.$$

**7.6.7** 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. 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

$$\mathbf{D}^T \cdot \mathbf{D} = \mathbf{E} \text{ und } \det \mathbf{D} = 1.$$

Man kann auch zeigen: wenn

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{E} \text{ und } \det \mathbf{A} = -1$$

gilt, gibt es eine Drehmatrix  $\mathbf{D}$  mit Drehachse  $\mathbf{a}$ , ohne Einschränkung  $\|\mathbf{a}\| = 1$ , so dass

$$\mathbf{A} = \mathbf{D} \cdot \mathbf{M}(s) = \mathbf{M}(s) \cdot \mathbf{D} ,$$

dabei ist  $\mathbf{M}(s)$  die obige Spiegelungsmatrix.

**7.6.8. 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.

Man sieht leicht, dass

$$\text{Lin}(\mathbf{c}_1, \dots, \mathbf{c}_r) = \text{Lin}(\mathbf{b}_1, \dots, \mathbf{b}_r) \quad r = 1, \dots, k .$$

- Sei  $(\mathbf{c}_1, \dots, \mathbf{c}_k)$  paarweise senkrecht und nicht  $\mathbf{0}$ . Dann sind  $(\mathbf{c}_1, \dots, \mathbf{c}_k)$  linear unabhängig.
- Sei  $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$  eine Orthonormalbasis des  $\mathbb{R}^n$ . Dann gilt für alle  $\mathbf{x} \in \mathbb{R}^n$ , daß

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{c}_1) \cdot \mathbf{c}_1 + \dots + (\mathbf{x} \cdot \mathbf{c}_n) \cdot \mathbf{c}_n .$$

Insbesondere wird die Koordinatenabbildung  $k_C$  aus Definition 7.3.12 gegeben durch

$$k_C(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{c}_1, \dots, \mathbf{x} \cdot \mathbf{c}_n)^T .$$

## 7.7 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  $f(\mathbf{x}_0) \stackrel{i=1}{=} \alpha_0$ ,  $\mathbf{a} := (D_1 f(\mathbf{x}_0), \dots, D_n f(\mathbf{x}_0))^T$  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.

**7.7.1 Definition.** 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**.

**7.7.2. 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 (3.4) 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  $\mathbf{B}$ .

**7.7.3 Definition.** 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  $\mathbb{R}^n$  mit der Eigenschaft, dass  $q(\mathbf{x})$  bezüglich der Basis  $\mathbf{B}$  rein quadratisch ist.

- Nach 7.7.2 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.

**7.7.4 Satz.** Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix. Dann gilt:

- Alle Eigenwerte von  $\mathbf{A}$  sind reell.
- Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.
- Algebraische und geometrische Vielfachheit jedes Eigenwertes sind gleich.

**7.7.5 Satz** 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 orthonormale 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

$$\mathbf{B}^T \mathbf{A} \mathbf{B} = \text{Diag} \left( \underbrace{\lambda_1, \dots, \lambda_1}_{k_1\text{-mal}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{k_r\text{-mal}} \right)$$

gilt. Demzufolge gilt

$$q(\mathbf{x}) = k_{\mathbf{B}}(\mathbf{x})^T \mathbf{B}^T \mathbf{A} \mathbf{B} k_{\mathbf{B}}(\mathbf{x}) = \lambda_1 (k_{\mathbf{B}}(\mathbf{x}))_1^2 + \dots + \lambda_r (k_{\mathbf{B}}(\mathbf{x}))_r^2.$$

Da  $\mathbf{B}$  eine Orthonormalbasis ist, lässt sich  $k_{\mathbf{B}}(\mathbf{x})_i = \mathbf{x}^T \cdot \mathbf{b}_i$  leicht bestimmen.

**7.7.6 Definition.** 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.

**7.7.7 Satz.**

- a)  $\mathbf{D} = \text{Diag}(\alpha_1, \dots, \alpha_n)$  ist genau dann positiv definit, wenn alle  $\alpha_i$  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.

**7.7.8. 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:

$$\mathbf{A} = \mathbf{A}^T \text{ positiv definit} \Rightarrow a_{ii} > 0, \quad i = 1, \dots, n. \quad (7.1)$$

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:

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ positiv definit} \Leftrightarrow a > 0, \det \mathbf{A} = ad - b^2 > 0 \quad (7.2)$$

**7.7.9 Satz.**  $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{n \times n}$ ,  $\mathbf{A} = (a_{ij})$  ist genau dann positiv definit, wenn die Determinanten der **Hauptuntermatrizen**  $\mathbf{H}_i$  positiv sind:

$$\mathbf{H}_1 = a_{11} \quad \mathbf{H}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \dots, \mathbf{H}_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{1k} & \cdots & a_{kk} \end{pmatrix}, \mathbf{H}_n = \mathbf{A}.$$

**7.7.10 Satz.**  $\mathbf{A} = \mathbf{A}^T$  ist genau dann negativ definit, wenn für die Determinanten der Hauptuntermatrizen gilt:

$$(-1)^k \det \mathbf{H}_k > 0.$$

- Diese Kriterien funktionieren nicht für semidefinite Matrizen. Beispielsweise sind bei

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

alle Hauptunterdeterminanten  $\geq 0$ , aber die Matrix ist offensichtlich nicht positiv semidefinit.

# Kapitel 8

## Differentiation von Funktionen in mehreren 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**,  $\mathbb{R}^m$  Bildraum und das **Bild von  $\mathbf{f}$**  ist

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

### 8.1 Kurven im $\mathbb{R}^n$

**8.1.1** Parameterdarstellung. *Kurven stellt man sich am besten als die Bahn eines bewegten Punktes vor, wobei man jedem Zeitpunkt  $t$  einen Ortsvektor  $\mathbf{x}(t)$  zuordnet. Sei also  $I \subseteq \mathbb{R}$  ein Intervall und  $\mathbf{x} : I \rightarrow \mathbb{R}^n$  eine gegebene vektorwertige Funktion. Die Funktion  $x_i : i \rightarrow \mathbb{R}, 1 \leq i \leq n$  heißt **Komponentenfunktion** von  $\mathbf{x}$ .*

**8.1.2 Definition.** *Sei  $\mathbf{x} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$  eine Funktion. Die Funktion  $\mathbf{x}$  hat im Punkt  $t_0 \in I$  den **Grenzwert  $\mathbf{c}$** , in Zeichen  $\lim_{t \rightarrow t_0} \mathbf{x}(t) = \mathbf{c}$ , genau dann, wenn*

$$\lim_{t \rightarrow t_0} x_i(t) = c_i, \quad 1 \leq i \leq n.$$

- *Der Grenzwert ist komponentenweise definiert und wird somit zurückgeführt auf Grenzwerte von Funktionen einer Variablen.*

- Analog heißt  $\mathbf{x} : I \rightarrow \mathbb{R}^n$  **stetig** bzw. **differenzierbar**, wenn alle Komponentenfunktionen stetig bzw. differenzierbar sind. Die Ableitung wird komponentenweise berechnet, d.h.

$$\dot{\mathbf{x}}(t) = \frac{d}{dt} \mathbf{x}(t) = \lim_{h \rightarrow 0} \frac{1}{h} (\mathbf{x}(t+h) - \mathbf{x}(t)) = \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}. \quad (1.1)$$

Man nennt  $\dot{\mathbf{x}}(t)$  den **Tangentenvektor** von  $\mathbf{x}$  an der Stelle  $t$ .

**8.1.3 Lemma.** Es gelten folgende Rechenregeln für  $\mathbf{x}, \mathbf{y} : I \rightarrow \mathbb{R}^n$  und  $\alpha : I \rightarrow \mathbb{R}$ :

- $\frac{d}{dt} (\alpha \mathbf{x}(t) + \beta \mathbf{y}(t)) = \alpha \dot{\mathbf{x}}(t) + \beta \dot{\mathbf{y}}(t)$ ,  $\alpha, \beta \in \mathbb{R}$ ,
- $\frac{d}{dt} (\mathbf{x}(t) \cdot \mathbf{y}(t)) = \dot{\mathbf{x}}(t) \cdot \mathbf{y}(t) + \mathbf{x}(t) \cdot \dot{\mathbf{y}}(t)$ ,
- $\frac{d}{dt} (\mathbf{x}(t) \times \mathbf{y}(t)) = \dot{\mathbf{x}}(t) \times \mathbf{y}(t) + \mathbf{x}(t) \times \dot{\mathbf{y}}(t)$ ,  $n = 3$ ,
- $\frac{d}{dt} (\alpha(t) \mathbf{x}(t)) = \dot{\alpha}(t) \mathbf{x}(t) + \alpha(t) \dot{\mathbf{x}}(t)$ .

- Aus  $\|\mathbf{x}(t)\|^2 = \mathbf{x}(t) \cdot \mathbf{x}(t)$  folgt:

$$\|\mathbf{x}(t)\| = \text{const.} \quad \Rightarrow \quad \mathbf{x}(t) \cdot \dot{\mathbf{x}}(t) = 0. \quad (1.2)$$

**8.1.4 Definition.** Sei  $G \subseteq \mathbb{R}^n$  und  $[a, b] \subseteq \mathbb{R}$ . Wir nennen jede stetig differenzierbare Funktion  $\mathbf{x} : [a, b] \rightarrow G$  eine **Parameterdarstellung** des Kurvenstücks  $\{\mathbf{x}(t); t \in [a, b]\}$  mit Anfangspunkt  $\mathbf{x}(a)$  und Endpunkt  $\mathbf{x}(b)$ . Ein Kurvenstück heißt **regulär**, wenn  $\dot{\mathbf{x}}(t) \neq \mathbf{0}$  für alle  $t \in [a, b]$  gilt.

- Eine **Kurve** ist eine Kette von Kurvenstücken. Diese muss in den Berührungspunkten nicht differenzierbar sein.
- Oft wird sowohl die Parameterdarstellung  $\mathbf{x}(t)$  als auch das Kurvenstück  $\{\mathbf{x}(t), t \in [a, b]\}$  einfach als Kurvenstück oder Kurve bezeichnet.

**8.1.5 Definition.** Sei  $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$  ein reguläres Kurvenstück. Dann heißt

$$s(t) := \int_a^t \|\dot{\mathbf{x}}(t)\| dt$$

die **Bogenlänge** des Kurvenstücks über  $[a, \tau]$ .

- Nach dem Hauptsatz der Differential- und Integralrechnung (Kapitel 4, Satz 4.1.10) gilt

$$\dot{s}(t) = \|\dot{\mathbf{x}}(t)\| . \quad (1.3)$$

- Analog zu Kapitel 4, (5.11) nennen wir

$$\begin{aligned} ds &:= \|\dot{\mathbf{x}}(t)\| dt && \text{das } \mathbf{Bogenelement} \text{ von } \mathbf{x}, \text{ und} \\ d\mathbf{x} &:= \dot{\mathbf{x}}(t)dt && \text{das } \mathbf{vektorielle Bogenelement} \text{ von } \mathbf{x}. \end{aligned}$$

**8.1.6 Definition.** Sei  $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$  stetig differenzierbar mit  $\dot{\mathbf{x}}(t) \neq \mathbf{0}$ . Dann heißt

$$\mathbf{T}(t) := \frac{\dot{\mathbf{x}}(t)}{\|\dot{\mathbf{x}}(t)\|} \quad \text{der } \mathbf{Tangenten(einheits)vektor}. \quad (1.4)$$

- Sei  $\mathbf{x}$  jetzt zweimal stetig differenzierbar. Der Grenzwert der Änderungsrate  $\frac{\Delta T}{\Delta s} = \frac{T(t_1) - T(t_2)}{s(t_1) - s(t_2)}$  des Tangentialvektors heißt **Krümmungsvektor** und ist durch  $\frac{\dot{\mathbf{T}}(t)}{\dot{s}(t)}$  gegeben. Seine Länge heißt Krümmung

$$\kappa(t) := \frac{\|\dot{\mathbf{T}}(t)\|}{\dot{s}(t)} = \frac{\|\dot{\mathbf{T}}(t)\|}{\|\dot{\mathbf{x}}(t)\|}. \quad (1.5)$$

**8.1.7 Lemma.** Für eine Kurve im  $\mathbb{R}^3$  haben wir die Darstellungen

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \dot{s}(t)\mathbf{T}(t), \\ \kappa(t) &= \frac{\|\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)\|}{\|\dot{\mathbf{x}}(t)\|^3} \end{aligned}$$

**Beispiel:** Eine Schraubenkurve hat die Darstellung

$$\mathbf{x}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ ht \end{pmatrix} .$$

Somit erhalten wir:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, & \ddot{\mathbf{x}}(t) &= \begin{pmatrix} -r \cos t \\ -r \sin t \\ 0 \end{pmatrix}, \\ \|\dot{\mathbf{x}}(t)\| &= \sqrt{r^2 + h^2} = \dot{s}(t) =: R, \\ \mathbf{T}(t) &= \frac{1}{R} \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, \\ \kappa(t) &= \frac{\|\dot{\mathbf{T}}\|}{\|\dot{\mathbf{x}}\|} = \frac{r}{R^2} = \frac{r}{r^2 + h^2}.\end{aligned}$$

## 8.2 Reellwertige Funktionen mehrerer Veränderlicher

**8.2.1 Grundlagen.** Sei  $D \subseteq \mathbb{R}^n$  und  $f : D \rightarrow \mathbb{R}$  eine reellwertige Funktion. Dieser Funktion ist ein **Graph**

$$\Gamma_f := \{(\mathbf{x}, f(\mathbf{x})); \mathbf{x} \in D\} \subseteq \mathbb{R}^{n+1}$$

zugeordnet. Zur Veranschaulichung von  $f$  betrachtet man **Niveaumengen**, d.h. für  $c \in \mathbb{R}$

$$\mathbf{N}_c := \{\mathbf{x} \in D; f(\mathbf{x}) = c\} \subseteq \mathbb{R}^n,$$

oder Schnitte mit zu Koordinatenachsen parallelen Geraden, d.h. man betrachtet die „partielle“ Funktion von  $f$ , die gegeben ist durch

$$x_i \rightarrow f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n),$$

mit  $\mathbf{a} = (a_1, \dots, a_n)^T \in D$ .

**8.2.2 Definition.** 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 (oder auch  $r$ -Ball) von  $\mathbf{a}$ .

**8.2.3 Definition.** Sei  $D$  eine Teilmenge des  $\mathbb{R}^n$ .

- a) 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.  $B_r(\mathbf{a}) \subseteq D$ .
- b)  $D$  heißt **offen**, wenn jeder Punkt von  $D$  ein innerer Punkt ist.
- c) 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  $\partial D$  bezeichnet.
- d) 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  $(x - a_1)^2 + (y - a_2)^2 = r^2$ .
- $C_r(\mathbf{a}) = \{(x, y)^T \in \mathbb{R}^2; (x - a_1)^2 + (y - a_2)^2 \leq r^2\}$  ist abgeschlossen.
- $D = B_r(\mathbf{a}) \cap \{(x, y)^T \in \mathbb{R}^2; x \geq 0\}$  weder offen noch abgeschlossen.

**8.2.4 Definition.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\mathbf{a} \in D \cup \partial D$ .

- a)  $f$  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, Produkte und Quotienten stetiger Funktionen sind stetig.
- Die Koordinatenfunktion  $p_i(\mathbf{x}) = x_i$  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  $f$ .

### 8.2.5 Definition.

a) Eine Menge  $D \subseteq \mathbb{R}^n$  heißt *beschränkt*, wenn es eine Konstante  $K > 0$  gibt mit

$$\|\mathbf{x}\| < K \quad \text{für alle } \mathbf{x} \in D.$$

b) Die abgeschlossenen und beschränkten Mengen des  $\mathbb{R}^n$  nennt man **kompakt**.

**8.2.6 Satz.** 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

$$f(\mathbf{a}) \leq f(\mathbf{x}) \leq f(\mathbf{b}) \quad \text{für alle } \mathbf{x} \in D.$$

**8.2.7 Satz.** 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 8.2.6 und 8.2.7 sind im Allgemeinen falsch, wenn  $D$  nicht kompakt ist.
- Sei  $f : D \rightarrow \mathbb{R}$  stetig, aber  $D$  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 8.2.6 und Satz 8.2.7 gelten.

- 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

$$f(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in D \cap B_r(\mathbf{a}).$$

**8.2.8 Definition.** 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.

$f$  heißt **partiell differenzierbar** bzw. **stetig (partiell) differenzierbar** wenn alle partiellen Ableitungen  $\frac{\partial f}{\partial x_i}$  existieren bzw. stetig sind.

- Die partielle Ableitung ist die Ableitung der „partiellen Funktionen“

$$x_i \rightarrow f(x_1, \dots, x_i, \dots, x_n),$$

alle anderen Argumente  $x_j, j \neq i$  werden konstant gehalten.

- Eine andere Bezeichnung ist  $f_{x_i}; f_x, f_y$  falls die Variablen  $x$  und  $y$  sind.
- **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. Die Begriffe  **$k$ -fach partiell differenzierbar** und  **$k$ -fach stetig differenzierbar** werden entsprechend definiert.

**8.2.9 Definition.** Sei  $f: D \rightarrow \mathbb{R}$  partiell differenzierbar an der Stelle  $\mathbf{x}$ . Der **Gradient** einer Funktion  $f$  ist der Vektor bestehend aus den partiellen Ableitungen, d.h.

$$\text{grad } f(\mathbf{x}) := \left( \frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)^T \in \mathbb{R}^n.$$

- Eine andere Bezeichnung ist:  $\nabla f(\mathbf{x})$ .
- Sei  $D \subseteq \mathbb{R}^n$  offen und  $k \in \mathbb{N}_0$ . Dann bezeichnen wir

$$C^k(D, \mathbb{R}) := \{f : D \rightarrow \mathbb{R}; f \text{ } k\text{-mal stetig partiell differenzierbar.}\}$$

Diese Menge ist ein Vektorraum.

- Im Allgemeinen gilt:  $\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$ .

**8.2.10 Satz.** Für jede  $C^2$ -Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen gilt:

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left( \frac{\partial f}{\partial x_i} \right), \quad 1 \leq i, j \leq n.$$

- Der Satz gilt analog für  $k$ -te gemischte partielle Ableitungen, falls  $f \in C^k(D, \mathbb{R})$ .

**8.2.11 Definition.** Für zwei Funktionen  $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\mathbf{x}_0 \in D$ ,  $k \in \mathbb{N}_0$  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 3.2.3 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.

**8.2.12 Definition.** 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  $\mathbf{a} \in \mathbb{R}^n$  mit

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|), \quad (2.1)$$

für  $\mathbf{x}$  nahe  $\mathbf{x}_0$ , gibt.

**8.2.13 Satz.** 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 und  $\mathbf{a}$  ist eindeutig bestimmt als  $\mathbf{a} = \text{grad } f(\mathbf{x}_0)$ .

- Wenn  $f$  total differenzierbar ist, so ist

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|) \quad (2.2)$$

die beste lineare Approximation von  $f$  nahe  $\mathbf{x}_0$ .

**8.2.14 Satz.** 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.

**8.2.15 Definition.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\mathbf{x}_0 \in D$ . 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. Wenn  $\mathbf{v} \in \mathbb{R}^n$  ein Einheitsvektor ist, d.h.  $\|\mathbf{v}\| = 1$ , dann heißt  $\partial_{\mathbf{v}} f(\mathbf{x}_0)$  **Richtungsableitung** von  $f$  an der Stelle  $\mathbf{x}_0$  in Richtung  $\mathbf{v}$ .

- Dies ist eine Verallgemeinerung von partiellen Ableitungen welche der Wahl  $\mathbf{v} = \mathbf{e}_i$  entspricht.
- $\partial_{\mathbf{v}} f(\mathbf{x}_0)$  beschreibt das Verhalten von  $f$  längs der Geraden  $\mathbf{x}_0 + t\mathbf{v}$ .

**8.2.16 Satz.** 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:

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{i=1}^n f_{x_i}(\mathbf{x}_0) v_i$$

• Wir hatten bereits gesehen, dass sich Funktionen längs verschiedener Kurven verschieden verhalten können (siehe Beispiel stetiger „partieller“ Funktionen). Deshalb ist es sinnvoll, auch Ableitungen längs von Kurven und nicht nur längs von Geraden (Richtungsableitungen) zu betrachten.

**8.2.17 Satz** (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:

$$\frac{d}{dt} f(\mathbf{x}(t)) = \text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t). \quad (2.3)$$

**8.2.18 Polarkoordinaten.** Im  $\mathbb{R}^2$  besteht folgende Beziehung zwischen den Polarkoordinaten  $(r, \varphi)$  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  $(r, \varphi)$  betrachtet werden. Die partiellen Ableitungen von  $F$  berechnen sich durch die Kettenregel als

$$\begin{aligned} F_r &= f_x \cos \varphi + f_y \sin \varphi, \\ F_\varphi &= -f_x r \sin \varphi + f_y r \cos \varphi. \end{aligned} \quad (2.4)$$

Dieses Gleichungssystem kann man nach  $f_x$  und  $f_y$  auflösen und erhält für  $r \neq 0$

$$\begin{aligned} f_x &= F_r \cos \varphi - \frac{1}{r} F_\varphi \sin \varphi, \\ f_y &= F_r \sin \varphi + \frac{1}{r} F_\varphi \cos \varphi. \end{aligned} \quad (2.5)$$

Es gibt Formeln auch für höhere partielle Ableitungen. Insbesondere haben wir

$$\Delta f := f_{xx} + f_{yy} = F_{rr} + \frac{1}{r} F_r + \frac{1}{r^2} F_{\varphi\varphi}, \quad (2.6)$$

welches die Formeln für den **Laplace - Operator**  $\Delta$  in kartesischen Koordinaten und in Polarkoordinaten sind. Anloge Formeln gelten für die Polar- oder Kugelkoordinaten im  $\mathbb{R}^3$ .

## 8.3 Anwendungen

**8.3.1** Richtung des stärksten Anstiegs. Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion und  $\mathbf{x}_0 \in D$ . Für die Richtungsableitung in Richtung  $\mathbf{v}$  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  $\alpha$  der Winkel zwischen  $\mathbf{v}$  und  $\text{grad } f(\mathbf{x}_0)$  ist. Somit ist der Wert der Richtungsableitung unter der Annahme  $\|\mathbf{v}\| = 1$  maximal, wenn  $\cos \alpha = 1$ , d.h.  $\mathbf{v}$  ist parallel zu  $\text{grad } f(\mathbf{x}_0)$ . Somit haben wir:

Richtung von  $\text{grad } f(\mathbf{x}_0)$  = Richtung des maximalen Anstiegs von  $f$  in  $\mathbf{x}_0$ .

**8.3.2** Gradientenverfahren. Zur Bestimmung des Maximums einer Funktion  $f : D \rightarrow \mathbb{R}$  wählt man einen Startwert  $\mathbf{x}_0$ , bestimmt dann die Richtung des maximalen Anstiegs  $\text{grad } f(\mathbf{x}_0)$  und berechnet einen neuen Punkt

$$\mathbf{x}_1 = \mathbf{x}_0 + h \text{grad } f(\mathbf{x}_0)$$

mit einer Schrittweite  $h$ . Falls  $f(\mathbf{x}_1) < f(\mathbf{x}_0)$  ist man zu weit gegangen und versucht es mit halber Schrittweite  $\frac{h}{2}$ . Falls  $f(\mathbf{x}_1) > f(\mathbf{x}_0)$  wiederholt man das Verfahren mit  $\mathbf{x}_1$  als Startwert. Auf diese Weise nähert man sich einer Maximalstelle von  $f$  in  $D$ , welche vom Startwert abhängt und nicht das globale Maximum sein muß. Analog für Minimum.

**8.3.3** Tangentialebene. Sei  $D \subseteq \mathbb{R}^n$  offen und  $f : D \rightarrow \mathbb{R}$  eine  $C^1$  Funktion. Die Niveaumenge  $\mathbf{N}_c = \{\mathbf{x} : f(\mathbf{x}) = c\}$  ist eine Hyperfläche. Für jede Kurve  $\mathbf{x} : [a, b] \rightarrow \mathbf{N}_c$  in  $\mathbf{N}_c$  gilt also  $f(\mathbf{x}(t)) = c$  und somit

$$\text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) = 0, \quad (3.1)$$

d.h. der Gradient  $\text{grad } f$  steht orthogonal zu allen Kurventangenten  $\dot{\mathbf{x}}$ .

Für  $n = 2$  ist also die Normalengleichung der Tangente an die Niveaukurve  $\mathbf{N}_c$  durch den Punkt  $(x_0, y_0) \in \mathbf{N}_c$  gegeben durch

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

Für  $n = 3$  spannen die Tangentialvektoren aller Niveaukurven durch einen Punkt  $(x_0, y_0, z_0) \in \mathbf{N}_c$  eine Ebene auf, die **Tangentialebene**. Die Normalengleichung der Tangentialebene ist

$$\text{grad } f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0.$$

- 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?

**8.3.4 Definition.** 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  $(x, y) \in D$  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.

**8.3.5 Satz** (über implizite Funktionen). Sei  $D \subseteq \mathbb{R}^2$  offen und  $f : D \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion. Ist  $(x_0, y_0) \in D$  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 dass gilt:

- $R = \{(x, y) : x \in I, y \in K\} \subseteq D$ ,  $f_y(x, y) \neq 0$  für alle  $(x, y) \in R$ .
- Durch  $f(x, y) = 0$  ist auf  $I$  eindeutig eine differenzierbare implizite Funktion  $g : I \rightarrow K$  erklärt, für deren Ableitung gilt:

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} = -\frac{f_x(x, y)}{f_y(x, y)} \quad \forall x \in I.$$

- Nachdem wir wissen, dass  $g$  differenzierbar ist, kann man die Formeln 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.

$$f_{xx} + f_{xy}g' + g_{yx}g' + f_{yy}(g')^2 + f_y g'' = 0.$$

Also gilt:

$$\begin{aligned} g'(x) &= \frac{-f_x}{f_y} \\ g''(x) &= -\frac{1}{f_y} (f_{xx} + f_{xy}g' + g_{yx}g' + f_{yy}(g')^2) \\ &= -\frac{1}{(f_y)^3} (f_{xx} (f_y)^2 - 2f_{xy}f_x f_y + f_{yy} (f_x)^2). \end{aligned}$$

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

$$\frac{f_{xx}}{f_y} > 0$$

und  $x$  ist ein lokales Minimum falls  $\frac{f_{xx}}{f_y} < 0$ .

**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 8.3.5 b) definierte Funktion, so gilt:

$$g'(x) = -\frac{f_x}{f_y} = -\frac{3x^2 + 2x}{e^y + 3y^2}.$$

Hat  $g$  an der Stelle  $x_0$  eine Extremstelle, d.h.  $g'(x_0) = 0$ , so gilt:

$$0 = \frac{f_x(x_0, g(x_0))}{f_y(x_0, g(x_0))} = \frac{3x_0^2 + 2x_0}{e^{y_0} + 3y_0^2}.$$

Dies ergibt  $x_0 = 0$  oder  $x_0 = -\frac{2}{3}$ . Wir haben

$$g''(x_0) = -\frac{f_{xx}(x_0, y_0)}{f_y(x_0, y_0)} = -\frac{6x_0 + 2}{e^{y_0} + 3y_0^2}$$

und also ist  $g''(0) = -2$  also ist 0 ein Maximum. Weiter gilt  $g''(-\frac{2}{3}) = 2$  und also ist  $-\frac{2}{3}$  ein Minimum.

- Ein Gebiet  $D$  heißt **konvex**, wenn  $D$  offen ist und gilt:  
 $\mathbf{x}, \mathbf{y} \in D \Rightarrow t\mathbf{y} + (1-t)\mathbf{x} \in D \quad \forall t \in [0, 1]$ .

**8.3.6 Satz** (Taylor-Formel). *Ist  $D \in \mathbb{R}^n$  ein konvexes Gebiet,  $f \in C^{k+1}(D)$ ,  $\mathbf{x} \in D$ , dann gilt mit  $\mathbf{x} + \mathbf{v} \in D$  die Taylorformel*

$$f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + \partial_{\mathbf{v}} f(\mathbf{x}) + \frac{1}{2!} \partial_{\mathbf{v}}^2 f(\mathbf{x}) + \cdots + \frac{1}{k!} \partial_{\mathbf{v}}^k f(\mathbf{x}) + R_k(\mathbf{x}, \mathbf{v})$$

mit dem Restglied

$$R_k(\mathbf{x}, \mathbf{v}) = \frac{1}{(k+1)!} \partial_{\mathbf{v}}^{k+1} f(\mathbf{x} + \xi_k \mathbf{v}) = \frac{1}{(k+1)!} \partial_{\mathbf{v}}^{k+1} f(\mathbf{x}) + o(\|\mathbf{v}\|^{k+1}),$$

wobei  $\xi_k \in (0, 1)$  und  $\partial_{\mathbf{v}} f = \sum_{i=0}^n v_i \frac{\partial f}{\partial x_i}$ .

- Sei  $f \in C^k(D)$ , dann ist

$$p(\mathbf{x}_0 + \mathbf{v}) := \sum_{i_1 + \cdots + i_n \leq k} \frac{\partial^{i_1 + \cdots + i_n} f}{\partial x_1^{i_1} \cdots \partial x_n^{i_n}}(\mathbf{x}_0) \frac{v_1^{i_1} \cdots v_n^{i_n}}{i_1! \cdots i_n!} \quad (3.2)$$

das **Taylor-Polynom** von  $f$  an der Stelle  $\mathbf{x}_0$ . Es gilt:

$$f(\mathbf{x}) = p(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{x}_0\|^k). \quad (3.3)$$

**8.3.7 Definition.** *Für eine  $C^2$ -Funktion  $f$  heißt die symmetrische Matrix*

$$H_f(\mathbf{x}) := \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \right)_{i,j=1,\dots,n}$$

die **Hesse-Matrix** von  $f$  an der Stelle  $\mathbf{x}$ .

**8.3.8 Folgerung.** Sei  $f: D \rightarrow \mathbb{R}$ , wobei  $D$  ein konvexes Gebiet ist und  $\mathbf{x}, \mathbf{x}_0 \in D$ . Dann gilt mit Punkten  $\tilde{\mathbf{x}}$  und  $\mathbf{x}^*$  zwischen  $\mathbf{x}$  und  $\mathbf{x}_0$ :

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\tilde{\mathbf{x}}) \cdot (\mathbf{x} - \mathbf{x}_0) && \text{falls } f \in C^1(D; \mathbb{R}), \\ f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0)^T \cdot (\mathbf{x} - \mathbf{x}_0) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)H_f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}_0) \\ &= f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0)^T \cdot (\mathbf{x} - \mathbf{x}_0) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^2) \quad \text{falls } f \in C^2(D; \mathbb{R}). \end{aligned}$$

**8.3.9 Definition.** 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.

**8.3.10 Satz** (Lokale Extrema im Innern). Ist  $f$  auf  $B_r(\mathbf{a})$ ,  $\mathbf{a} \in \mathbb{R}^n$ , eine  $C^1$ -Funktion, so gilt:

$$\mathbf{a} \text{ ist lokale Extremstelle} \Rightarrow \text{grad } f(\mathbf{a}) = 0.$$

- Ein Punkt  $\mathbf{a} \in \mathbb{R}^n$  für den  $\text{grad } f(\mathbf{a}) = 0$  gilt, heißt **stationärer Punkt**. Ein stationärer Punkt, der kein Extrempunkt ist, heißt **Sattelpunkt**.
- Satz 8.3.10 besagt, dass man Extrempunkte im Innern unter den stationären Punkten suchen muss.

**8.3.11 Satz.** Ist  $U \subseteq \mathbb{R}^n$  offen,  $\mathbf{a} \in U$  ein stationärer Punkt einer Funktion  $f \in C^2(U, \mathbb{R})$  und  $H_f(\mathbf{a}) = (f_{x_i x_j}(\mathbf{a}))$  die Hesse-Matrix von  $f$  in  $\mathbf{a}$ . Dann gilt:

- Ist  $H_f(\mathbf{a})$  positiv definit, so ist  $\mathbf{a}$  ein lokales Minimum.
- Ist  $H_f(\mathbf{a})$  negativ definit, so ist  $\mathbf{a}$  ein lokales Maximum.

c) Ist  $H_f(\mathbf{a})$  indefinit, so ist  $\mathbf{a}$  ein Sattelpunkt.

Um festzustellen, ob  $H_f(\mathbf{a})$  positiv oder negativ definit ist, können wir die Sätze 7.7.8 und 7.7.9 benutzen.

**8.3.12 Folgerung.** Sei  $f : U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  ein  $C^2$ -Funktion und  $\mathbf{a}$  ein stationärer Punkt. Dann gilt:

a)  $f_{xx}(\mathbf{a}) > 0$ , und  $\det H_f(\mathbf{a}) > 0$ , so ist  $\mathbf{a}$  lokales Minimum,

b)  $f_{xx}(\mathbf{a}) < 0$ , und  $\det H_f(\mathbf{a}) > 0$ , so ist  $\mathbf{a}$  lokales Maximum,

c)  $\det H_f(\mathbf{a}) < 0$ , so ist  $\mathbf{a}$  Sattelpunkt.

**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:

$$\begin{aligned} f_x &= 6y - 6x^2 \\ f_y &= 6x - 6y \\ \text{grad } f = 0 &\Leftrightarrow x = y \text{ und } x^2 = y \end{aligned}$$

Mögliche Kandidaten sind also  $(0, 0)$  und  $(1, 1)$ . Zur genauen Bestimmung berechnen wir die Hessematrix

$$\begin{aligned} H_f(x, y) &= \begin{pmatrix} -12x & 6 \\ 6 & -6 \end{pmatrix} \\ H_f(0, 0) &= \begin{pmatrix} 0 & 6 \\ 6 & -6 \end{pmatrix} && \text{indefinit} \\ H_f(1, 1) &= \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix} && \text{negativ definit} \end{aligned}$$

Damit liegt bei  $(0, 0)$  ein Sattelpunkt und bei  $(1, 1)$  ein Maximum vor.

- In Extremwertaufgaben  $f(\mathbf{x}) = \text{Extr}!$  ist oft die Menge der zulässigen Punkte  $\mathbf{x}$  durch Nebenbedingungen der Form

$$g_1(\mathbf{x}) = 0, \dots, g_k(\mathbf{x}) = 0$$

eingeschränkt. Eine solche Aufgabe bezeichnen wir mit

$$f(\mathbf{x}) = \text{Extr}! \quad \text{NB } g_i(\mathbf{x}) = 0 \quad i = 1, \dots, k.$$

**8.3.13 Definition.** Sei  $M = \{\mathbf{x} \in \mathbb{R}^n; g(\mathbf{x}) = 0\}$ . Man sagt,  $\mathbf{a} \in M$  ist ein (lokales) **Maximum (Minimum) von  $f$  unter der Nebenbedingung  $g(\mathbf{x}) = 0$** , wenn es eine Umgebung  $B_r(\mathbf{a})$  gibt so, dass  $f(\mathbf{x}) \leq f(\mathbf{a})$ , ( $f(\mathbf{a}) \leq f(\mathbf{x})$ ) gilt für alle  $\mathbf{x} \in B_r(\mathbf{a}) \cap M$ .

**8.3.14 Satz (Lagrange Multiplikator).** Zu jeder Lösung  $\mathbf{a}$  des Extremalproblems mit Nebenbedingung

$$f(\mathbf{x}) = \text{Extr}! \quad \text{NB } g(\mathbf{x}) = 0$$

mit  $C^1$ -Funktionen  $f, g$  und  $\text{grad } g(\mathbf{a}) \neq 0$  gibt es eine Zahl  $\lambda_0 \in \mathbb{R}$ , den **Lagrange Multiplikator**, so dass gilt:

$$\text{grad } f(\mathbf{a}) + \lambda_0 \text{grad } g(\mathbf{a}) = 0.$$

- Satz 8.3.13 besagt, dass in Extremalstellen die Gradienten von  $f$  und  $g$  parallel sind.

**8.3.15 Lagrange Multiplikatorregel.** Satz 8.3.14 liefert folgendes Verfahren zur Bestimmung von Extremalstellen mit Nebenbedingungen:

1. Bilde die Hilfsfunktion

$$L(x_1, \dots, x_n, \lambda) := f(x_1, \dots, x_n) + \lambda g(x_1, \dots, x_n)$$

und berechne  $\text{grad } L$ .

2. Bestimme die Lösungen  $(\mathbf{a}, \lambda_0)$  von  $\text{grad } L(\mathbf{x}, \lambda) = 0$ , d.h. löse das System

$$\begin{aligned} L_{x_i}(\mathbf{x}, \lambda) &= f_{x_i}(\mathbf{x}) + \lambda g_{x_i}(\mathbf{x}) = 0, & i &= 1, \dots, n, \\ L_{\lambda}(\mathbf{x}, \lambda) &= g(\mathbf{x}) = 0. \end{aligned}$$

3. Man untersucht welche der gefundenen Werte  $\mathbf{a}$  tatsächlich Extremalstellen sind. Dies ist oft sehr kompliziert. Am Ende sind noch die Punkte  $\mathbf{b}$ , in denen  $f, g$  nicht differenzierbar sind bzw.  $\text{grad } g(\mathbf{b}) = 0$  ist, zu betrachten.

**Beispiel:** Sei  $f(x, y) = \frac{1}{p}x^p + \frac{1}{q}y^q$ , mit  $\frac{1}{p} + \frac{1}{q} = 1$ , und  $g(x, y) = 1 - xy$  die Nebenbedingung. Dann bilden wir die Lagrange Funktion  $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$  und berechnen den Gradienten:

$$\begin{aligned} L_x &= x^{p-1} - \lambda y, \\ L_y &= y^{q-1} - \lambda x, \\ L_\lambda &= 1 - xy. \end{aligned}$$

Die stationären Punkte, d.h.  $\text{grad } L = 0$ , berechnen sich wie folgt: Die ersten beiden Gleichungen liefern:

$$x^{p-1} = \lambda y \quad y^{q-1} = \lambda x \quad \Rightarrow \quad \frac{x^{p-1}}{y} = \frac{y^{q-1}}{x} \quad \Rightarrow \quad x = y^{\frac{q}{p}}.$$

Aus der 3. Gleichung erhalten wir

$$y^{\frac{p+q}{p}} = 1 \quad \Rightarrow \quad y = 1 \quad \Rightarrow \quad x = 1 \quad \Rightarrow \quad \lambda = 1.$$

Somit ist der Punkt  $(1, 1)$  ein Kandidat für ein Extremum.

Die Nebenbedingung  $g(x, y) = 0$  ist äquivalent zu  $y = \frac{1}{x}$ . Also hat die Funktion  $h(x) = f(x, \frac{1}{x}) = \frac{1}{p}x^p + \frac{1}{q}x^{-q}$  in  $x = 1$  einen stationären Punkt. Wir haben:

$$\begin{aligned} h'(x) &= x^{p-1} - x^{-q-1}, \\ h''(x) &= (p-1)x^{p-2} + (q+1)x^{-q-2}, \\ h''(1) &= p-1 + q+1 = p+q > 0. \end{aligned}$$

Also ist  $x = 1$  ein globales Minimum, d.h.  $f(x, y) \geq 1$  für alle  $x, y > 0$  mit  $xy = 1$ . Die spezielle Wahl  $x = \frac{u}{(uv)^{1/p}}$  und  $y = \frac{v}{(uv)^{1/q}}$  erfüllt die Bedingung  $xy = 1$  und somit erhalten wir

$$\frac{1}{p} \frac{u^p}{uv} + \frac{1}{q} \frac{v^q}{uv} \geq 1 \quad \Leftrightarrow \quad uv \leq \frac{1}{p}u^p + \frac{1}{q}v^q.$$

Dies ist die sogenannte **Youngsche Ungleichung**. Es gibt wesentlich direktere Beweise dieser Ungleichung.

- Im Falle mehrerer Nebenbedingungen geht man analog vor. Seien  $f, g_i, i = 1, \dots, k$   $C^1$ -Funktionen und seien für alle  $\mathbf{x} \in \mathbf{M}$

grad  $g_i(\mathbf{x}), i = 1, \dots, k$  linear unabhängig, wobei  $\mathbf{M} = \{\mathbf{x} \in \mathbb{R}^n; g_i(\mathbf{x}) = 0, i = 1, \dots, k\}$  ist. Dann findet man die Lösung des Extremalproblems mit Nebenbedingungen

$$f(\mathbf{x}) = \text{Extr} ! \quad \text{NB } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, k$$

unter den stationären Punkten der Lagrange Funktion

$$L(\mathbf{x}, \lambda_1, \dots, \lambda_k) := f(\mathbf{x}) + \sum_{i=1}^k \lambda_i g_i(\mathbf{x}).$$

**8.3.16** Extremwertbestimmung. Sei  $f$  auf der Menge

$$U := \{\mathbf{x} \in \mathbb{R}^n; g_i(\mathbf{x}) \leq 0, i = 1, \dots, r\}$$

definiert. Kandidaten für Extremwerte sind:

a) Die stationären Punkte im Innern von  $U$ . Diese Menge ist gegeben durch:

$$\{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) < 0, i = 1, \dots, r \text{ und } \text{grad } f(\mathbf{x}) = \mathbf{0}\}.$$

b) Die Extremalstellen auf dem Rand von  $U$ , d.h.  $g_i(\mathbf{x}) = 0$  für ein  $i$ , die man mit 8.3.15 berechnet. Achtung: Selbst wenn  $\mathbf{x}$  lokales Extremum von  $f|_{\partial U}$  ist, muss man noch prüfen, dass  $\mathbf{x}$  auch lokales Extremum von  $f|_U$  ist.

c) Die Ecken von  $U$  (falls vorhanden sind sie die eindeutigen Lösungen von Kombinationen  $g_i(\mathbf{x}) = \dots = g_k(\mathbf{x}) = 0$  sonst  $g_l(\mathbf{x}) < 0$ ). Gegebenenfalls gehen wir wie unter 8.3.15 mit mehreren Nebenbedingungen vor.

d) Die Punkte in  $U$  in denen  $f$  nicht differenzierbar ist, oder  $g_i = 0$  für ein  $i$  gilt, aber  $g_i$  nicht differenzierbar ist oder aber  $\text{grad } g_i = \mathbf{0}$  gilt.

**Beispiel:** Sei  $f(x, y) = 3x^2 - 2xy + y^2$  und  $g(x, y) = x^2 + y^2 - 1$ . Die Punkte c) und d) in 8.3.16 entfallen. Die Menge  $U = B_1(\mathbf{0})$  ist die abgeschlossene Kreisscheibe mit Radius 1.

- Im Innern von  $U$  d.h. Punkte  $(x, y)$  für die gilt:  $x^2 + y^2 < 1$ . Für die partiellen Ableitungen gilt:

$$f_x = 6x - 2y, \quad f_y = -2x + 2y.$$

Es gibt nur einen stationären Punkt, nämlich  $(x, y) = (0, 0)$ . Für die Hesse Matrix gilt:

$$H_f(0, 0) = \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix},$$

Und somit ist  $6 > 0$ ,  $\det H = 12 - 4 > 0$ . Also ist  $H_f(0, 0)$  positiv definit und  $(0, 0)$  ist ein lokales Minimum.

- Auf dem Rand, d.h. Punkte  $(x, y)$  für die  $x^2 + y^2 = 1$  gilt:  
Die Lagrange Funktion ist

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

Deren Ableitungen berechnen sich als

$$\begin{aligned} L_x &= 6x - 2y + 2\lambda x, \\ L_y &= 2y - 2x + 2\lambda y, \\ L_\lambda &= x^2 + y^2 - 1. \end{aligned}$$

Das System der ersten zwei Gleichungen kann man als  $\mathbf{A}(x, y)^T = -\lambda(x, y)^T$  schreiben mit der Matrix

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}.$$

Dies ist ein Eigenwertproblem und wir erhalten:

$$\begin{aligned} \chi_{\mathbf{A}}(\lambda) &= \det \begin{pmatrix} \lambda + 3 & -1 \\ -1 & \lambda + 1 \end{pmatrix} = \lambda^2 + 4\lambda + 2, \\ \lambda_{1,2} &= -2 \pm \sqrt{2}. \end{aligned}$$

Ein Eigenvektor zu  $\lambda_1 = -2 - \sqrt{2}$  ist  $(1, 1 - \sqrt{2})^T$ . Die Bedingung  $x^2 + y^2 = 1$  bedeutet, dass wir einen Einheitsvektor suchen. Dieser ist  $\pm \mathbf{v}_1$ , mit

$$\mathbf{v}_1 = \frac{1}{\sqrt{4 - 2\sqrt{2}}} \begin{pmatrix} 1 \\ 1 - \sqrt{2} \end{pmatrix}$$

Analog erhalten wir  $\pm \mathbf{v}_2$  für  $\lambda_2 = -2 + \sqrt{2}$  mit

$$\mathbf{v}_2 = \frac{1}{\sqrt{4 + 2\sqrt{2}}} \begin{pmatrix} 1 \\ 1 + \sqrt{2} \end{pmatrix}$$

- Kandidaten für Extremstellen sind also  $\mathbf{0}$  und  $\pm \mathbf{v}_1, \pm \mathbf{v}_2$ . Einsetzen ergibt:

$$\begin{aligned} f(\mathbf{v}_1) &= f(-\mathbf{v}_1) = 2 + \sqrt{2} && \text{Maximum,} \\ f(\mathbf{v}_2) &= f(-\mathbf{v}_2) = 2 - \sqrt{2}, \\ f(\mathbf{0}) &= 0 && \text{Minimum.} \end{aligned}$$

Beachte insbesondere: Der Wert  $f(\pm \mathbf{v}_2)$  ist das globale Minimum, wenn man  $f$  auf den Rand  $S^1$  einschränkt, aber im Inneren gibt es noch kleinere Werte.

## 8.4 Vektorwertige Funktionen

Für Funktionen  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  werden die Begriffe aus Paragraph 8.2 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.

**8.4.1 Definition.** Sei  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $\mathbf{x}, \mathbf{x}_0 \in D$ .

- a) Der **Grenzwert**, die **partiellen Ableitungen** und die „**klein-o**“-**Notation** sind jeweils komponentenweise erklärt:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) := \begin{pmatrix} \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_1(\mathbf{x}) \\ \vdots \\ \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_m(\mathbf{x}) \end{pmatrix}; \quad \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(\mathbf{x}) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(\mathbf{x}) \end{pmatrix}$$

$$\mathbf{f}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|) \quad \Leftrightarrow \quad f_j(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|), j = 1, \dots, m.$$

- 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  $\mathbf{A} \in \mathbb{R}^{m \times n}$  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:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{A}(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|). \quad (4.1)$$

- Aus (4.1) 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  $\mathbf{A}$  in (4.1) eindeutig bestimmt als

$$\begin{aligned} \mathbf{A} = J_{\mathbf{f}}(\mathbf{x}_0) &:= \begin{pmatrix} \text{grad } f_1(\mathbf{x}_0)^T \\ \vdots \\ \text{grad } f_m(\mathbf{x}_0)^T \end{pmatrix} \\ &= \left( \frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{x}_0) \right) = \left( \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right)_{i,j}. \end{aligned}$$

$J_{\mathbf{f}}(\mathbf{x}_0) \in \mathbb{R}^{m \times n}$  heißt **Jacobi-Matrix** von  $\mathbf{f}$  an der Stelle  $\mathbf{x}_0$ .

- Umgekehrt ist  $\mathbf{f}$  total differenzierbar, wenn die Komponentenfunktionen  $f_1, \dots, f_m$  differenzierbar sind.

**8.4.2 Satz.** Jede  $C^1$ -Funktion  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $D$  offen, ist auf  $D$  total differenzierbar.

**8.4.3 Newton - Verfahren.** Es sei  $D \subset \mathbb{R}^n$  und  $\mathbf{f} : D \rightarrow \mathbb{R}^n$ . Zur näherungsweise Lösung des nichtlinearen Gleichungssystems

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \quad (*)$$

für  $f, g \in C^1$  ersetzt man (\*) durch die lineare Approximation

$$\mathbf{f}(\mathbf{x}_0) + J_{\mathbf{f}}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = \mathbf{0}$$

aus Definition 8.4.1 c). Falls  $J_{\mathbf{f}}(\mathbf{x}_0)$  invertierbar ist, erhält man eine verbesserte Lösung zum Startwert  $(f(x_0, y_0), g(x_0, y_0))^T$  durch

$$\mathbf{x}_1 = \mathbf{x}_0 - J_{\mathbf{f}}(\mathbf{x}_0)^{-1} \cdot \mathbf{f}(\mathbf{x}_0).$$

In der Praxis löst man dazu das Gleichungssystem

$$J_{\mathbf{f}}(\mathbf{x}_0) \cdot \mathbf{v} = -\mathbf{f}(\mathbf{x}_0)$$

und setzt  $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{v}$ . Dieses Vorgehen kann wiederholt werden und liefert oft gute Ergebnisse bei entsprechender Wahl des Startpunktes.

In Analogie von Satz 3.3.18 kann man zeigen, dass das Newton-Verfahren für eine  $C^2$ -Funktion  $\mathbf{f}$  mit Nullstelle  $\mathbf{a}$  mit Startwerten in einer Umgebung von  $\mathbf{a}$  gegen  $\mathbf{a}$  konvergiert, wenn  $J_{\mathbf{f}}(\mathbf{a})$  invertierbar ist.

**8.4.4 Methode der kleinsten Quadrate.** Aus experimentellen Daten sind die Werte gewisser Konstanten  $x_1, \dots, x_n$  zu bestimmen. Da Messungen mit Fehlern behaftet sind, werden mehr Messdaten  $y_1, \dots, y_m$  ermittelt als nötig wären um die  $x_i$  zu bestimmen. Oft sind die experimentellen Daten  $y_1, \dots, y_m$  nur indirekt mit den Konstanten  $x_i$  verbunden. Also ergibt sich ein überbestimmtes nichtlineares Gleichungssystem

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) \quad (4.2)$$

für eine Funktion  $\mathbf{f}: D \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \in D \subset \mathbb{R}^n$ . Um die Messfehler zu berücksichtigen, wird stattdessen eine Lösung  $\mathbf{x}$  gesucht so, dass  $y_i - f_i(\mathbf{x})$  minimal ist. Dazu wird die **Methode der kleinsten Quadrate** benutzt, d.h. finde  $\mathbf{x} \in \mathbb{R}^n$  mit

$$F(\mathbf{x}) := \|\mathbf{y} - \mathbf{f}(\mathbf{x})\|^2 = \sum_{i=1}^m (y_i - f_i(\mathbf{x}))^2 = \text{Min} ! \quad (4.3)$$

Im nächsten Satz nehmen wir an, dass  $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$  linear ist mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$ .

**8.4.5 Satz.** Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $m > n$ .

a) Das Minimalproblem

$$F(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 = \text{Min} ! \quad (4.4)$$

ist immer lösbar.

b) Lösungen von (4.4) und von

$$\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T \mathbf{y} \quad (4.5)$$

stimmen immer überein.

c) Für zwei Lösungen  $\mathbf{x}, \mathbf{x}_0$  von (4.5) gilt stets  $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}_0$ . Ist  $\text{Rang } \mathbf{A} = n$ , so ist die Lösung von (4.5) eindeutig bestimmt.

- Formel 4.5 liefert eine praktische Methode zur Lösung von

$$\|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 = \text{Min} !$$

Man berechnet  $\mathbf{B} := \mathbf{A}^T \mathbf{A}$  und  $\mathbf{z} := \mathbf{A}^T \mathbf{y}$  und löst mittels Gaußelimination das System  $\mathbf{B}\mathbf{x} = \mathbf{z}$ , welches nach Satz 8.4.5 immer lösbar ist.

- Falls die Funktionen  $f_i$  in (4.3) nichtlinear sind kann (4.3) oft nur approximativ gelöst werden. Sukzessive Näherungslösungen werden mit Hilfe einer Variante des Gradientenverfahrens 8.3.2 berechnet. Sei  $\mathbf{x}_0$  ein Startwert. Dann wird anstelle von (4.3) das lineare Problem

$$\|\mathbf{y} - f(\mathbf{x}_0) - J_{\mathbf{f}}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|^2 = \text{Min} ! \quad (*)$$

gelöst.

Sei  $\mathbf{x}_1$  eine Lösung von (\*). Dann wird wie beim Gradientenverfahren durch  $\mathbf{x}_0 + h(\mathbf{x}_1 - \mathbf{x}_0)$  eine verbesserte Lösung konstruiert.

**8.4.6 Satz** (Taylor-Formel). Ist  $D \subseteq \mathbb{R}^n$  ein konvexes Gebiet,  $\mathbf{f} \in C^{k+1}(D, \mathbb{R}^m)$ , und  $\mathbf{x}, \mathbf{x} + \mathbf{v} \in D$ , dann gilt:

$$f_i(\mathbf{x} + \mathbf{v}) = f_i(\mathbf{x}) + \sum_{j=1}^k \frac{1}{j!} \partial_{\mathbf{v}}^j f_i(\mathbf{x}) + R_{i,k}(\mathbf{x}, \mathbf{v}), \quad i = 1, \dots, m$$

mit dem Restglied

$$R_{i,k}(\mathbf{x}, \mathbf{v}) = \frac{1}{(k+1)!} \partial_{\mathbf{v}}^{k+1} f_i(\mathbf{x} + \xi_{i,k} \mathbf{v})$$

mit  $\xi_{i,k} \in (0, 1)$  welches von  $i$  abhängt.

- Achtung! Die Punkte im Restglied zwischen  $\mathbf{x}, \mathbf{x} + \mathbf{v}$  hängen von der Komponentenfunktion  $f_i$  ab!

**8.4.7 Folgerung.** Für eine  $C^1$ -Funktion  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  gilt:

$$f_i(\mathbf{x}) = f_i(\mathbf{x}_0) + \text{grad } f_i(\mathbf{x}_i^*) \cdot (\mathbf{x} - \mathbf{x}_0), \quad i = 1, \dots, m,$$

mit  $\mathbf{x}_i^*$  zwischen  $\mathbf{x}$  und  $\mathbf{x}_0$ . Gilt insbesondere  $\|\text{grad } f_i(\mathbf{x})\| \leq M$  für alle  $\mathbf{x} \in D$  und  $i = 1, \dots, m$ , so gilt:

$$\|f(\mathbf{x}) - f(\mathbf{x}_0)\| \leq M \|\mathbf{x} - \mathbf{x}_0\| \quad \forall \mathbf{x}, \mathbf{x}_0 \in D.$$

**Beispiele:**

- Sei  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear. Dann existiert ein  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mit  $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ . Also gilt  $J_{\mathbf{f}}(\mathbf{x}) = \mathbf{A}$ .
- Polarkoordinaten: Die Menge  $D = \{(r, \varphi); 0 \leq r < \infty, 0 \leq \varphi < 2\pi\}$  wird mittels

$$\mathbf{x}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

auf die  $\mathbf{x}$ -Ebene abgebildet, d.h.  $\mathbf{x} : D \rightarrow \mathbb{R}^2$ . Die Jacobi Matrix ist

$$J_{\mathbf{x}}(r, \varphi) = \left( \frac{\partial \mathbf{x}}{\partial r}, \frac{\partial \mathbf{x}}{\partial \varphi} \right) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

**8.4.8 Lemma.** Für  $D \subset \mathbb{R}^n$ ,  $\mathbf{v}, \mathbf{w} : D \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \in D$ ,  $f : D \rightarrow \mathbb{R}$  und  $\alpha, \beta \in \mathbb{R}$  gelten:

- $J_{\alpha\mathbf{v} + \beta\mathbf{w}}(\mathbf{x}) = \alpha J_{\mathbf{v}}(\mathbf{x}) + \beta J_{\mathbf{w}}(\mathbf{x})$ ,
- $J_{f\mathbf{v}}(\mathbf{x}) = f(\mathbf{x})J_{\mathbf{v}}(\mathbf{x}) + \mathbf{v}(\mathbf{x})(\text{grad } f(\mathbf{x}))^T$ ,
- $J_{\mathbf{v}^T\mathbf{w}}(\mathbf{x}) = \mathbf{v}^T J_{\mathbf{w}}(\mathbf{x}) + \mathbf{w}^T J_{\mathbf{v}}(\mathbf{x})$ ,
- $J_{\mathbf{v} \times \mathbf{w}} = \mathbf{v}(\mathbf{x}) \times J_{\mathbf{w}}(\mathbf{x}) - \mathbf{w}(\mathbf{x}) \times J_{\mathbf{v}}(\mathbf{x})$  falls  $m = 3$ .

• In obiger Formel ist zu beachten:

- $\mathbf{v}(\text{grad } f)^T$  ist eine Matrix
- Vektor  $\times$  Matrix  $\hat{=}$  Vektor  $\times$  Spaltenvektoren.

**8.4.9 Satz** (Kettenregel). Sind  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  in  $\mathbf{x}_0 \in D$   $\mathbf{g} : G \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^q$ , 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:

$$J_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}_0) = J_{\mathbf{g}}(\mathbf{f}(\mathbf{x}_0)) J_{\mathbf{f}}(\mathbf{x}_0) .$$

**8.4.10 Basiswechsel.** Ein (orientiertes) kartesisches Koordinatensystem im  $\mathbb{R}^n$  hat die Gestalt  $K = (\mathbf{P}, \mathbf{b}_1, \dots, \mathbf{b}_n)$ . Dabei ist  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  eine (orientierte) Orthonormalbasis des  $\mathbb{R}^n$  (das heißt,  $\mathbf{B}^{-1} = \mathbf{B}^T$  und ggf.  $\det \mathbf{B} = 1$ ) und  $\mathbf{p} = \overrightarrow{\mathbf{0P}}$  der Fußpunkt. Ein Vektor  $\mathbf{x} \in \mathbb{R}^n$  hat im Koordinatensystem  $K$  den Koordinatenvektor  $\mathbf{y}$ , der durch

$$\mathbf{x} = \mathbf{B}\mathbf{y} + \mathbf{p} \quad (k_{\mathbf{S}}(\mathbf{x}) = \mathbf{B} k_K(\mathbf{x}) + \mathbf{p})$$

bestimmt ist. Umgekehrt ist  $\mathbf{y} = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{p}) = \mathbf{B}^T(\mathbf{x} - \mathbf{p})$ .

Sei  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_m)$  eine Basis des  $\mathbb{R}^m$ . Ein Vektor  $\mathbf{v} \in \mathbb{R}^m$  hat bezüglich der Basis  $\mathbf{W}$  den Koordinatenvektor  $\mathbf{w}$

$$\mathbf{v} = \mathbf{W}\mathbf{w} . \quad (k_{\mathbf{S}}(\mathbf{v}) = \mathbf{W} k_{\mathbf{W}}(\mathbf{v}))$$

Eine Funktion  $\mathbf{f} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  besitzt nach diesem Koordinaten- und Basiswechsel eine Darstellung  $\mathbf{g} : D^* \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ , die gegeben ist durch

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= k_{\mathbf{S}, \mathbb{R}^m}(\mathbf{f}(k_{\mathbf{S}, \mathbb{R}^n}(\mathbf{x}))) \\ &= \mathbf{W} k_{\mathbf{W}}(\mathbf{f}(\mathbf{B} k_K(\mathbf{x}) + k_{\mathbf{S}, \mathbb{R}^n}(\mathbf{p}))) \\ &=: \mathbf{W} \mathbf{g}(k_K(\mathbf{x})) . \end{aligned}$$

Mit der Bezeichnung  $\mathbf{y} = k_K(\mathbf{x})$   $\mathbf{g}(\mathbf{y}) = k_{\mathbf{W}}(\mathbf{f}(\mathbf{x}))$  erhalten wir

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{B}\mathbf{y} + \mathbf{p}) = \mathbf{W} \mathbf{g}(\mathbf{y}) ,$$

d.h.

$$\mathbf{g}(\mathbf{y}) = \mathbf{W}^{-1} \mathbf{f}(\mathbf{B}\mathbf{y} + \mathbf{p})$$

und für die Jacobi-Matrix gilt:

$$J_{\mathbf{g}}(\mathbf{y}_0) = \mathbf{W}^{-1} J_{\mathbf{f}}(\mathbf{x}_0) \mathbf{B}, \quad \text{mit } \mathbf{x}_0 = \mathbf{B}\mathbf{y}_0 + \mathbf{p} .$$

Oftmals sind die Basen  $\mathbf{B}$  und  $\mathbf{W}$  identisch.

**8.4.11** Skalaren- und Vektorfelder. Ein **Skalarenfeld**  $f : D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$  ordnet jedem Punkt  $\mathbf{x} \in D$  eine Zahl  $f(\mathbf{x})$  zu. Ein **Vektorfeld**  $\mathbf{v} : D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ordnet jedem Punkt  $\mathbf{x} \in D$  einen Vektor  $\mathbf{v}(\mathbf{x}) \in \mathbb{R}^3$  zu. Eine Kurve in  $D$  heißt **Feldlinie** des Vektorfeldes  $\mathbf{v}$ , wenn der Vektor  $\mathbf{v}(\mathbf{x})$  in jedem Kurvenpunkt  $\mathbf{x}$  parallel zur Kurventangente ist.

**8.4.12** Die starre Drehung. Eine Rechtsdrehung mit konstanter Winkelgeschwindigkeit  $\omega$  um eine Achse durch den Nullpunkt in Richtung  $\mathbf{a} = (a_1, a_2, a_3)^T$ ,  $\|\mathbf{a}\| = 1$  ist gegeben durch (Kapitel 7 (7.6.7))

$$\mathbf{x}(t) = \cos(\omega t)\mathbf{x}_0 + (1 - \cos \omega t)(\mathbf{x}_0 \cdot \mathbf{a})\mathbf{a} + \sin(\omega t)\mathbf{a} \times \mathbf{x}_0,$$

wobei  $\mathbf{x}_0 = \mathbf{x}(0)$  die Lage zum Zeitpunkt Null bezeichnet. Für die Geschwindigkeit  $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$  gilt also

$$\begin{aligned} \mathbf{v}(t) &= \omega(-\sin(\omega t)\mathbf{x}_0 + \sin(\omega t)(\mathbf{x}_0 \cdot \mathbf{a})\mathbf{a} + \cos(\omega t)\mathbf{a} \times \mathbf{x}_0) \\ &= \omega \mathbf{a} \times \mathbf{x}(t), \end{aligned}$$

dabei haben wir  $\mathbf{a} \times (\mathbf{a} \times \mathbf{x}) = (\mathbf{x} \cdot \mathbf{a})\mathbf{a} - \mathbf{x}$  benutzt. Man nennt  $\mathbf{w} := \omega \mathbf{a}$  die **vektorielle Winkelgeschwindigkeit** und somit hat das Geschwindigkeitsfeld einer gleichförmigen Drehbewegung die Darstellung

$$\mathbf{x} \rightarrow \mathbf{v}(\mathbf{x}) = \mathbf{w} \times \mathbf{x} = \omega \mathbf{a} \times \mathbf{x}.$$

Dies ist eine lineare Abbildung mit der Abbildungsmatrix  $\mathbf{M}$

$$J_{\mathbf{v}}(\mathbf{x}) = \mathbf{M} = (\mathbf{v}(\mathbf{e}_1), \mathbf{v}(\mathbf{e}_2), \mathbf{v}(\mathbf{e}_3)) = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix},$$

die schiefsymmetrisch ist, d.h.  $\mathbf{M}^T = -\mathbf{M}$ , und deren Spur (Kapitel 7 (5.8))  $\text{Sp } J_{\mathbf{v}}(\mathbf{x})$  Null ist.

**8.4.13** Das Zentralkraftfeld. Eine Punktmasse  $M$  im Ursprung zieht die Punktmasse  $m$  in  $\mathbf{x} = (x_1, x_2, x_3)^T$  mit der Gravitationskraft

$$\mathbf{K}(\mathbf{x}) = c \frac{\mathbf{x}}{\|\mathbf{x}\|^3} \quad \mathbf{x} \neq \mathbf{0} \quad (*)$$

an, mit  $c = -\gamma m M$ ,  $\gamma > 0$ . Falls es sich um Punktladungen der Stärke  $Q$  bzw.  $q$  handelt gilt (\*) mit  $c = kqQ$ ,  $k > 0$ . Das zentrale Kraftfeld  $\mathbf{K}(\mathbf{x})$

besitzt die Jacobi Matrix für  $\mathbf{x} \neq \mathbf{0}$

$$\begin{aligned} J_{\mathbf{K}}(\mathbf{x}) &= \mathbf{x} \left( \text{grad} \frac{c}{\|\mathbf{x}\|^3} \right)^T + \frac{c}{\|\mathbf{x}\|^3} \mathbf{E}_3 \\ &= \frac{c}{\|\mathbf{x}\|^5} (-3\mathbf{x}\mathbf{x}^T + \mathbf{x} \cdot \mathbf{x} \mathbf{E}_3) \\ &= \frac{c}{\|\mathbf{x}\|^5} \begin{pmatrix} x_2^2 + x_3^2 - 2x_1^2 & -3x_1x_2 & -3x_1x_3 \\ -3x_1x_2 & x_1^2 + x_3^2 - 2x_2^2 & -3x_2x_3 \\ -3x_1x_3 & -3x_2x_3 & x_1^2 + x_2^2 - 2x_3^2 \end{pmatrix}, \end{aligned}$$

welche symmetrisch ist und deren Spur Null ist.

**8.4.14** Der Wirbel. Ein gerader stromdurchflossener Draht in Richtung der  $x_3$ -Achse besitzt das magnetische Feld

$$\begin{aligned} \mathbf{H}(\mathbf{x}) &= \frac{c}{\|\mathbf{x}\|^2 - (\mathbf{e}_3 \cdot \mathbf{x})^2} \mathbf{e}_3 \times \mathbf{x} \\ &= \frac{c}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \end{aligned}$$

$(x_1, x_2) \neq (0, 0), c > 0$ . Die Jacobi Matrix ist

$$J_{\mathbf{H}}(\mathbf{x}) = \frac{c}{(x_1^2 + x_2^2)^2} \begin{pmatrix} 2x_1x_2 & x_2^2 - x_1^2 & 0 \\ x_2^2 - x_1^2 & -2x_1x_2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

welche symmetrisch ist und deren Spur Null ist.

**8.4.15** Die laminare Rohrströmung. Das Geschwindigkeitsprofil einer zähen Flüssigkeit, die durch ein zur  $x_2$ -Achse koaxiales Rohr vom Radius  $r$  strömt, ist

$$\mathbf{v}(\mathbf{x}) = c(0, r^2 - x_1^2 - x_3^2, 0)^T, \quad x_1^2 + x_3^2 \leq r, c > 0.$$

Die Jacobi Matrix ist

$$J_{\mathbf{v}}(\mathbf{x}) = c \begin{pmatrix} 0 & 0 & 0 \\ -2x_1 & 0 & -2x_3 \\ 0 & 0 & 0 \end{pmatrix},$$

welche weder symmetrisch noch schiefsymmetrisch ist aber Spur Null hat.

**8.4.16 Definition.** 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**  $\Delta$  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:

$$\text{div } \mathbf{v}(\mathbf{x}) = \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i}(\mathbf{x}),$$

$$\text{rot } \mathbf{v}(\mathbf{x}) = \begin{pmatrix} \frac{\partial v_3}{\partial x_2}(\mathbf{x}) - \frac{\partial v_2}{\partial x_3}(\mathbf{x}) \\ \frac{\partial v_1}{\partial x_3}(\mathbf{x}) - \frac{\partial v_3}{\partial x_1}(\mathbf{x}) \\ \frac{\partial v_2}{\partial x_1}(\mathbf{x}) - \frac{\partial v_1}{\partial x_2}(\mathbf{x}) \end{pmatrix}.$$

- Mit diesen Begriffen gelten für die Beispiele 8.4.12 - 8.4.15:
  - Die Divergenz ist Null (Divergenz = Spur  $J$ ).
  - Die Rotation des zentralen Kraftfeldes und des Wirbels sind Null. Die Rotation der gleichförmigen Drehung ist  $2\omega \mathbf{a}$  und der laminaren Strömung ist  $2c(x_3, 0, -x_1)^T$ .
- Wir wollen zeigen, dass die Werte von  $\Delta f$ ,  $\text{div } f$  und die durch  $\text{rot } \mathbf{v}$ ,  $\text{grad } \mathbf{v}$  dargestellten Vektoren in einem festen Punkt nicht von der Wahl des Koordinatensystems abhängen.

**8.4.17 Lemma.** *Es gilt:*

$$\text{grad } f(\mathbf{x}_0) = (J_f(\mathbf{x}_0))^T,$$

$$\text{div } \mathbf{v}(\mathbf{x}_0) = \text{Sp } J_{\mathbf{v}}(\mathbf{x}_0),$$

$$\text{rot } \mathbf{v}(\mathbf{x}_0) \times (\mathbf{x} - \mathbf{x}_0) = (J_{\mathbf{v}}(\mathbf{x}_0) - J_{\mathbf{v}}(\mathbf{x}_0)^T) (\mathbf{x} - \mathbf{x}_0).$$

- Nach Lemma 8.4.17 ist  $\mathbf{x} \mapsto \mathbf{S}(\mathbf{x}) = \mathbf{w} \times \mathbf{x}$  das Geschwindigkeitsfeld einer Drehung, daher der Name Rotation.

**8.4.18 Satz.** *Sind  $I = (\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  und  $K = (\mathbf{P}, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$  zwei kartesische Koordinatensysteme und ist  $\mathbf{p} = \overrightarrow{\text{OP}}$ , und  $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$  die Darstellung von  $K$  bezüglich  $I$ , so gilt:*

System	$I$	$K$
Basis	$\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$	$\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3), \mathbf{B}^T = \mathbf{B}^{-1}$
Punkt $\mathbf{x}$	$\mathbf{x} = (x_1, x_2, x_3)^T$	$\mathbf{y} = (y_1, y_2, y_3)^T = \mathbf{B}^T(\mathbf{x} - \mathbf{p})$
Vektor $\overrightarrow{\mathbf{xy}}$	$\mathbf{v} = \mathbf{y} - \mathbf{x}$	$\mathbf{w} = \mathbf{B}^T \mathbf{v}$
Skalarenfeld $f : D \rightarrow \mathbb{R}$	$f(\mathbf{x})$ $J_f(\mathbf{x})$	$g(\mathbf{y}) = f(\mathbf{B}\mathbf{y} + \mathbf{p})$ $J_g(\mathbf{y}) = J_f(\mathbf{B}\mathbf{y} + \mathbf{p})\mathbf{B}$
Gradient	$\text{grad } f(\mathbf{x}) = J_f(\mathbf{x})^T$	$\text{grad } g(\mathbf{y}) = \mathbf{B}^T \text{grad } f(\mathbf{x})$
Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^3$	$\mathbf{v}(\mathbf{x})$ $J_{\mathbf{v}}(\mathbf{x})$	$\mathbf{w}(\mathbf{y}) = \mathbf{B}^T \mathbf{v}(\mathbf{B}\mathbf{y} + \mathbf{p})$ $J_{\mathbf{w}}(\mathbf{y}) = \mathbf{B}^T J_{\mathbf{v}}(\mathbf{B}\mathbf{y} + \mathbf{p})\mathbf{B}$
Divergenz	$\text{div } \mathbf{v}(\mathbf{x}) = \text{Sp } J_{\mathbf{v}}(\mathbf{x})$	$\text{div } \mathbf{w}(\mathbf{y}) = \text{div } \mathbf{v}(\mathbf{B}\mathbf{y} + \mathbf{p})$
Rotation	$\text{rot } \mathbf{v}(\mathbf{x})$	$\text{rot } \mathbf{w}(\mathbf{y}) = \mathbf{B}^T \text{rot } \mathbf{v}(\mathbf{B}\mathbf{y} + \mathbf{p})$

**8.4.19 Folgerung.** *Der Wert von  $\text{div } \mathbf{v}(\mathbf{x})$  und  $\Delta f(\mathbf{x})$  ist vom Koordinatensystem unabhängig. Dasselbe gilt für die Länge und die Richtung von  $\text{rot } \mathbf{v}(\mathbf{x})$  und  $\text{grad } f(\mathbf{x})$ .*

- Die einzelnen Koordinaten von  $\text{grad } f$  und  $\text{rot } \mathbf{v}$  hängen natürlich doch von der Basis  $\mathbf{B}$  ab.

**8.4.20 Lemma.** *Es gelten folgende Rechenregeln:*

a)  $\text{rot}(\text{grad } f) = 0,$

b)  $\text{div}(\text{rot } \mathbf{v}) = 0,$

$$c) \operatorname{div}(\operatorname{grad} f) = \Delta f,$$

$$d) \operatorname{div}(f\mathbf{v}) = \operatorname{grad} f \cdot \mathbf{v} + f \operatorname{div} \mathbf{v},$$

$$e) \operatorname{rot}(f\mathbf{v}) = \operatorname{grad} f \times \mathbf{v} + f \operatorname{rot} \mathbf{v},$$

$$f) \operatorname{rot}(\operatorname{rot} \mathbf{v}) = \operatorname{grad}(\operatorname{div} \mathbf{v}) - \Delta \mathbf{v}.$$

**8.4.21** Nablakalkül. 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

$$\operatorname{grad} f = \nabla f,$$

$$\operatorname{div} \mathbf{v} = \nabla \cdot \mathbf{v},$$

$$\Delta f = \nabla \cdot \nabla f,$$

$$\operatorname{rot} \mathbf{v} = \nabla \times \mathbf{v}.$$

Der Nablakalkül erlaubt die Formeln aus 8.4.20 formal zu berechnen. Dies sei am folgenden Beispiel illustriert:

$$\operatorname{div}(f\mathbf{v}) = \nabla \cdot (f\mathbf{v}) = \nabla \cdot (\overline{f}\mathbf{v}) + \nabla \cdot (f\overline{\mathbf{v}})$$

*umformen bis alle Größen ohne  
Querstrich links von  $\nabla$  stehen*

$$= \mathbf{v} \cdot \nabla \overline{f} + f \nabla \cdot \overline{\mathbf{v}}$$

$$= \mathbf{v} \cdot \operatorname{grad} f + f \operatorname{div} \mathbf{v}$$

Dabei sind folgende Identitäten hilfreich

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

$$\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = 0,$$

$$\mathbf{a} \times \mathbf{a} = 0,$$

$$\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2,$$

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}),$$

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d}).$$



# Kapitel 9

## Integration von Funktionen mehrerer Variablen

### 9.1 Parameterintegrale

Viele Funktionen der höheren Analysis besitzen Integraldarstellungen, die von einem Parameter abhängen, d.h. es gibt eine Darstellung

$$F(x) = \int_c^d f(x, y) dy. \quad (1.1)$$

wobei  $x$  auf der rechten Seite ein Parameter ist. Integrale dieser Form heißen **Parameterintegrale**.

Typische Beispiele sind:

Gammafunktion:  $\Gamma(x) := \int_0^{\infty} t^{x-1} e^{-t} dt, \quad x > 0,$

Besselfunktion:  $J_n(x) := \frac{1}{\pi} \int_0^{\pi} \cos(x \sin t - nt) dt, \quad n \in \mathbb{Z},$

Laplace Transformierte:  $F(x) := \int_0^{\infty} f(t) e^{-xt} dt.$

**9.1.1 Satz.** Sei  $D = [a, b] \times [c, d] = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$  ein abgeschlossenes Rechteck im  $\mathbb{R}^2$  und  $f : D \rightarrow \mathbb{R}$  stetig. Dann gilt für die

Integralfunktion

$$F(x) = \int_c^d f(x, y) dy, \quad x \in [a, b],$$

a)  $F$  ist in  $[a, b]$  stetig.

b) Satz von Fubini

$$\int_c^b F(x) dx = \int_a^b \left( \int_c^d f(x, y) dy \right) dx = \int_c^d \left( \int_a^b f(x, y) dx \right) dy, \quad (1.2)$$

c) ist  $f$  auf  $[a, b]$  stetig partiell nach  $x$  differenzierbar, dann ist  $F$  differenzierbar mit der Ableitung

$$F'(x) = \frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d \frac{\partial f}{\partial x}(x, y) dy. \quad (1.3)$$

**9.1.2 Satz.** Seien  $g, h$   $C^1$ -Funktionen und  $f$  stetig partiell nach  $x$  differenzierbar. Dann gilt:

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, y) dy = \int_{g(x)}^{h(x)} f_x(x, y) dy + f(x, h(x))h'(x) - f(x, g(x))g'(x).$$

**Beispiel:** Sei  $x(t) = \frac{1}{k} \int_0^t f(u) \sin(k(t-u)) du$ , mit  $f: \mathbb{R} \rightarrow \mathbb{R}$  stetig. Dann

gilt:

$$\dot{x}(t) = \int_0^t f(u) \cos(k(t-u)) du + \underbrace{f(t) \sin 0}_{=0} - 0$$

$$\ddot{x}(t) = -k \int_0^t f(u) \sin(k(t-u)) du - f(t),$$

d.h.  $x$  ist eine Lösung der Schwingungsgleichung

$$\ddot{x}(t) + k^2 x(t) = f(t).$$

**9.1.3 Satz.** *Ist  $f$  auf dem einseitig offenen Rechteck*

$$D = \{(x, y); a \leq x \leq b, c \leq y < d\}, \quad d \in \mathbb{R} \cup \{\infty\}$$

*stetig und stetig partiell nach  $x$  differenzierbar und gibt es Funktionen  $g, h : [c, d) \rightarrow \mathbb{R}$  für die gilt*

$$1) |f(x, y)| \leq g(y), \quad |f_x(x, y)| \leq h(y) \quad \forall (x, y) \in D,$$

2) *Die uneigentlichen Integrale*

$$\int_c^d g(y) dy, \quad \int_c^d h(y) dy$$

*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*

$$F'(x) = \int_c^d f_x(x, y) dy.$$

- Falls das Integral in (1.1) uneigentlich ist, zum Beispiel

$$\int_c^d f(x, y) dy = \lim_{\varepsilon \rightarrow 0} \int_c^{d-\varepsilon} f(x, y) dy$$

oder

$$\int_c^\infty f(x, y) dy = \lim_{d \rightarrow \infty} \int_c^d f(x, y) dy,$$

gilt Satz 9.1.1 nur unter zusätzlichen Voraussetzungen. Analoges gilt, falls das Integral am unteren Ende uneigentlich ist.

## 9.2 Kurvenintegrale

**9.2.1 Definition.** Sei  $D \subseteq \mathbb{R}^n$  offen,  $\mathbf{w} : [a, b] \rightarrow D$  ein Kurvenstück und  $f$  ein Skalarfeld auf dem Kurvenstück, so daß  $t \rightarrow f(\mathbf{w}(t))$  stetig ist. Dann heißt

$$\int_{\mathbf{w}} f ds := \int_a^b f(\mathbf{w}(t)) \|\dot{\mathbf{w}}(t)\| dt \quad (2.1)$$

das **Kurvenintegral** von  $f$  längs  $\mathbf{w}$ .

**9.2.2 Interpretation.** Sei  $\mathbf{w}(t)$  ein Kurvenstück und  $f$  ein skalares Feld. Das Kurvenstück wird durch eine Zerlegung des Parameterintervalls  $a = t_0 < t_1 < \dots < t_m = b$  in Bögen über  $[t_{i-1}, t_i]$  der Länge

$$\Delta s_i = \int_{t_{i-1}}^{t_i} \|\dot{\mathbf{w}}(t)\| dt = \|\dot{\mathbf{w}}(\tau_i)\| \Delta t_i$$

mit  $\tau_i \in (t_{i-1}, t_i)$  zerlegt. Durch

$$\sum_{i=1}^m f(\mathbf{w}(\tau_i)) \Delta s_i = \sum_{i=1}^n f(\mathbf{w}(\tau_i)) \|\dot{\mathbf{w}}(\tau_i)\| \Delta t_i \quad (*)$$

ist eine Riemann-Summe gegeben, die gegen

$$\int_a^b f(\mathbf{w}(t)) \|\dot{\mathbf{w}}(t)\| dt$$

konvergiert. Andererseits approximiert die linke Seite in (\*) das Kurvenintegral  $\int_{\mathbf{w}} f ds$  (Vergleiche Kapitel 4, 4.5.17). Mit der Sprechweise aus Kapitel 4 ist  $ds = \|\dot{\mathbf{w}}(t)\|$  das Längenelement, siehe auf Definition 8.1.5. Das Kurvenintegral hängt nicht von der Parametrisierung der Kurve ab.

### 9.2.3 Definition.

- a) Unter einer **Kurve**  $\mathbf{w}$  in  $D \subseteq \mathbb{R}^n$  verstehen wir eine endliche Folge  $\mathbf{w}_1, \dots, \mathbf{w}_k$  von Kurvenstücken  $\mathbf{w}_i : [a_i, b_i] \rightarrow D, i = 1, \dots, k$ , mit  $\mathbf{w}_i(b_i) = \mathbf{w}_{i+1}(a_{i+1}), 1 \leq i \leq k-1$ , siehe Definition 8.1.4.

- b) Für eine aus den Kurvenstücken  $\mathbf{w}_1, \dots, \mathbf{w}_k$  bestehende Kurve und für jede stetige Funktion  $f : \cup_{i=1}^k \{\mathbf{w}_i(t), t \in [a_i, b_i]\} \rightarrow \mathbb{R}$  ist das **Kurvenintegral** von  $f$  längs  $\mathbf{w}$  definiert durch

$$\int_{\mathbf{w}} f \, ds := \sum_{i=1}^k \int_{\mathbf{w}_i} f \, ds$$

- Für jedes Kurvenstück  $\mathbf{w} : [a_i, b_i] \rightarrow \mathbb{R}^n$  ergibt sich eine „Durchlaufrichtung“ vom Anfangspunkt  $\mathbf{w}_i(a_i)$  zum Endpunkt  $\mathbf{w}_i(b_i)$ . Man kann die Kurve auch in umgekehrter Richtung durchlaufen, das heißt, man betrachtet  $\mathbf{w}_i^* : [a_i, b_i] \rightarrow \mathbb{R}^n$ , definiert durch

$$\mathbf{w}^*(t) := \mathbf{w}(a + b - t).$$

Dieses Kurvenstück hat den Anfangspunkt  $\mathbf{w}_i^*(a_i) = \mathbf{w}_i(b_i)$  und den Endpunkt  $\mathbf{w}_i^*(b_i) = \mathbf{w}_i(a_i)$ . In die Definition des Integrals ist die Richtung nicht eingeflossen. Also gilt:

$$\int_{\mathbf{w}} f \, ds = \int_{\mathbf{w}^*} f \, ds.$$

**9.2.4 Lemma.** Das Kurvenintegral ist **linear** im Integranden, das heißt, für alle Kurven  $\mathbf{w}$  in  $D$ , stetige Funktionen  $f, g$  und  $\alpha \in \mathbb{R}$  gilt:

$$\begin{aligned} \int_{\mathbf{w}} \alpha f \, ds &= \alpha \int_{\mathbf{w}} f \, ds, \\ \int_{\mathbf{w}} (f + g) \, ds &= \int_{\mathbf{w}} f \, ds + \int_{\mathbf{w}} g \, ds. \end{aligned}$$

**9.2.5 Anwendungen.**

- a) **Bogenlänge:** Die Länge  $L(\mathbf{w})$  eines Kurvenstückes  $\mathbf{w} : [a, b] \rightarrow \mathbb{R}^3$  beträgt (Kapitel 8, Definition 8.1.5)

$$L(\mathbf{w}) = \int_{\mathbf{w}} ds = \int_a^b \|\dot{\mathbf{w}}(t)\| \, dt = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \, dt.$$

Eine aus den Kurvenstücken  $\mathbf{w}_1, \dots, \mathbf{w}_k$  bestehende Kurve  $\mathbf{w}$  hat somit die Länge

$$L = \sum_{i=1}^k L(\mathbf{w}_i) = \int_{\mathbf{w}} ds.$$

b) **Schwerpunkt:** Der Schwerpunkt  $\bar{S} = (\bar{x}, \bar{y}, \bar{z})^T$  einer Kurve  $\mathbf{w}$  hat bei gleichverteilter Masse die Koordinaten

$$\bar{x} = \frac{1}{L(\mathbf{w})} \int_{\mathbf{w}} x \, ds, \quad \bar{y} = \frac{1}{L(\mathbf{w})} \int_{\mathbf{w}} y \, ds, \quad \bar{z} = \frac{1}{L(\mathbf{w})} \int_{\mathbf{w}} z \, ds.$$

Wir definieren jetzt das Integral von Vektorfeldern über Kurven. Es spielt sowohl in der Mathematik als auch in der Physik eine wichtige Rolle.

**9.2.6 Definition.** Sei  $D \subseteq \mathbb{R}^n$  offen,  $\mathbf{w} : [a, b] \rightarrow D$  eine  $C^1$ -Kurve und  $\mathbf{v} : D \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld. Man nennt

$$\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} := \int_a^b \mathbf{v}(\mathbf{w}(t)) \cdot \dot{\mathbf{w}}(t) \, dt$$

das **Kurvenintegral von  $\mathbf{v}$  längs  $\mathbf{w}$** .

- Das Kurvenintegral ist linear im Vektorfeld, das heißt, für Vektorfelder  $\mathbf{v}$ ,  $\mathbf{u}$  und  $\alpha \in \mathbb{R}$  gilt

$$\begin{aligned} \int_{\mathbf{w}} (\mathbf{u} + \mathbf{v}) \cdot d\mathbf{x} &= \int_{\mathbf{w}} \mathbf{u} \cdot d\mathbf{x} + \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}, \\ \int_{\mathbf{w}} (\alpha \mathbf{v}) \cdot d\mathbf{x} &= \alpha \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}, \end{aligned} \tag{2.2}$$

- Mit der Durchlaufrichtung wechselt dieses Kurvenintegral das Vorzeichen, also

$$\int_{\mathbf{w}^*} \mathbf{v} \cdot d\mathbf{x} = - \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}.$$

- Analog zu Definition 9.2.3 werden Kurvenintegrale von Vektorfeldern längs stückweise differenzierbarer Kurven definiert.
- Man nennt eine Kurve **geschlossen**, wenn der Anfangspunkt mit dem Endpunkt zusammenfällt. In diesem Falle schreiben wir:

$$\oint_{\mathbf{w}} f \, ds \quad \text{und} \quad \oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}$$

### 9.2.7 Berechnung des Kurvenintegrals im $\mathbb{R}^3$ .

Zur Berechnung des Kurvenintegrals  $A = \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}$  gehen wir wie folgt vor:

1) Das Kurvenstück parametrisieren

$$\mathbf{w}(t) = (x(t), y(t), z(t))^T \quad t \in [a, b].$$

2) Das vektorielle Bogenelement berechnen

$$d\mathbf{x} = \dot{\mathbf{w}}(t) \, dt = (\dot{x}(t), \dot{y}(t), \dot{z}(t))^T \, dt$$

3) Einsetzen

$$A = \int_a^b (v_1(\mathbf{w}(t))\dot{x}(t) + v_2(\mathbf{w}(t))\dot{y}(t) + v_3(\mathbf{w}(t))\dot{z}(t)) \, dt$$

und ausrechnen.

**Beispiel:** Ein Wirbel im  $\mathbb{R}^2$  (vergleiche Beispiel 8.4.14) hat die Darstellung:

$$\mathbf{v}(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix}, \quad (x, y) \neq (0, 0).$$

Das Integral des Wirbels längs des positiv durchlaufenen Einheitskreises  $\mathbf{w}$  um  $(0, 0)^T$  berechnet sich wie folgt:

$$1. \quad x = \cos t, y = \sin t, \quad 0 \leq t \leq 2\pi$$

$$2. \quad d\mathbf{x} = (-\sin t, \cos t)^T \, dt$$

$$3. A = \oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = \int_0^{2\pi} \frac{-\sin t \cdot (-\sin t) + \cos t \cos t}{\cos^2 t + \sin^2 t} dt = 2\pi.$$

### 9.2.8 Anwendungen.

- a) **Arbeit.** Die von einer Kraft  $\mathbf{K}$  im Kurvenpunkt  $\mathbf{w}(t)$  längs des Weges  $d\mathbf{x} = \dot{\mathbf{w}}(t) dt$  geleistete Arbeit beträgt  $dA = \mathbf{K} \cdot d\mathbf{x} = \mathbf{K} \cdot \mathbf{T} ds$ . Also ist die Arbeit von  $\mathbf{K}$  längs  $\mathbf{w}$  gegeben durch

$$A = \int_{\mathbf{w}} dA = \int_{\mathbf{w}} \mathbf{K} \cdot d\mathbf{x} = \int_{\mathbf{w}} \mathbf{K} \cdot \mathbf{T} ds.$$

In einem elektrischen Spannungsfeld  $\mathbf{E}$  liefert das „negative Arbeitsintegral“ den Spannungsabfall längs  $\mathbf{w}$ . Also gilt

$$U = - \int_{\mathbf{w}} \mathbf{E} \cdot d\mathbf{x}.$$

- b) **Zirkulation** In der Strömungsmechanik heißt das Integral eines Geschwindigkeitsfeldes  $\mathbf{v}$  längs einer geschlossenen Kurve  $\mathbf{w}$

$$\oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = \oint_{\mathbf{w}} \mathbf{v} \cdot \mathbf{T} ds$$

die **Zirkulation** von  $\mathbf{v}$  längs  $\mathbf{w}$ .

**Beispiel:** Für die ebene Gegenströmung  $\mathbf{v} = (cy, 0)$ ,  $c > 0$  und den positiv durchlaufenen Kreis  $\mathbf{w}(t) = (\cos t, 1 + \sin t)^T$ ,  $t \in [0, 2\pi]$  gilt:

$$\begin{aligned} \oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} &= \int_0^{2\pi} \begin{pmatrix} c(1 + \sin t) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt \\ &= c \left[ \cos t - \frac{1}{2}t + \frac{1}{4} \sin 2t \right]_0^{2\pi} = -c\pi. \end{aligned}$$

Die mittlere Tangentialkomponente ist  $-\frac{c}{2}$ , d.h. die Strömung erfolgt überwiegend im Uhrzeigersinn.

**9.2.9 Definition.** Eine Teilmenge  $G \subseteq \mathbb{R}^n$  heißt **Gebiet**, wenn gilt:

1.  $G$  ist **offen**, d.h. für alle  $\mathbf{x}_0 \in G$  existiert eine  $\varepsilon$ -Umgebung  $B_\varepsilon(\mathbf{x}_0)$  so, dass  $B_\varepsilon(\mathbf{x}_0) \subseteq G$ ,
2.  $G$  ist **zusammenhängend**, d.h. zu je zwei Punkten  $\mathbf{x}_0, \mathbf{y}_0 \in G$  gibt es eine Kurve  $\mathbf{w} : [a, b] \rightarrow G$  mit  $\mathbf{w}(a) = \mathbf{x}_0, \mathbf{w}(b) = \mathbf{y}_0$ .

**9.2.10 Definition.** Sei  $G \subseteq \mathbb{R}^n$  ein Gebiet. Man nennt ein Vektorfeld  $\mathbf{v} \in C^0(G, \mathbb{R}^n)$  **konservativ** oder ein **Potential-** bzw. **Gradientenfeld**, wenn es eine Funktion  $f \in C^1(G, \mathbb{R})$  gibt, mit

$$\mathbf{v}(\mathbf{x}) = \text{grad } f(\mathbf{x}) \quad \text{für alle } \mathbf{x} \in G.$$

In diesem Fall heißt  $f$  **Stammfunktion** und  $U := -f$  ein **Potential** von  $\mathbf{v}$ .

- Falls  $n = 1$  wissen wir, dass

$$\int_0^x f'(x) dx \quad (*)$$

eine Stammfunktion von  $f'(x)$  ist. Im allgemeinen Fall  $G \subseteq \mathbb{R}^n$  wird das Integral in (\*) durch ein Kurvenintegral ersetzt.

- In einem konservativen Kraftfeld ist das Potential die potentielle Energie.
- Für jede durch ein konservatives Kraftfeld  $\mathbf{K} = -\text{grad } U$  hervorgerufene Bewegung  $\mathbf{x}(t)$  leitet man aus den **Newtonschen Bewegungsgleichungen** die Energieerhaltung ab:

$$\begin{aligned} \mathbf{0} &= m\ddot{\mathbf{x}} - \mathbf{K} = m\ddot{\mathbf{x}} + \text{grad } U, \\ 0 &= m\ddot{\mathbf{x}} \cdot \dot{\mathbf{x}} + \text{grad } U \cdot \dot{\mathbf{x}}, \\ &= \frac{d}{dt} \left( \frac{m}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} + U(\mathbf{x}(t)) \right), \end{aligned}$$

d.h. die Summe aus kinetischer und potentieller Energie ist konstant.

**9.2.11 Satz.** Ist  $\mathbf{v} : G \rightarrow \mathbb{R}^n$  ein stetiges Potentialfeld auf dem Gebiet  $G \subseteq \mathbb{R}^n$  mit Stammfunktion  $f$ , dann gilt für jede stückweise  $C^1$ -Kurve  $\mathbf{w}$  in  $G$  mit Anfangspunkt  $\mathbf{w}(a)$  und Endpunkt  $\mathbf{w}(b)$ :

$$\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = f(\mathbf{w}(b)) - f(\mathbf{w}(a)).$$

**9.2.12 Satz.** Für ein Vektorfeld  $\mathbf{v} \in C^0(G, \mathbb{R}^n)$  auf einem Gebiet  $G \subseteq \mathbb{R}^n$  sind äquivalent:

- $\mathbf{v}$  ist ein Potentialfeld,
- Für alle stückweisen  $C^1$ -Kurven  $\mathbf{w}$  in  $G$  hängt  $\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}$  nur vom Anfangs- und Endpunkt von  $\mathbf{w}$  ab. Man sagt, das Integral ist **wegunabhängig**.
- Für geschlossene stückweise  $C^1$ -Kurven  $\mathbf{w}$  in  $G$  gilt

$$\oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = 0.$$

- Wir suchen ein praktisches Kriterium, um für ein gegebenes  $C^1$  Vektorfeld  $v$  zu entscheiden ob es ein Potentialfeld ist. Nahe eines Punktes  $\mathbf{x} \in G$  kann man dies direkt aus der Jacobi Matrix  $\left(\frac{\partial v_i}{\partial x_j}(\mathbf{x})\right)$  ablesen, für das gesamte Gebiet braucht man mehr.

**9.2.13 Definition.** Ein Gebiet  $G \in \mathbb{R}^n$  heißt **einfach zusammenhängend**, wenn jede geschlossene doppelpunktfreie Kurve in  $G$  stetig auf einem Punkt in  $G$  zusammengezogen werden kann, ohne dass  $G$  verlassen wird.

- Im  $\mathbb{R}^2$  sind Gebiete ohne Loch einfach zusammenhängend und Gebiete mit Loch nicht einfach zusammenhängend. Dasselbe gilt im  $\mathbb{R}^3$  für Gebiete mit bzw. ohne „Henkel“ (Tassen sind nicht einfach zusammenhängend). Ein Gebiet im  $\mathbb{R}^3$  bleibt einfach zusammenhängend, wenn man endlich viele Punkte herausnimmt, z.B.  $\mathbb{R}^3 \setminus \{0\}$ . Wenn man allerdings eine Gerade herausnimmt, ist es nicht mehr einfach zusammenhängend, z.B.  $\mathbb{R}^3 \setminus \text{Gerade}$ .

**9.2.14 Satz.** Ein  $C^1$ -Vektorfeld  $\mathbf{v} : G \rightarrow \mathbb{R}^n$  auf dem einfach zusammenhängenden Gebiet  $G \subseteq \mathbb{R}^n$  ist genau dann ein Potentialfeld, wenn die „Integrabilitätsbedingung“

$$J_{\mathbf{v}}(\mathbf{x}) = J_{\mathbf{v}}(\mathbf{x})^T$$

für alle  $\mathbf{x} \in G$  erfüllt ist.

**Gegenbeispiel:** Sei  $G = \mathbb{R}^2 \setminus \{0\}$  und  $\mathbf{v} = (x, y) = \frac{1}{x^2+y^2} \begin{pmatrix} -y \\ x \end{pmatrix}$  ein Wirbel.

Es gilt:  $\frac{\partial v_1}{\partial y} = \frac{\partial v_2}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2}$ , aber  $\mathbf{v}$  besitzt auf  $G$  **kein** Potential! Wenn dem so wäre, müsste nach Satz 9.2.12 für einen Kreis mit Radius 1 um den Nullpunkt  $\oint \mathbf{v} \cdot d\mathbf{x} = 0$  gelten. Aber das Beispiel nach 9.2.7 liefert

$$\oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = 2\pi \neq 0.$$

Immerhin liefert aber der Winkel  $f = \arctan \frac{y}{x}$  zur  $x$ -Achse auf einer großen offenen Menge eine Stammfunktion.

**9.2.15 Folgerung.** Ein  $C^1$ -Vektorfeld  $\mathbf{v} : G \rightarrow \mathbb{R}^3$  auf dem einfach zusammenhängenden Gebiet  $G \subseteq \mathbb{R}^3$  ist genau dann ein Potentialfeld, wenn  $\operatorname{rot} \mathbf{v} = \mathbf{0}$  für alle  $\mathbf{x} \in G$ .

**9.2.16** Berechnung des Potentials. Sei  $\mathbf{v}$  ein  $C^1$ -Vektorfeld definiert auf  $G \subseteq \mathbb{R}^3$ .

- Überprüfe ob  $\operatorname{rot} \mathbf{v} = \mathbf{0}$ . Falls es ein  $\mathbf{x} \in G$  gibt mit  $\operatorname{rot} \mathbf{v}(\mathbf{x}) \neq \mathbf{0}$  (allgemeiner:  $J_{\mathbf{v}}(\mathbf{x}) \neq J_{\mathbf{v}}(\mathbf{x})^T$ ), dann hat  $\mathbf{v}$  kein Potential.
- Falls  $\operatorname{rot} \mathbf{v} = \mathbf{0}$  und  $G$  einfach zusammenhängend, dann wählt man ein  $\mathbf{x}_0 \in G$  und für beliebige  $\mathbf{x} \in G$  eine  $\mathbf{x}_0$  und  $\mathbf{x}$  verbindende Kurve. Dann ist

$$f(\mathbf{x}) := \int_{\mathbf{x}_0}^{\mathbf{x}} \mathbf{v} \cdot d\mathbf{x}$$

ein Potential.

Man sollte die Wege so wählen, dass die Integrale so einfach wie möglich werden, zum Beispiel:

- Strecke  $\mathbf{w}(t) := \mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)$ ,  $0 \leq t \leq 1$ , falls sie in  $G$  verläuft.  
Dann gilt:

$$f(\mathbf{x}) = \int_0^1 \mathbf{v}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \cdot (\mathbf{x} - \mathbf{x}_0) dt.$$

- Streckenzug parallel zu den Koordinatenachsen, z.B.

$$f(x, y, z) = \int_{x_0}^x v_1(t, y_0, z_0) dt + \int_{y_0}^y v_2(x, t, z_0) dt + \int_{z_0}^z v_3(x, y, t) dt.$$

**9.2.17 Ansatzmethode.** Man kann die 3 Differentialgleichungen  $\text{grad } f = \mathbf{v}$  durch dreimaliges unbestimmtes Integrieren lösen.

- a) Überprüfe ob  $G$  einfach zusammenhängend ist. Mache den Ansatz  $f_x = v_1, f_y = v_2, f_z = v_3$ .
- b) Unbestimmte Integration nach  $x$  ( $y, z$  fest) von  $v_1$  liefert

$$f(x, y, z) = \int v_1(x, y, z) dx + c(y, z), \quad (*)$$

wobei die „Integrationskonstante“  $c$  von  $y$  und  $z$  abhängt.

- c) (\*) nach  $y$  ableiten und durch Ansatz  $f_y = v_2$  eine Gleichung für  $c_y(y, z)$  herleiten, d.h.:

$$\frac{\partial}{\partial y} \int v_1(x, y, z) dx + \frac{\partial c}{\partial y}(y, z) = \frac{\partial f}{\partial y} = v_2. \quad (+)$$

Tritt in der Gleichung für  $c$  noch  $x$  explizit auf dann ist  $\text{rot } \mathbf{v} \neq \mathbf{0}$  und  $\mathbf{v}$  besitzt kein Potential.

d) Durch unbestimmte Integration nach  $y$  nun  $c(y, z)$  bestimmen, d.h.:

$$c(y, z) = \int h(y, z) dy + q(z),$$

wobei  $h(y, z) = v_2 - \int \partial_y v_1 dx$  und die „Integrationskonstante“  $q$  von  $z$  abhängt. Das Ergebnis in (\*) eingesetzt liefert:

$$f(x, y, z) = \int v_1(x, y, z) dx + \int h(y, z) dy + q(z) \quad (**)$$

e) Eine Gleichung für  $q(z)$  aus (\*\*) durch Ableitung nach  $z$  herleiten, d.h.:

$$\frac{\partial}{\partial z} \int v_1(x, y, z) dx + \frac{\partial}{\partial z} \int h(y, z) dy + q'(z) = v_3. \quad (++)$$

Tritt in der Gleichung für  $q$  noch  $y$  explizit auf, so ist  $\operatorname{rot} \mathbf{v} \neq \mathbf{0}$  und  $\mathbf{v}$  besitzt kein Potential.

f) Berechne  $q(z)$  durch unbestimmte Integration von (++) nach  $z$ . Einsetzen des Ergebnisses in (\*\*) liefert das Potential.

### Beispiele:

a) **Zentralkraftfelder** haben die Form

$$\mathbf{v}(\mathbf{x}) = \varphi(\|\mathbf{x} - \mathbf{z}\|)(\mathbf{x} - \mathbf{z}),$$

mit einer Funktion  $\varphi \in C^1(\mathbb{R} \setminus \{0\}, \mathbb{R})$ , die meist in „0“ einen Pol hat.  $\mathbf{v}$  ist auf  $\mathbb{R}^3 \setminus \{z\}$  ein  $C^1$ -Vektorfeld. Wir setzen der Einfachheit halber  $z = 0$ , d.h.

$$\mathbf{v}(\mathbf{x}) = \varphi(\|\mathbf{x}\|)\mathbf{x}.$$

Da  $G = \mathbb{R}^3 \setminus \{0\}$  ein einfach zusammenhängendes Gebiet ist und nach Kapitel 8 Lemma 8.4.8 gilt

$$J_{\mathbf{v}}(\mathbf{x}) = \varphi'(\|\mathbf{x}\|) \frac{\mathbf{x}\mathbf{x}^T}{\|\mathbf{x}\|} + \varphi(\|\mathbf{x}\|)\mathbf{E} = J_{\mathbf{v}}(\mathbf{x})^T,$$

besitzt  $\mathbf{v}$  ein Potential. Berechnung nach 9.2.16:

Sei  $\mathbf{x}_0$  fest und  $\mathbf{x}$  beliebig. Der Integrationsweg  $\mathbf{w}$  besteht aus zwei Stücken  $\mathbf{w}_1, \mathbf{w}_2$ .  $\mathbf{w}_1$  ist ein Teil der Sphäre um den Nullpunkt mit dem Radius  $\|\mathbf{x}_0\|$  der  $\mathbf{x}_0$  und  $\mathbf{x}_1 = \frac{\|\mathbf{x}_0\|}{\|\mathbf{x}\|} \mathbf{x}$  verbindet.  $\mathbf{w}_2$  ist die Strecke  $\mathbf{w}_2(t) = t \frac{\mathbf{x}}{\|\mathbf{x}\|}, \|\mathbf{x}_0\| \leq t \leq \|\mathbf{x}\|$ . Es gilt  $\mathbf{v}(\mathbf{w}_1(t)) \cdot \dot{\mathbf{w}}_1(t) = 0$

$$f(\mathbf{x}) = \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = \int_{\mathbf{w}_2} \mathbf{v} \cdot d\mathbf{x} = \int_{\|\mathbf{x}_0\|}^{\|\mathbf{x}\|} \varphi(t) t dt.$$

Sei  $\varphi(t) = \frac{1}{t^3}$ , das entspricht dem Newtonschen Gravitationsfeld, dann gilt:

$$U(\mathbf{x}) = -f(\mathbf{x}) = - \int_{\|\mathbf{x}_0\|}^{\|\mathbf{x}\|} \frac{t}{t^3} dt = \frac{1}{\|\mathbf{x}\|} - \frac{1}{\|\mathbf{x}_0\|}.$$

b) Sei

$$\mathbf{v}(x, y, z) = \begin{pmatrix} y^2 \cos x \\ 2y \sin x + e^{2z} \\ 2ye^{2z} \end{pmatrix}, \quad G = \mathbb{R}^3.$$

$G$  ist einfach zusammenhängend. Also können wir das Potential nach 9.2.17 berechnen:

a)  $f_x = y^2 \cos x, f_y = 2y \sin x + e^{2z}, f_z = 2ye^{2z}$ .

b)

$$\begin{aligned} f(x, y, z) &= \int y^2 \cos x dx + c(y, z) \\ &= y^2 \sin x + c(y, z). \end{aligned}$$

c) Die Gleichung für  $c(y, z)$  lautet  $f_y = 2y \sin x + c_y(y, z) = 2y \sin x + e^{2z}$  woraus folgt

$$c_y(y, z) = e^{2z}.$$

d) Somit haben wir  $c(y, z) = ye^{2z} + q(z)$  und erhalten

$$f(x, y, z) = y^2 \sin x + ye^{2z} + q(z).$$

e) Die Gleichung für  $q$  lautet

$$q'(z) + 2ye^{2z} = 2ye^{2z},$$

woraus folgt  $q'(z) = 0$ , also  $q(z) = \text{const.}$  Also ist

$$f(x, y, z) = y^2 \sin x + ye^{2z} + \text{const.}$$

ein Potential von  $\mathbf{v}$ .

## 9.3 Integration über ebene Bereiche

**9.3.1** Der Flächeninhalt. Der Flächeninhalt von Mengen im  $\mathbb{R}^2$  wird durch Zerlegung auf Flächeninhalte von Rechtecken zurückgeführt. Sei  $k \in \mathbb{N}$ . Durch das Gitter achsenparalleler Koordinatenlinien

$$x = n \times 2^{-k}, \quad y = n \times 2^{-k}, \quad n = 0, \pm 1, \pm 2, \dots,$$

wird die  $(x, y)$ -Ebene in Quadrate mit dem Flächeninhalt  $2^{-2k}$  zerlegt. Sei  $M \subseteq \mathbb{R}^2$  eine beliebige beschränkte Menge. Sei  $s_k(M)$  die Summe der Flächeninhalte aller Quadrate, die einschließlich des Randes in  $M$  liegen und  $S_k(M)$  der Flächeninhalt der Quadrate die mindestens einen Punkt aus  $M$  enthalten. Dann gilt:

$$s_k(M) \leq s_{k+1}(M), \quad s_k(M) \leq S_k(M), \quad S_{k+1}(M) \leq S_k(M).$$

Also sind  $s_k(M)$  und  $S_k(M)$  monotone Folgen, und es gilt:

$$\lim_{k \rightarrow \infty} s_k(M) =: F_i(M) \leq F_a(M) := \lim_{k \rightarrow \infty} S_k(M),$$

wobei  $F_i(M)$  der innere Inhalt von  $M$  und  $F_a(M)$  der äußere Inhalt von  $M$  ist.

**9.3.2 Definition.** Man nennt eine beschränkte Menge  $M$  des  $\mathbb{R}^2$  **Riemann-messbar**, wenn  $F_i(M) = F_a(M)$  gilt. In diesem Falle ist  $F(M) = F_i(M) = F_a(M)$  der **Flächeninhalt** von  $M$ .

- $M = \{(x, y); x, y \in [0, 1], x, y \in \mathbb{Q}\}$  ist **nicht** Riemann-messbar, da  $F_i(M) = 0$  und  $F_a(M) = 1$ .

- Nachdem wir das Integral eingeführt haben, betrachten wir eine Reihe von Beispielen Riemann-messbarer Mengen.
- Sei  $M$  Riemann-messbar und  $N$  bestehe aus endlich vielen Punkten und endlich vielen regulären Kurvenstücken. Dann gilt:

$$F(M) = F(M \cup N) = F(M \setminus N). \quad (3.1)$$

Insbesondere ist  $N$  Riemann-messbar mit  $F(N) = 0$ , daher heißt  $N$  **Nullmenge**.

**9.3.3** Das Doppelintegral. Sei  $M$  eine Riemann-messbare Menge und  $f : M \rightarrow \mathbb{R}$  eine beschränkte, auf  $M$  stetige Funktion. Wir zerlegen die  $(x, y)$ -Ebene wie oben in achsenparallele Quadrate mit Flächeninhalt  $2^{-2k}$ . Es sei  $q_k(M)$  die Menge dieser Quadrate, die ganz in  $M$  liegen. Für jedes Quadrat  $Q_i = [x_i, x_i + 2^{-k}] \times [y_i, y_i + 2^{-k}] \in q_k(M)$  mit  $x_i, y_i \in 2^{-k}\mathbb{Z}$  sei  $(x_i^*, y_i^*) \in Q_i$ . Dann bilden wir die **Riemann-Summe**

$$\sum_i f(x_i^*, y_i^*) 2^{-2k}. \quad (3.2)$$

Für  $k \rightarrow \infty$  konvergieren diese Riemann-Summen gegen einen Grenzwert unabhängig von der Wahl der Punkte  $(x_i^*, y_i^*)$ , der als **Doppelintegral** bezeichnet wird,

$$\iint_M f \, dF := \lim_{k \rightarrow \infty} \sum_i f(x_i^*, y_i^*) 2^{-2k}. \quad (3.3)$$

- Wenn es eine Nullmenge  $N \subset M$  gibt, auf der  $f$  nicht stetig ist, ist  $M \setminus N$  immer noch Riemann-messbar mit  $F(M \setminus N) = F(M)$ , und wir setzen

$$\iint_M f \, df = \iint_{M \setminus N} f \, dF. \quad (3.4)$$

### 9.3.4 Integration.

- a) **Flächeninhalt:** Mit  $f(x, y) = 1$  approximieren die Riemann-Summen (5.1) den Flächeninhalt von  $B$ , also

$$F = \iint_B dF.$$

b) **Volumen:** Ist  $f(x, y) \geq 0$  für alle  $(x, y) \in B$ , dann ist

$$v = \iint_B f(x, y) dF$$

das Volumen des senkrecht auf der  $(x, y)$ -Ebene stehenden Zylinderabschnitts mit der Grundfläche  $B$  und der Deckfläche  $z = f(x, y)$ . Die Summanden in (5.1) sind die Volumina von Quadern mit Grundfläche  $Q_i$  und Höhe  $f(x_i^*, y_i^*)$ , also approximiert die Summe das Gesamtvolumen.

**9.3.5 Lemma.** Es gelten folgende Rechenregeln:

$$a) \iint_B (\alpha f + \beta g) dF = \alpha \iint_B f dF + \beta \iint_B g dF,$$

$$b) f(x, y) \leq g(x, y) \Rightarrow \iint_B f dF \leq \iint_B g dF,$$

$$c) \iint_B f dF = \iint_{B_1} f dF + \iint_{B_2} f dF,$$

falls  $B$  durch eine stückweise reguläre Kurve in zwei Bereichen  $B_1$  und  $B_2$  zerlegt wird.

- Als **Mittelwert** der Funktion  $f$  auf  $B$  bezeichnet man die Zahl

$$\frac{1}{F(B)} \iint_B f dF,$$

wobei  $F(B) = \iint_B dF$  der Inhalt von  $B$  ist.

**9.3.6 Satz.** Ist  $B$  zusammenhängend und abgeschlossen, dann gibt es zu jeder stetigen Funktion  $f : B \rightarrow \mathbb{R}$  einen Punkt  $(x^*, y^*) \in B$  mit

$$\frac{1}{F(B)} \iint_B f dF = f(x^*, y^*).$$

- Für viele Gebiete lässt sich die Berechnung eines Doppelintegrals auf zwei Berechnungen von „Einfachintegralen“ zurückführen.

**9.3.7 Definition.** Wir nennen  $B_1 \subseteq \mathbb{R}^2$  einen **Normalbereich** vom Typ 1, wenn es  $a, b \in \mathbb{R}$  und stetige Funktionen  $g, h : [a, b] \rightarrow \mathbb{R}$  gibt mit  $g(x) \leq h(x)$  und

$$B_1 = \{ (x, y) \mid a \leq x \leq b, g(x) \leq y \leq h(x) \}.$$

Ein Normalbereich vom Typ 2 hat die Form

$$B_2 = \{ (x, y) \mid c \leq y \leq d, l(y) \leq x \leq r(y) \}$$

mit  $c, d \in \mathbb{R}$  und stetigen Funktionen  $l, r : [c, d] \rightarrow \mathbb{R}$  mit  $l(y) \leq r(y)$ .

**9.3.8 Satz.**

a) Für jede stetige Funktion  $f : B_1 \rightarrow \mathbb{R}$  auf einem Normalbereich vom Typ 1 gilt:

$$\iint_{B_1} f(x, y) \, dx \, dy = \int_a^b \left( \int_{g(x)}^{h(x)} f(x, y) \, dy \right) dx.$$

b) Für jede stetige Funktion  $f : B_2 \rightarrow \mathbb{R}$  auf einem Normalbereich vom Typ 2 gilt:

$$\iint_{B_2} f(x, y) \, dx \, dy = \int_c^d \left( \int_{l(y)}^{r(y)} f(x, y) \, dx \right) dy.$$

- Aus Satz 9.3.8a) folgt sofort eine Formel für den Flächeninhalt eines Normalbereiches vom Typ 1, nämlich

$$F = \iint_B dF = \int_a^b \int_{g(x)}^{h(x)} dy \, dx = \int_a^b (h(x) - g(x)) \, dx.$$

**9.3.9 Definition.** Ein Bereich  $B \subseteq \mathbb{R}^2$  heißt **regulär**, wenn

- der Rand  $\partial B$  aus endlich vielen regulären Kurvenstücken besteht,
- das Innere  $B \setminus \partial B$  ein nicht leeres Gebiet des  $\mathbb{R}^2$  ist,
- $B$  abgeschlossen und beschränkt ist, d.h.  $\partial B \subseteq B$ .

**9.3.10 Lemma.** Reguläre Bereiche sind Riemann-messbar.

- Ein regulärer Integrationsbereich  $B$  wird durch achsenparallele Schnitte in Normalbereiche  $B_1, \dots, B_n$  zerlegt. Damit haben wir:

$$\iint_B f \, dF = \sum_{i=1}^n \iint_{B_i} f \, dF.$$

- Besitzt  $B$  sowohl eine Darstellung als Normalbereich vom Typ 1 als auch vom Typ 2, dann gilt

$$\iint_B f \, dF = \int_a^b \int_{g(x)}^{h(x)} f(x, y) \, dy \, dx = \int_c^d \int_{l(y)}^{r(y)} f(x, y) \, dx \, dy. \quad (*)$$

Im Spezialfall eines Rechteckes ist (\*) nichts anderes als der Satz von Fubini (Satz 9.1.1b).

**Beispiel:**  $B$  sei von  $y = x$ ,  $xy = 1$  und  $y = 2$  berandet.

- a) Der Flächeninhalt: Darstellung als ein Normalgebiet vom Typ 2

$$B = \{(x, y); 1 \leq y \leq 2, \frac{1}{y} \leq x \leq y\}$$

$$F = \int_1^2 \int_{\frac{1}{y}}^y dx \, dy = \int_1^2 y - \frac{1}{y} \, dy = \frac{y^2}{2} - \ln y \Big|_1^2 = \frac{3}{2} - \ln 2$$

- b) Das Volumen des Zylinders mit der Grundfläche  $B$  und der Deckfläche  $z = \frac{x^2}{y^2}$  berechnet sich durch:

$$V = \int_1^2 \int_{\frac{1}{y}}^y \frac{x^2}{y^2} \, dx \, dy = \int_1^2 y^2 \left( -\frac{1}{x} \right) \Big|_{x=\frac{1}{y}}^{x=y} dy$$

$$= \int_1^2 -y + y^3 \, dy = \frac{9}{4}$$

**9.3.11 Anwendungen.** Sei  $f : B \rightarrow \mathbb{R}$  eine Belegungsfunktion. Mit den Sprechweisen aus Kapitel 4 haben wir folgendes Vorgehen:

- 1) Zerlegung von  $B$  in Flächenelemente  $dF$ ,
- 2) Das Flächenelement im Punkt  $(x, y)$  trägt die Belegung  $dM = f(x, y) dF$ ,
- 3) Gesamtbelegung von  $B$  ist

$$M = \iint_B dM = \iint_B f(x, y) dF.$$

- a) **Ladung** Die Ladung einer Platte  $B$  mit Ladungsdichte  $\varepsilon(x, y)$  ist

$$Q = \iint_B \varepsilon(x, y) dx dy.$$

- b) **Massenmittelpunkt** Ein Flächenelement  $dF$  besitzt bei der Massendichte  $\mu(x, y)$  die Masse  $dM = \mu dF$ . Die **axialen Momente** sind  $x\mu dF$  bzw.  $y\mu dF$ , d.h.

$$M_x = \iint_B x\mu(x, y) dF, \quad M_y = \iint_B y\mu(x, y) dF.$$

Der **Massenmittelpunkt** der Fläche  $B$  ist derjenige Punkt  $\mathbf{S} = (x_s, y_s)^T$ , in dem eine Punktmasse der Größe  $M = \iint \mu dF$  dieselben Momente besitzt wie  $B$ , d.h.

$$x_s = \frac{1}{M} \iint_B x\mu(x, y) dF, \quad y_s = \frac{1}{M} \iint_B y\mu(x, y) dF. \quad (3.5)$$

Der **geometrische Schwerpunkt**  $\bar{\mathbf{S}} = (\bar{x}, \bar{y})^T$  ergibt sich aus (3.5) mit  $\mu(x, y) = \mu_0 = \text{const.}$ , d.h.

$$\bar{x} = \frac{1}{F} \iint_B x dF, \quad \bar{y} = \frac{1}{F} \iint_B y dF.$$

- c) Sei  $Z$  ein Zylinder über einem regulären Bereich  $B$ , dessen obere und untere Grenzfläche beschrieben werden durch stetige Funktionen  $g \leq h: B \rightarrow \mathbb{R}$ . Dann hat der Schwerpunkt die Komponenten

$$\begin{aligned} x_s &= \frac{1}{V(Z)} \iint_B x (h(x, y) - g(x, y)) dF, \\ y_s &= \frac{1}{V(Z)} \iint_B y (h(x, y) - g(x, y)) dF, \\ z_s &= \frac{1}{V(Z)} \iint_B \frac{h(x, y)^2 - g(x, y)^2}{2} dF, \end{aligned}$$

wobei das Volumen  $V(Z)$  wie im Beispiel b) zu 9.3.10 berechnet wird.

**Beispiel:** Sei  $B$  begrenzt durch  $x = 2$ ,  $y = 1$  und  $y = x^2$ . Die Massenverteilung sei  $\mu(x, y) = x^2 + y^2$ .  $B$  ist ein Normalbereich mit der Darstellung  $B = \{(x, y); 1 \leq x \leq 2, 1 \leq y \leq x^2\}$

$$\begin{aligned} M &= \iint_B x^2 + y^2 \, dx \, dy = \int_2^1 \int_1^{x^2} x^2 + y^2 \, dy \, dx = \int_1^2 yx^2 + \frac{1}{3}y^3 \Big|_1^{x^2} \, dx \\ &= \int_1^2 \frac{1}{3}x^6 + x^4 - x^2 - \frac{1}{3} \, dx = \frac{1006}{105} \end{aligned}$$

$$\begin{aligned} M_x &= \int_1^2 \int_1^{x^2} x^3 + xy^2 \, dy \, dx = \int_1^2 yx^3 + \frac{1}{3}xy^3 \Big|_1^{x^2} \, dx \\ &= \int_1^2 \frac{1}{3}x^7 + x^5 - \frac{1}{3}x - x^3 \, dx = \frac{135}{8} \end{aligned}$$

Analog gilt

$$M_y = \iint_B yx^2 + y^3 \, dx \, dy = \frac{2753}{126}.$$

Für den geometrischen Schwerpunkt haben wir

$$\begin{aligned} F &= \int_1^2 x^2 - 1 \, dx = \frac{1}{3}x^3 - x \Big|_1^2 = \frac{4}{3} \\ \bar{x} &= \frac{3}{4} \int_1^2 \int_1^{x^2} x \, dy \, dx = \frac{3}{4} \int_1^2 xy \Big|_1^{x^2} \, dx = \frac{3}{4} \int_1^2 x^3 - x \, dx = \frac{27}{16} \end{aligned}$$

$$\bar{y} = \frac{3}{4} \int_1^2 \int_1^{x^2} y \, dy \, dx = \frac{3}{4} \int_1^2 \frac{1}{2}y^2 \Big|_1^{x^2} \, dx = \frac{3}{4} \int_1^2 \frac{x^4}{2} - \frac{1}{2} \, dx = \frac{39}{20}$$

- Sei  $B$  ein regulärer Bereich, dessen Rand  $\partial B$  aus endlich vielen geschlossenen stückweise glatten Kurven  $\mathbf{w}_1, \dots, \mathbf{w}_n$  besteht. Die Parametrisierung sei so, dass  $B$  stets **links** zur Durchlaufrichtung liegt.

Das Kurvenintegral eines Vektorfeldes  $\mathbf{v}$  über  $\partial B$  ist

$$\int_{\partial B} \mathbf{v} \cdot d\mathbf{x} = \sum_{i=1}^n \int_{\mathbf{w}_i} \mathbf{v} \cdot d\mathbf{x}.$$

**9.3.12 Satz (Green).** Seien  $B$  und  $\partial B$  wie oben beschrieben und  $D \subseteq \mathbb{R}^2$  eine offene Menge mit  $B \subseteq D$ , dann gilt für jedes  $C^1$ -Vektorfeld  $\mathbf{v} : D \rightarrow \mathbb{R}^2$ :

$$\int_{\partial B} \mathbf{v} \cdot d\mathbf{x} = \iint_B \left( \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dx dy.$$

- Der Integrand rechts gibt — analog zu  $\operatorname{rot} \mathbf{v}$  im  $\mathbb{R}^3$  — die **Wirbelstärke** des Vektorfeldes  $\mathbf{v}$  an. Also sagt der Satz von Green: die durchschnittliche Umlaufgeschwindigkeit auf dem Rand von  $B$  ist das Integral über die Wirbelstärke von  $\mathbf{v}$  im Inneren von  $B$ .

Es sei wieder  $B$  ein regulärer Bereich, dessen Rand aus Kurven  $\mathbf{w}_1, \mathbf{w}_n$  mit  $\mathbf{w}_i : [a_i, b_i] \rightarrow \mathbb{R}^2$  besteht, so dass  $B$  jeweils links zur Durchlaufrichtung liegt. Wir definieren den **äußeren Normaleneinheitsvektor**  $\mathbf{n}$  und das **äußere Normalelement**  $d\mathbf{n}$  von  $\mathbf{w}_i(t) = (x(t), y(t))^T$  durch

$$\mathbf{n}(t) = \frac{1}{\|\dot{\mathbf{w}}_i(t)\|} \begin{pmatrix} \dot{y}(t) \\ -\dot{x}(t) \end{pmatrix},$$

$$d\mathbf{n} = \mathbf{n} ds = \begin{pmatrix} \dot{y}(t) \\ -\dot{x}(t) \end{pmatrix} dt$$

Das **normale Kurvenintegral** eines Vektorfeldes  $\mathbf{v}$  ist dann definiert als

$$\begin{aligned} \int_{\partial B} \mathbf{v} \cdot d\mathbf{n} &= \sum_i \int_{\mathbf{w}_i} \mathbf{v} \cdot d\mathbf{n} = \sum_i \int_{a_i}^{b_i} \mathbf{v}(\mathbf{w}_i(t)) \cdot d\mathbf{n}(t) \\ &= \sum_i \int_{a_i}^{b_i} \mathbf{v}(\mathbf{w}_i(t)) \cdot \begin{pmatrix} \dot{y}(t) \\ -\dot{x}(t) \end{pmatrix} dt. \end{aligned}$$

**9.3.13 Satz (Divergenzsatz in der Ebene).** Seien  $B, \partial B$  wie oben und sei  $\mathbf{v}$  ein  $C^1$ -Vektorfeld auf  $B$ , dann gilt

$$\int_{\partial B} \mathbf{v} \cdot d\mathbf{n} = \iint_B \operatorname{div} \mathbf{v} dF.$$

- Wir stellen uns  $\mathbf{v}$  als Geschwindigkeitsvektorfeld eines Flusses vor. Dann gibt die linke Seite an, wieviel pro Zeit aus dem Gebiet  $B$  insgesamt hinausfließt. Die Divergenz  $\operatorname{div} \mathbf{v}$  gibt an, wie stark sich das Material an einer festen Stelle ausdehnt ( $\operatorname{div} \mathbf{v} > 0$ ) beziehungsweise zusammenzieht ( $\operatorname{div} \mathbf{v} < 0$ ). Somit bedeutet die obige Gleichheit, dass genau soviel „Fläche“ aus  $B$  hinausfließt, wie im Inneren durch Ausdehnung neu entsteht.
- Jede Funktion auf  $\mathbb{R}^2$  kann sowohl als Wirbelstärke wie auch als Divergenz eines jeweils geeigneten Vektorfeldes geschrieben werden. Auf diese Weise lassen sich Doppelintegrale immer durch Kurvenintegrale ersetzen. Genauso haben wir eindimensionale Integrale mit dem Hauptsatz zurückgeführt auf das Auswerten einer Stammfunktion an den Intervallenden. So gesehen sind die beiden obigen Sätze höherdimensionale Verallgemeinerungen des Hauptsatzes 4.1.10 b).
- **Sonderfälle:** Mit  $v_1 = 0, v_2 = x$  bzw.  $v_2 = 0, v_1 = -y$  ergibt sich aus dem Satz 9.3.12 von Green eine Formel für den Flächeninhalt:

$$F = \iint_B dx dy = \int_{\partial B} x dy = - \int_{\partial B} y dx = \frac{1}{2} \int_{\partial B} x dx - \frac{1}{2} \int_{\partial B} y dx .$$

Diese Formel wurde früher in mechanischen Geräten zum Ausmessen des Flächeninhalts benutzt (Planimeter). Für den geometrischen Schwerpunkt gilt:

$$\begin{aligned} \bar{x} &= -\frac{1}{F} \int_{\partial B} xy dx = \frac{1}{2F} \int_{\partial B} x^2 dy, \\ \bar{y} &= \frac{1}{F} \int_{\partial B} xy dy = -\frac{1}{2F} \int_{\partial B} y^2 dx, \end{aligned}$$

wobei Satz 9.3.12 jeweils mit  $\mathbf{v} = \begin{pmatrix} xy \\ 0 \end{pmatrix}$ ,  $\frac{1}{2} \begin{pmatrix} 0 \\ x^2 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ xy \end{pmatrix}$  beziehungsweise  $\frac{1}{2} \begin{pmatrix} y^2 \\ 0 \end{pmatrix}$  benutzt wurde.

**9.3.14 Satz** (Gauß in der Ebene). *Ist  $f \in C^2(D, \mathbb{R})$ , so gilt unter den Voraussetzungen der Sätze 9.3.12, 9.3.13*

$$\iint_B \Delta f dx dy = \int_{\partial B} \partial_{\mathbf{n}} f ds .$$

**Beispiel:** Sei  $B$  begrenzt durch die Strecke  $\mathbf{w}_1(t) = (t, 0)^T$ ,  $0 \leq t \leq 2\pi a$  und die Zykloide  $\mathbf{w}_2(t) = a(2\pi - t + \sin t, 1 - \cos t)^T$ ,  $0 \leq t \leq 2\pi$ . Wir wollen den geometrischen Mittelpunkt berechnen.

$$\begin{aligned} F &= - \int_{\partial B} y \, dx = -a^2 \int_0^{2\pi} (1 - \cos t)(-1 + \cos t) \, dt + \int_0^{2\pi a} 0 \, dt \\ &= a^2 \left[ t - 2 \sin t + \frac{t}{2} + \frac{1}{4} \cos 2t \right]_0^{2\pi} \\ &= a^2 3\pi, \end{aligned}$$

$$\begin{aligned} \int_{\partial B} x^2 \, dy &= a^3 \int_0^{2\pi} (2\pi - t + \sin t)^2 \sin t \, dt + \int_0^{2\pi a} 0 \, dt \\ &= a^3 \int_0^{2\pi} 4\pi^2 \sin t - 4\pi t \sin t - 2t \sin^2 t + \sin^3 t + t^2 \sin t + 4\pi \sin^2 t \, dt \\ &= a^3 \left[ -4\pi^2 \cos t - 4\pi \sin t + 4\pi t \cos t - \frac{t^2}{2} - t \frac{\sin 2t}{2} + \frac{\cos^2 t}{2} \right. \\ &\quad \left. - \cos t + \frac{\cos^2 t}{2} t^2 \cos t + 2 \cos t + 2t \sin t + 2\pi t - \pi \sin 2t \right]_0^{2\pi} \\ &= 6\pi^2 a^3. \end{aligned}$$

Also gilt  $\bar{x} = \frac{1}{2F} \int_{\partial B} x^2 \, dy = \pi a$ . Analog haben wir:

$$\begin{aligned} \int_{\partial B} y^2 \, dx &= a^3 \int_0^{2\pi} (1 - \cos t)^3 (-1) \, dt + \int_0^{2\pi a} 0 \, dt \\ &= -a^3 \left[ \frac{5}{2} t - \frac{11}{3} \sin t + \frac{3}{2} \cos t \sin t - \frac{1}{3} \cos^2 t \sin t \right]_0^{2\pi} = -a^3 5\pi, \end{aligned}$$

und demzufolge  $\bar{y} = -\frac{1}{2F} \int_{\partial B} y^2 \, dx = \frac{5}{6} a$ .

**9.3.15 Definition.** Es seien  $B, D \subset \mathbb{R}^2$  reguläre Gebiete. Eine **Koordinatentransformation** ist eine Abbildung  $(x, y)^T: D \rightarrow B$  mit folgenden Eigenschaften:

- 1)  $x = x(u, v), y = y(u, v)$  sind  $C^1$ -Funktionen auf  $D$ ,
- 2) zu jedem Punkt  $(x, y)^T \in B$  existiert genau ein Punkt  $(u, v) \in D$  mit  $x = x(u, v)$  und  $y = y(u, v)$ ,
- 3)  $|x_u y_v - x_v y_u| \neq 0$  für alle  $(u, v) \in D$ .

Man nennt  $(u, v)$  auch ein **krummliniges Koordinatensystem** auf  $B$ .

- Das Gebietsintegral  $\iint_B f dF$  ist unabhängig von der Zerlegung von  $B$ . Die Zerlegung durch kartesische Koordinatenlinien liefert (siehe (5.1), (3.3))

$$\iint_B f dF = \iint_B f dx dy.$$

andererseits liefert die Zerlegung durch die krummlinigen Koordinatenlinien zur Abbildung  $(x, y): D \rightarrow B$

$$dF = dO = \|\mathbf{x}_u \times \mathbf{x}_v\| du dv$$

wobei  $\mathbf{x}(u, v) = (x(u, v), y(u, v), 0)^T \in \mathbb{R}^3$  ist. Also haben wir

$$\begin{aligned} dF &= |x_u y_v - x_v y_u| du dv = \left| \det \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \right| du dv \\ &=: \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv, \end{aligned}$$

wobei  $\frac{\partial(x, y)}{\partial(u, v)}$  die **Funktionaldeterminante** heißt. Also gilt

$$\iint_B f dF = \iint_D f \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv.$$

Wir haben also bewiesen:

**9.3.16 Satz** (Transformationsformel). *Entsteht ein regulärer Bereich  $B \subseteq \mathbb{R}^2$  unter der Koordinatentransformation  $x = x(u, v)$   $y = y(u, v)$  aus  $D$ , dann gilt für jede Funktion  $f: B \rightarrow \mathbb{R}$ , dass*

$$\iint_B f(x, y) dx dy = \iint_D f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv. \quad (3.6)$$

- Die Transformationsformel ist eine höherdimensionale Verallgemeinerung der Substitutionsregel 4.2.5. Wie auch die Substitutionsregel lässt sich die Transformationsformel sowohl „von links nach rechts“ als auch „von rechts nach links“ anwenden.
- Wegen (3.4) verändern sich die Integrale in (3.6) nicht wenn  $B$  bzw.  $D$  durch  $B \setminus N$  bzw.  $D \setminus M$  ersetzt werden, wobei  $N$  und  $M$  Nullmengen sind. Insbesondere sind also  $C^1$ -Transformationen  $\mathbf{x} : D \rightarrow S$  die nur auf  $D \setminus M$ , mit einer Nullmenge  $M$ , die Bedingungen aus 9.3.15 erfüllen, zulässig.
- Folgende Fälle treten besonders häufig auf:

a) affine Koordinaten:

$$x = au + bv + x_0, y = cu + dv + y_0, \text{ mit } ad - bc \neq 0;$$

$$dF = |ad - bc| du dv$$

$$\iint_B f(x, y) dx dy = \iint_D f(au+bv+x_0, cu+dv+y_0) |ad - bc| du dv .$$

b) Polarkoordinaten

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad dF = r dr d\varphi$$

$$\iint_B f(x, y) dx dy = \iint_D f(r \cos \varphi, r \sin \varphi) r dr d\varphi .$$

### Beispiele:

- a) Durch die lineare Transformation  $x = au, y = bv, a, b > 0$  geht der Kreis  $K : u^2 + v^2 \leq 1$  in die Ellipse  $E : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$  über. Also gilt:

$$F(E) = \iint_E dx dy = \iint_K ab du dv = abF(K) = ab\pi .$$

- b) Für die Gaußverteilung  $F(x) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-t^2} dt$  gilt:  $\lim_{x \rightarrow \infty} F(x) = 1$ , d.h.

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi} .$$

## 9.4 Integration über Flächen im Raum

- Flächen im Raum können als Graph  $z = h(x, y)$  oder implizit durch  $f(x, y, z) = 0$  dargestellt werden.

**9.4.1 Definition.** Sei  $D$  ein regulärer Bereich in einem Gebiet  $G$  der  $(u, v)$ -Ebene. Unter der **Parameterdarstellung** eines **regulären** Flächenstücks verstehen wir die Einschränkung einer  $C^1$ -Funktion  $\mathbf{x} : G \rightarrow \mathbb{R}^3$ ,  $\mathbf{x}(u, v) = (x(u, v), y(u, v), z(u, v))^T$  auf  $D$  mit den Eigenschaften:

- 1)  $(u, v) \neq (u', v')$  in  $D$ ,  $\Rightarrow \mathbf{x}(u, v) = \mathbf{x}(u', v')$ ,
- 2)  $\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v) \neq \mathbf{0}$  für alle  $(u, v) \in D$ .

Die Punktmenge  $S = \{\mathbf{x}(u, v); (u, v) \in D\}$  ist das dargestellte Flächenstück.

- Aus 1) folgt:  $\forall (x, y, z) \in S \quad \exists!(u, v) \in D : (x, y, z)^T = \mathbf{x}(u, v)$ .
- Durch die Abbildungen

$$\begin{array}{ll} u \rightarrow \mathbf{x}(u, v) & v = \text{const.}, \\ v \rightarrow \mathbf{x}(u, v) & u = \text{const.} \end{array}$$

sind **Parameterlinien**  $v = \text{const.}$ , bzw.  $u = \text{const.}$  gegeben.

- Aus 2) folgt, dass die Tangentialvektoren der Parameterlinien in jedem Punkt  $(u, v) \in D$  linear unabhängig sind.
- Die Tangentialvektoren  $\mathbf{x}_u(u, v)$ ,  $\mathbf{x}_v(u, v)$  spannen im Flächenpunkt  $\mathbf{x}(u, v)$  die Tangentialebene auf, die die Parameterdarstellung

$$v(\lambda, \mu) = \mathbf{x}(u, v) + \lambda \mathbf{x}_u(u, v) + \mu \mathbf{x}_v(u, v)$$

hat. Der auf der Tangentialebene senkrecht stehende Vektor

$$\mathbf{n} := \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} \quad (4.1)$$

heißt **Flächeneinheitsnormalenvektor**, kurz **Flächennormale**.

- Ein reguläres Kurvenstück  $t \rightarrow (u(t), v(t))$ ,  $t_0 \leq t \leq t_1$  im Parameterbereich  $D$  induziert durch

$$\mathbf{w}(t) = \mathbf{x}(u(t), v(t))$$

eine reguläre Flächenkurve in  $S$ , deren Tangentialvektor  $\dot{\mathbf{w}} = \mathbf{x}_u \dot{u} + \mathbf{x}_v \dot{v}$  in der Tangentialebene durch  $\mathbf{x}(u, v)$  liegt.

- Die Bogenlänge

$$s(t) = \int_{t_0}^{t_1} \|\dot{\mathbf{w}}(t)\| dt$$

ergibt sich wegen

$$\|\dot{\mathbf{w}}\|^2 = \dot{\mathbf{w}} \cdot \dot{\mathbf{w}} = (\mathbf{x}_u \cdot \mathbf{x}_u) \dot{u}^2 + 2(\mathbf{x}_u \cdot \mathbf{x}_v) \dot{u} \dot{v} + (\mathbf{x}_v \cdot \mathbf{x}_v) \dot{v}^2$$

aus den **metrischen Fundamentalgrößen**

$$E := \mathbf{x}_u \cdot \mathbf{x}_u, \quad F := \mathbf{x}_u \cdot \mathbf{x}_v, \quad G := \mathbf{x}_v \cdot \mathbf{x}_v.$$

Diese bestimmen auch die Winkelmessung (Seien  $\mathbf{w}_1, \mathbf{w}_2$  zwei reguläre Flächenkurven, die sich in einem Punkt  $\mathbf{x}(u, v)$  schneiden. Der Winkel zwischen den beiden Flächenkurven ist der Winkel zwischen den beiden Tangentialvektoren), und die Flächeninhalte (siehe 8.4.9). Aufgrund von  $\|\mathbf{a} \times \mathbf{b}\|^2 = \|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2$  gilt

$$\|\mathbf{x}_u \times \mathbf{x}_v\|^2 = EG - F^2 \tag{4.2}$$

- Wenn sich in jedem Flächenpunkt  $\mathbf{x}(u_0, v_0)$  die Parameterlinien  $u = u_0, v = v_0$  senkrecht schneiden, d.h.  $F = 0$ , dann sagt man, daß die Parameterlinien ein **orthogonales Netz** bilden.
- Der **Rand** des Flächenstücks  $S$  ist das  $\mathbf{x}$ -Bild des Randes von  $D$ , d.h.

$$\partial S := \{\mathbf{x}(u, v); (u, v) \in \partial D\}.$$

- Eine **stückweise reguläre Fläche**  $S$  ist die Vereinigung endlich vieler regulärer Flächenstücke  $S_i$ , von denen je zwei längs einer oder mehrerer gemeinsamer Randstücke aneinandertreffen. Der **Rand**  $\partial S$  von  $S$  wird aus allen Randstücken gebildet die nur zu einem regulären Flächenstück gehören. Man nennt  $S$  geschlossen, wenn  $\partial S = \emptyset$ .

**Beispiele:**

- a) Für  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$  mit  $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$  ist

$$\mathbf{x}(u, v) = \mathbf{x}_0 + u\mathbf{a} + v\mathbf{b}$$

mit  $(u, v) \in D$  ein Ebenenstück. Die Parameterlinien  $u = \text{const.}, v = \text{const.}$  sind Geraden parallel zu  $\mathbf{a}$  bzw.  $\mathbf{b}$ . Wir haben

$$\begin{aligned} \mathbf{x}_u &= \mathbf{a}, & \mathbf{x}_v &= \mathbf{b} \\ E &= \|\mathbf{a}\|^2, & F &= \mathbf{a} \cdot \mathbf{b}, & G &= \|\mathbf{b}\|^2. \end{aligned}$$

- b) Aus der expliziten Darstellung  $z = h(x, y)$ , mit  $h \in C^1(D, \mathbb{R})$  erhält man sofort die Parameterdarstellung

$$\mathbf{x}(u, v) = (u, v, h(u, v))^T, \quad (u, v) \in D.$$

Die Parameterlinien sind Schnitte mit den Ebenen  $u = \text{const.}, v = \text{const.}$  und es gilt:

$$\begin{aligned} \mathbf{x}_u &= (1, 0, h_u)^T, & \mathbf{x}_v &= (0, 1, h_v)^T \\ \mathbf{x}_u \times \mathbf{x}_v &= (-h_u, -h_v, 1)^T \\ \|\mathbf{x}_u \times \mathbf{x}_v\| &= \sqrt{1 + h_u^2 + h_v^2}. \end{aligned}$$

- c) Aus einem regulären Kurvenstück  $t \rightarrow \mathbf{w}(t) = (x(t), 0, z(t))^T, t_0 \leq t \leq t_1$ , ohne Doppelpunkte in der  $x$ - $z$ -Ebene mit  $x(t) > 0$  entsteht (siehe Kapitel 7 7.6.7) durch Drehung um die  $z$ -Achse und den Winkel  $\varphi_0$  ein reguläres Flächenstück mit der Darstellung

$$\begin{aligned} \mathbf{x}(t, \varphi) &= \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x(t) \\ 0 \\ z(t) \end{pmatrix} \\ &= \begin{pmatrix} x(t) \cos \varphi \\ x(t) \sin \varphi \\ z(t) \end{pmatrix} \end{aligned}$$

definiert auf  $D = \{(t, \varphi); t_0 \leq t \leq t_1, 0 \leq \varphi \leq \varphi_0\}$ . Die Parameterlinien  $t = \text{const.}$  sind Breitenkreise,  $\varphi = \text{const.}$  die Meridiane. Es gilt:

$$\mathbf{x}_t = (\dot{x}(t) \cos \varphi, \dot{x}(t) \sin \varphi, \dot{z}(t))^T$$

$$\begin{aligned}\mathbf{x}_\varphi &= (-x(t) \sin \varphi, x(t) \cos \varphi, 0)^T \\ \mathbf{x}_t \times \mathbf{x}_\varphi &= (-x\dot{z} \cos \varphi, -x\dot{z} \sin \varphi, x\dot{x})^T \\ \|\mathbf{x}_t \times \mathbf{x}_\varphi\| &= x \sqrt{\dot{z}^2 \cos^2 \varphi + \dot{z}^2 \sin^2 \varphi + \dot{x}^2} = x \cdot \|\dot{\mathbf{w}}\| .\end{aligned}$$

Falls  $\varphi_0 = 2\pi$  ist das Flächenstück  $S$  nur stückweise regulär, da die Meridiane  $\varphi = 0$  und  $\varphi = 2\pi$  zusammenfallen. Der Rand von  $S$  sind die Breitenkreise

$$\varphi \rightarrow (x(t_i) \cos \varphi, x(t_i) \sin \varphi, z(t_i))^T, \quad i = 0, 1.$$

#### 9.4.2 Sonderfälle.

a) *Kegelstumpf*

$$\begin{aligned}\mathbf{x}(t, \varphi) &= \begin{pmatrix} t \cos \varphi \\ t \sin \varphi \\ b(1 - \frac{t}{a}) \end{pmatrix}, \quad a_0 \leq t \leq a, 0 \leq \varphi \leq 2\pi, \\ \mathbf{x}_t \times \mathbf{x}_\varphi &= \begin{pmatrix} \frac{b}{a}t \cos \varphi, \frac{b}{a}t \sin \varphi, t \end{pmatrix}^T\end{aligned}$$

b) *Zylinderstumpf*

$$\begin{aligned}\mathbf{x}(z, \varphi) &= \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}, \quad z_0 \leq z \leq z_1, 0 \leq \varphi \leq 2\pi, \\ \mathbf{x}_z \times \mathbf{x}_\varphi &= (-r \cos \varphi, -r \sin \varphi, 0)^T\end{aligned}$$

c) *Sphäre*

$$\begin{aligned}\mathbf{x}(\psi, \varphi) &= \begin{pmatrix} r \sin \psi \cos \varphi \\ r \sin \psi \sin \varphi \\ r \cos \psi \end{pmatrix}, \quad 0 \leq \psi \leq \pi, 0 \leq \varphi \leq 2\pi, \\ \mathbf{x}_\psi \times \mathbf{x}_\varphi &= (r^2 \sin^2 \psi \cos \varphi, r^2 \sin^2 \psi \sin \varphi, r^2 \sin \psi \cos \psi)^T .\end{aligned}$$

Für  $\psi = 0, \pi$  (Süd-, bzw. Nordpool) ist  $\mathbf{x}_\psi \times \mathbf{x}_\varphi = \mathbf{0}$ , längs des Halbkreises  $\varphi = 0$  ist die Darstellung nicht eindeutig. Beides sind Nullmengen.

d) *Torus*

$$\mathbf{x}(\psi, \varphi) = \begin{pmatrix} (R + r \sin \psi) \cos \varphi \\ (R + r \sin \psi) \sin \varphi \\ r \cos \psi \end{pmatrix}, \quad 0 \leq \psi \leq 2\pi.$$

*Der Torus ist eine geschlossene Fläche.*

**9.4.3** Der Flächeninhalt. Sei  $\mathbf{x} : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$\mathbf{x}(u, v) = (x(u, v), y(u, v), z(u, v))^T$$

eine Parameterdarstellung eines regulären Flächenstücks  $S$ . Das Flächenelement  $\Delta S$ , berandet von den Parameterlinien  $u = u_0, u = u_0 + \Delta u, v = v_0, v = v_0 + \Delta v$ , wird durch das Parallelogramm, das von  $\mathbf{x}_u \Delta u$  und  $\mathbf{x}_v \Delta v$  aufgespannt wird, approximiert, denn

$$\mathbf{x}(u, v) \approx (\mathbf{x}_u(u_0, v_0), \mathbf{x}_v(u_0, v_0)) \begin{pmatrix} u - u_0 \\ v - v_0 \end{pmatrix} + \mathbf{x}(u_0, v_0)$$

ist die beste lineare Approximation. Also gilt

$$\begin{aligned} \Delta S &\approx \Delta O, \\ \Delta O &= \|\mathbf{x}_u(u_0, v_0) \times \mathbf{x}_v(u_0, v_0)\| \Delta u \Delta v. \end{aligned}$$

Die Fläche  $S$  wird durch Parameterlinien in  $n$  Teilstücke  $S_i, i = 1, \dots, n$ , zerlegt und diese werden durch Parallelogramme  $O_i$  approximiert. Der Grenzwert  $n \rightarrow \infty, \Delta u^2 + \Delta v^2 \rightarrow 0$  liefert den **Flächeninhalt** von  $S$ , d.h.

$$\begin{aligned} O(S) &:= \lim_{\substack{n \rightarrow \infty \\ \Delta u^2 + \Delta v^2 \rightarrow 0}} \sum_{i=1}^n \|\mathbf{x}_u(u_i, v_i) \times \mathbf{x}_v(u_i, v_i)\| \Delta u \Delta v \\ &= \iint_D \|\mathbf{x}_u \times \mathbf{x}_v\| \, du \, dv. \end{aligned} \tag{4.3}$$

Mit den Sprechweisen aus Kapitel 4 haben wir

$$O(S) = \iint_D dO$$

mit

$$dO = \|\mathbf{x}_u \times \mathbf{x}_v\| \, du \, dv.$$

- Im Falle eines Graphen  $z = h(x, y)$  haben wir

$$O = \iint_D \sqrt{1 + h_x^2 + h_y^2} dx dy$$

und im Falle einer Drehfläche ( $\varphi_0 = 2\pi$ )

$$\begin{aligned} O &= \iint_D \sqrt{x^2(\dot{x}^2 + \dot{z}^2)} dt d\varphi \\ &= 2\pi \int_{t_0}^{t_1} x(t) \|\dot{\mathbf{w}}(t)\| dt. \end{aligned}$$

**Beispiele:**

- a) Kugeloberfläche. Nach 9.4.2 c) haben wir  $E = r^2, F = 0, G = r^2 \sin^2 \psi$  und  $D = \{(\psi, \varphi); \psi \in [0, \pi], \varphi \in [0, 2\pi]\}$ . Also

$$\begin{aligned} O &= \iint_D r^2 \sin \psi d\psi d\varphi \\ &= r^2 2\pi \int_0^\pi \sin \psi d\psi = 4\pi r^2. \end{aligned}$$

- b) Torusoberfläche. Nach 9.4.2 d) gilt:  $G = (R + r \sin \psi)^2, E = r^2, F = 0$  mit  $D = [0, 2\pi] \times [0, 2\pi]$ . Also ist

$$\begin{aligned} O &= \int_0^{2\pi} \int_0^{2\pi} r(R + r \sin \psi) d\psi d\varphi \\ &= 2\pi r \int_0^{2\pi} R + r \sin \psi d\psi = 4\pi^2 r R. \end{aligned}$$

**9.4.4 Definition.** Ist  $\mathbf{x} : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^3$  eine Parameterdarstellung eines regulären Flächenstücks  $S$  und  $f$  ein auf  $S$  stetiges Skalarfeld, dann nennt man

$$\iint_S f dO := \iint_S f(\mathbf{x}(u, v)) \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| du dv$$

das **Oberflächenintegral** von  $f$  über  $S$ . Für eine aus  $S_1, \dots, S_n$  bestehende stückweise reguläre Fläche  $S$  setzen wir

$$\iint_S f \, dO := \sum_{i=1}^n \iint_{S_i} f \, dO.$$

- Auf Grund der Definition gelten die üblichen Rechenregeln: Linearität, Monotonie und Additivität. Außerdem gilt der Mittelwertsatz, d.h. es existiert ein  $(u^*, v^*)$  mit

$$f(\mathbf{x}(u^*, v^*)) = \frac{1}{O(S)} \iint_S f \, dO. \quad (4.4)$$

**9.4.5** Berechnung des Oberflächenintegrals. Wir gehen wie folgt vor:

- 1) Parameterdarstellung  $\mathbf{x} : D \rightarrow \mathbb{R}^3$  der Fläche angeben.
- 2) Die partiellen Ableitungen  $\mathbf{x}_u, \mathbf{x}_v$  und das Oberflächenelement

$$dO = \|\mathbf{x}_u \times \mathbf{x}_v\| \, du \, dv = \sqrt{EG - F^2} \, du \, dv$$

berechnen.

- 3) Einsetzen

$$\iint_S f \, dO = \iint_A f(\mathbf{x}(u, v)) \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| \, du \, dv$$

und ausrechnen.

**Beispiele:**

- a) Die  $x, y$ -Koordinaten des geometrischen Schwerpunkts  $(\bar{x}, \bar{y}, \bar{z})$  des Teilstücks der Fläche  $3z = 2(x^{\frac{3}{2}} + y^{\frac{3}{2}})$ , der von den Ebenen  $x = 0$ ,  $y = 0$ ,  $x + y = 1$  ausgeschnitten wird berechnet sich wie folgt: Flächenstück ist bezüglich  $x$  und  $y$  symmetrisch, also ist  $\bar{x} = \bar{y}$ .

$$O(S) = \iint_S 1 \, dO = \int_0^1 \int_0^{1-x} \sqrt{1+x+y} \, dy \, dx$$

$$\begin{aligned}
&= \int_0^1 \frac{2}{3} \sqrt{(1+x+y)^3} \Big|_{y=0}^{y=1-x} dx \\
&= \frac{2}{3} \int_0^1 (2)^{\frac{3}{2}} - (1+x)^{\frac{3}{2}} dx \\
&= \frac{2}{3} 2^{\frac{3}{2}} x \Big|_0^1 - \frac{2}{3} \frac{2}{5} (1+x)^{\frac{5}{2}} \Big|_0^1 = \frac{4(\sqrt{2}+1)}{15}
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \int_0^{1-x} x \sqrt{1+x+y} dy dx &= \frac{2}{3} \int_0^1 2^{\frac{3}{2}} x - (1-x)^{\frac{3}{2}} x dx \\
&= \frac{2}{3} 2^{\frac{1}{2}} x^2 + \frac{8}{105} (1+x)^{\frac{7}{2}} - \frac{4}{15} (1+x)^{\frac{5}{2}} x \Big|_0^1 \\
&= \frac{3}{2} 2^{\frac{1}{2}} + \frac{8}{105} 2^{\frac{7}{2}} - \frac{4}{15} 2^{\frac{5}{2}} - \frac{8}{105} \\
&= \frac{22\sqrt{2} - 8}{105},
\end{aligned}$$

$$\bar{x} = \frac{(22\sqrt{2} - 8)15}{105 \cdot 4(\sqrt{2} + 1)} = \frac{(11\sqrt{2} - 4)(\sqrt{2} - 1)}{14} = \frac{26 - 15\sqrt{2}}{14}.$$

- b) Das elektrostatische Potential  $U(\mathbf{a})$  einer homogenen mit Dichte  $\rho$  geladenen Fläche ist

$$\iint_S \frac{\rho}{\|\mathbf{x} - \mathbf{a}\|} dO.$$

Falls  $S$  der Kegelmantel  $x^2 + y^2 = z^2, 0 \leq z \leq 1$  und  $\mathbf{a} = (0, 0, 1)^T$  ist, ergibt sich nach 9.4.2 a) die Parametrisierung  $\mathbf{x}(t, \varphi) = (t \cos \varphi, t \sin \varphi, t)^T, 0 \leq t \leq 1, 0 \leq \varphi \leq 2\pi, \mathbf{x}_t \times \mathbf{x}_\varphi = -t(\cos \varphi, \sin \varphi, -1)^T$  und das Oberflächenelement  $dO = t\sqrt{2} d\varphi dt$ . Also haben wir

$$U(\mathbf{a}) = \int_0^1 \int_0^{2\pi} \frac{\rho\sqrt{2}t}{\sqrt{t^2 + (t-1)^2}} d\varphi dt = 2\pi\sqrt{2}\rho \int_0^1 \frac{t}{\sqrt{2t^2 - 2t + 1}} dt$$

$$\begin{aligned}
&= \pi\sqrt{2}\varrho \int_0^1 \frac{4t-2}{2\sqrt{2t^2-2t+1}} + \frac{1}{\sqrt{2t^2-2t+1}} dt \\
&= \pi\sqrt{2}\varrho \sqrt{2t^2-2t+1} \Big|_0^1 + 2\pi\varrho \int_0^1 \frac{1}{\sqrt{(2t-1)^2+1}} dt \\
&= \sqrt{2}\varrho\pi\sqrt{2t^2-2t+1} + \varrho\pi \ln(2t-1 + \sqrt{(2t-1)^2+1}) \Big|_0^1 \\
&= \varrho\pi \left( \ln(1+\sqrt{2}) - \ln(\sqrt{2}-1) \right) \\
&= \varrho\pi \ln(3+2\sqrt{2}).
\end{aligned}$$

- Das Oberflächenintegral tritt bei Berechnungen von Flächeninhalten von Flächen im Raum und bei Berechnungen der Masse, der Ladung und des Massenmittelpunktes auf.
- Die Berechnung des Oberflächenintegrals ist oft schwierig, kann aber durch eine dem Problem und der Geometrie der Fläche angepaßte Koordinatenwahl vereinfacht werden.

**9.4.6 Definition.** Ist  $S$  ein durch die Parameterdarstellung  $\mathbf{x} : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^3$  gegebenes reguläres Flächenstück mit der Flächennormalen  $\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|}$  und  $\mathbf{v}$  ein auf  $S$  stetiges Vektorfeld, dann nennt man das Oberflächenintegral der skalaren Normalenkomponente  $\mathbf{v} \cdot \mathbf{n}$  den **Fluß** von  $\mathbf{v}$  durch  $S$ , d.h.

$$\iint_S \mathbf{v} \cdot d\mathbf{O} := \iint_S \mathbf{v} \cdot \mathbf{n} dO = \iint_S \det(\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v) du dv.$$

Man nennt

$$\begin{aligned}
d\mathbf{O} &:= \mathbf{n} dO = \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|} \|\mathbf{x}_u \times \mathbf{x}_v\| du dv \\
&= \mathbf{x}_u \times \mathbf{x}_v du dv
\end{aligned}$$

das **vektorielle Oberflächenelement** von  $S$ .

- **Integration:** Für das Geschwindigkeitsfeld  $\mathbf{v}$  einer stationären Strömung ist die Determinante

$$\det(\mathbf{v}, \mathbf{x}_u du, \mathbf{x}_v dv) = \det(\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v) du dv$$

das Volumen der Flüssigkeitsmenge die pro Zeiteinheit durch das Oberflächenelement  $dO$  fließt, denn

$$\begin{aligned} \det(\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v) du dv &= \mathbf{v} \cdot (\mathbf{x}_u \times \mathbf{x}_v) du dv \\ &= \mathbf{v} \cdot \mathbf{n} dO. \end{aligned}$$

- Fall  $S$  durch einen Graphen  $z = h(x, y)$  gegeben ist, haben wir für den Fluß eines Vektorfeldes  $\mathbf{v}$

$$\begin{aligned} \iint_S \mathbf{v} \cdot d\mathbf{O} &= \iint_D \det \begin{pmatrix} v_1 & 1 & 0 \\ v_2 & 0 & 1 \\ v_3 & h_x & h_y \end{pmatrix} dx dy \\ &= \iint_D (-v_1 h_x - v_2 h_y + v_3) dx dy, \end{aligned}$$

wobei  $v_i = v_i(x, y, h(x, y))$ .

**Beispiel:** Der Fluß des elektrischen Feldes  $D = \frac{cq}{\|\mathbf{x}\|^3} \mathbf{x}$  einer Punktladung  $q$  im Ursprung durch die Sphäre  $S : \|\mathbf{x}\| = R$  ist wegen

$$\begin{aligned} \mathbf{E} \cdot \mathbf{n} &= \frac{cq}{\|\mathbf{x}\|^3} \mathbf{x} \cdot \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{cq}{R^2} \\ \iint_S \mathbf{E} \cdot \mathbf{n} dO &= \iint_S \frac{cq}{R^2} dO = \frac{cq}{R^2} \iint_S dO = 4\pi cq. \end{aligned}$$

- Der Satz von Green hat eine Verallgemeinerung für **zweiseitige Flächen** im Raum, d.h. Flächen für die man eindeutig von einer Ober- bzw. Unterseite sprechen kann. Jedes reguläre Flächenstück  $\mathbf{x} : D \subseteq \mathbb{R}^2 \rightarrow S \subseteq \mathbb{R}^3$  ist zweiseitig. Die Oberseite ist die Seite in die die Flächennormale  $\mathbf{n}$  zeigt. Eine stückweise reguläre Fläche ist **zweiseitig**, wenn man die Oberseiten der Flächenstücke  $S_k$  so wählen kann, daß sich der Umlaufsinn über die Kanten  $S_k \cap S_j$  stetig fortsetzt.

**Gegenbeispiel:** Das Möbiusband, hat nur eine Seite.

**9.4.7 Satz (Stokes).** Sei  $S$  eine zweiseitige, stückweise reguläre Fläche mit überschneidungsfreier und geschlossener Randkurve  $\partial S$ , die so durchlaufen wird, das  $S$  links liegt und der Umlaufsinn zusammen mit der Normalenrichtung  $\mathbf{n}$  von  $S$  ein Rechtssystem ergibt. Sei  $U \subseteq \mathbb{R}^3$  eine offene Menge, die  $S$  enthält. Dann gilt für alle  $C^1$ -Vektorfelder  $\mathbf{v} : U \rightarrow \mathbb{R}^3$

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \iint_S (\text{rot } \mathbf{v} \cdot \mathbf{n}) dO. \quad (4.5)$$

- Falls  $S$  ein ebener Bereich der  $(x, y)$ -Ebene ist und  $\mathbf{v} = (v_1, v_2, 0)^T$  ein ebenes Vektorfeld ist, gilt:

$$\operatorname{rot} \mathbf{v} \cdot \mathbf{n} = \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y},$$

mit  $\mathbf{n} = (0, 0, 1)^T$ . Somit geht (4.5) in die Formel in (9.3.12) über.

**Beispiel:** Sei  $\mathbf{v} = (x, y, z) = (-y^3, x^3, -z^3)$  das Geschwindigkeitsfeld einer Strömung im Zylinder  $x^2 + y^2 \leq 1$ . Wir wollen die Zirkulation von  $\mathbf{v}$  längs des Schnittes des Zylindermantels mit der Ebene  $x + y + z = 1$  berechnen.

a) Methode aus 9.2.7

1) Parametrisierung

$$\varphi \rightarrow (\cos \varphi, \sin \varphi, 1 - \sin \varphi - \cos \varphi)^T, \quad 0 \leq \varphi \leq 2\pi.$$

2) Bogenelement

$$d\mathbf{x} = (-\sin \varphi, \cos \varphi, \sin \varphi - \cos \varphi)^T d\varphi.$$

3) Einsetzen

$$\int_0^{2\pi} (\sin \varphi)^3 \sin \varphi + (\cos \varphi)^3 \cos \varphi - (1 - \sin \varphi - \cos \varphi)^3 (\sin \varphi - \cos \varphi) d\varphi$$

Ausrechnen = viel Arbeit.

b) Satz 9.4.7

1) Parametrisierung

$$\mathbf{x}(u, v) = (u, v, 1 - u - v)^T \quad (u, v) \in B_1(0),$$

2) Berechne

$$\begin{aligned} \operatorname{rot} \mathbf{v} &= (0, 0, 3(x^2 + y^2))^T, \\ \mathbf{x}_u \times \mathbf{x}_v &= (1, 1, 1)^T, \end{aligned}$$

3) Einsetzen

$$\begin{aligned} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} &= \iint_S \operatorname{rot} \mathbf{v} \cdot (\mathbf{x}_u \times \mathbf{x}_v) \, du \, dv \\ &= 3 \iint_K (u^2 + v^2) \, du \, dv = 3 \int_0^{2\pi} \int_0^1 r^2 r \, dr \, d\varphi = \frac{3}{2}\pi. \end{aligned}$$

- In (4.5) kann  $S$  durch jede andere Fläche ersetzt werden die denselben Rand  $\partial S$  hat, d.h.

$$\iint_{S_1} \operatorname{rot} \mathbf{v} \cdot d\mathbf{O} = \iint_{S_2} \operatorname{rot} \mathbf{v} \cdot d\mathbf{O}$$

gilt für beliebige stückweise reguläre Flächen  $S_1, S_2$  mit Rand  $\partial S$ .

**9.4.8** Interpretation der Rotation. Sei  $S$  ein reguläres Flächenstück im Definitionsbereich von  $\mathbf{v}$  und  $\mathbf{x}$  ein Punkt in  $S$ . Sei  $S_r$  der Durchschnitt von  $S$  mit dem Ball  $B_r(\mathbf{x})$ . Dann existiert  $\tilde{\mathbf{x}}_r \in S_r$  mit

$$\begin{aligned} \oint_{\partial S_r} \mathbf{v} \cdot d\mathbf{x} &= \iint_{S_r} \operatorname{rot} \mathbf{v} \cdot \mathbf{n} \, dO \\ &= \operatorname{rot} \mathbf{v}(\tilde{\mathbf{x}}_r) \cdot \mathbf{n}(\tilde{\mathbf{x}}_r). \end{aligned}$$

Also gilt:

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \lim_{r \rightarrow 0} \frac{1}{L(\partial S_r)} \oint_{\partial S_r} \mathbf{v} \cdot d\mathbf{x}.$$

Rechts steht die „Zirkulation pro Flächeneinheit“, diese wird **Wirbelstärke** genannt. Sie ist am größten wenn die Rotation des Vektorfeldes  $\mathbf{v}$  in dieselbe Richtung wie die Normale  $\mathbf{n}$  zeigt.

## 9.5 Integration über dreidimensionale Bereiche

Die Methoden aus Abschnitt 9.3 werden auf 3-dimensionale Bereiche übertragen.

**9.5.1** Das Volumen. Der  $(x, y, z)$ -Raum wird durch achsenparallele Ebenen

$$x = n2^{-k}, \quad y = n2^{-k}, \quad z = n2^{-k}, \quad k = 0, \pm 1, \pm 2, \dots$$

in Würfel mit dem Volumen  $2^{-3k}$  zerlegt. Für eine beschränkte Menge  $M \subseteq \mathbb{R}^3$  bezeichne  $s_k(M)$  bzw.  $S_k(M)$  das Gesamtvolumen aller ganz in  $M$  enthaltenen Würfel bzw. aller Würfel die mindestens einen Punkt aus  $M$  enthalten. Man nennt  $M$  **Riemann-meßbar** und  $V(M)$  das **Volumen** von  $M$ , wenn

$$V(M) := \lim_{k \rightarrow \infty} s_k(M) = \lim_{k \rightarrow \infty} S_k(M).$$

Besteht eine Menge  $N \subseteq \mathbb{R}^3$  aus nur endlich vielen Punkten, regulären Kurvenstücken und regulären Flächenstücken, so bezeichnet man sie als **Nullmenge** und es gilt:

$$V(M) = V(M \cup N) = V(M \setminus N),$$

wobei  $M$  eine Riemann-meßbare Menge ist.

**9.5.2 Definition.** Ein Bereich  $B \subseteq \mathbb{R}^3$  heißt **regulär**, wenn:

- 1) Der Rand  $\partial B$  aus endlich vielen stückweise regulären Flächen besteht,
- 2) das Innere  $B \setminus \partial B$  ein nicht leeres Gebiet des  $\mathbb{R}^3$  ist,
- 3)  $B$  abgeschlossen und beschränkt ist, d.h.  $\partial B \subseteq B$ .

- Man kann zeigen, daß jeder reguläre Bereich  $B \subseteq \mathbb{R}^3$  Riemann-meßbar ist und ein Volumen  $V(B) \neq 0$  besitzt.

**9.5.3 Das Volumenintegral.** Sei  $B \subseteq \mathbb{R}^3$  ein Riemann-messbarer Bereich und  $f: B \rightarrow \mathbb{R}$  stetig. Wir zerlegen  $\mathbb{R}^3$  wie oben in achsenparallele Würfel mit Flächeninhalt  $2^{-3k}$ . Es sei  $w_k(M)$  die Menge dieser Würfel, die ganz in  $B$  liegen. Für jeden Würfel  $W_i = [x_i, x_i + 2^{-k}] \times [y_i, y_i + 2^{-k}] \times [z_i, z_i + 2^{-k}] \in w_k(M)$  mit  $x_i, y_i, z_i \in 2^{-k}\mathbb{Z}$  sei  $(x_i^*, y_i^*, z_i^*) \in W_i$ . Dann bilden wir die **Riemann-Summe**

$$\sum_i f(x_i^*, y_i^*, z_i^*) 2^{-3k}. \quad (5.1)$$

Für  $k \rightarrow \infty$  konvergieren diese Riemann-Summen gegen einen Grenzwert unabhängig von der Wahl der Punkte  $(x_i^*, y_i^*, z_i^*)$ , der **Volumenintegral** heißt und mit

$$\iiint_B f \, dV \quad (5.2)$$

bezeichnet wird. Das Symbol  $dV$  heißt **Volumenelement**.

- Es gelten die üblichen Rechenregeln: Linearität, Monotonie, Additivität und Mittelwertsatz.
- Man bezeichnet das Volumenintegral auch durch

$$\iiint_B f \, dV = \iiint_B f(\mathbf{x}) \, dV(\mathbf{x}) = \iiint_B f(x, y, z) \, dx \, dy \, dz. \quad (5.3)$$

- Läßt sich  $B$  als Abschnitt eines geraden Zylinders darstellen mit einem regulären Bereich  $D$  der  $(x, y)$ -Ebene als Querschnitt, unten und oben beschränkt von den Graphen  $z = g(x, y)$  bzw.  $z = h(x, y)$ , d.h.

$$B = \{(x, y, z); (x, y) \in D, g(x, y) \leq z \leq h(x, y)\},$$

dann kann man zeigen:

$$\iiint_B f \, dx \, dy \, dz = \iint_D \left( \int_{g(x,y)}^{h(x,y)} f(x, y, z) \, dz \right) dx \, dy. \quad (5.4)$$

Analoge Formeln gelten auch für Zylinderabschnitte in den anderen Richtungen.

- Ist der oben beschriebene Querschnitt  $D$  ein Normalbereich, z.B.

$$B = \{(x, y, z); a \leq x \leq b, u(x) \leq y \leq v(x), g(x, y) \leq z \leq h(x, y)\},$$

dann gilt wegen (5.4) und Satz 9.3.8

$$\iiint_B f \, dx \, dy \, dz = \int_a^b \left( \int_{u(x)}^{v(x)} \left( \int_{g(x,y)}^{h(x,y)} f(x, y, z) \, dz \right) dy \right) dx.$$

- Im Falle eines Quaders gilt der Satz von Fubini, d.h.

$$\begin{aligned} \iiint_B f \, dx \, dy \, dz &= \int_{x_0}^{x_1} \left( \int_{y_0}^{y_1} \left( \int_{z_0}^{z_1} f(x, y, z) \, dz \right) dy \right) dx \\ &= \int_{z_0}^{z_1} \left( \int_{x_0}^{x_1} \left( \int_{y_0}^{y_1} f(x, y, z) \, dy \right) dx \right) dz = \dots \end{aligned}$$

**9.5.4** Einfache Anwendungen. Das Volumen eines regulären Bereiches berechnet sich als

$$V(B) = \iiint_B dV. \quad (5.5)$$

Insbesondere gilt im oben beschriebenen Falle eines Abschnitts eines geraden Zylinders die Formel

$$V(B) = \iint_D (h(x, y) - g(x, y)) dx dy. \quad (5.6)$$

Die Masse eines Körpers mit Massendichte  $\varrho$  berechnet sich als

$$M = \iiint_B \varrho dV. \quad (5.7)$$

Die Momente sind wie folgt definiert:

$$\begin{aligned} M_{y,z} &:= \iiint_B x \varrho(x, y, z) dV, \\ M_{x,y} &:= \iiint_B z \varrho(x, y, z) dV, \\ M_{x,z} &:= \iiint_B y \varrho(x, y, z) dV. \end{aligned} \quad (5.8)$$

Der Schwerpunkt  $\mathbf{S} = (x_{\mathbf{S}}, y_{\mathbf{S}}, z_{\mathbf{S}})$  eines Bereiches  $B$  hat die Koordinaten

$$x_{\mathbf{S}} = \frac{1}{M} M_{y,z}, \quad y_{\mathbf{S}} = \frac{1}{M} M_{x,z}, \quad z_{\mathbf{S}} = \frac{1}{M} M_{x,y}. \quad (5.9)$$

Wenn wir das Integral von vektorwertigen Funktionen über  $B$  komponentenweise definieren, können wir auch kurz schreiben

$$\mathbf{S} = \iiint_B \varrho(\mathbf{x}) \mathbf{x} dV(\mathbf{x}) \in \mathbb{R}^3. \quad (5.10)$$

Mit  $\varrho = 1$  erhalten wir den geometrischen Mittelpunkt.

**9.5.5** Impuls, Trägheitstensor und Hauptachsen. Es sei  $B \in \mathbb{R}^3$  ein regulärer Bereich und  $\varrho: B \rightarrow \mathbb{R}$  eine Dichtefunktion. Wir fassen  $B$  als starren Körper auf und fragen, ob sich  $B$  um eine vorgegebene Achse durch den Nullpunkt in Richtung  $\mathbf{v} \in \mathbb{R}^3$  drehen kann, ohne dass ihn externe Kräfte auf der Bahn halten.

Dazu berechnen wir zunächst den Impuls. Ein Massepunkt  $\mathbf{x} \in B$  bewegt sich mit Geschwindigkeit  $\mathbf{v} \times \mathbf{x}$ , der Beitrag zum Impuls ist also  $\varrho(\mathbf{x}) \mathbf{v} \times \mathbf{x}$ . Somit erhalten wir den Gesamtimpuls

$$\mathbf{p} = \iiint_B \varrho(\mathbf{x}) \mathbf{v} \times \mathbf{x} dV(\mathbf{x}) = M \mathbf{v} \times \mathbf{S} \in \mathbb{R}^3$$

mit Formel (5.10). Wenn dieser Vektor nicht in Richtung der Drehachse  $\mathbf{v}$  zeigt, würde er bei der Drehung um  $\mathbf{v}$  verändert, was der Impulserhaltung widerspricht. Also muss der Schwerpunkt  $\mathbf{S}$  auf der Drehachse liegen, und es folgt  $\mathbf{p} = \mathbf{0}$ .

Da sich  $\mathbf{x} \in B$  in Richtung  $\mathbf{v} \times \mathbf{x}$  bewegt, trägt er mit  $\varrho(\mathbf{x}) \mathbf{x} \times (\mathbf{v} \times \mathbf{x})$  zum Drehimpuls bei. Der Gesamtdrehimpuls berechnet sich also als

$$\mathbf{L} = \iiint_B \varrho(\mathbf{x}) \mathbf{x} \times (\mathbf{v} \times \mathbf{x}) dV(\mathbf{x}) \in \mathbb{R}^3 .$$

Läge  $\mathbf{L}$  nicht auf der Drehachse  $\mathbf{v}$ , so würde er bei der Drehung um  $\mathbf{v}$  verändert im Widerspruch zur Drehimpulserhaltung. Wir müssen also untersuchen, unter welchen Umständen  $\mathbf{L}$  ein Vielfaches von  $\mathbf{v}$  ist.

Wir erkennen, dass  $\mathbf{L}(\mathbf{v})$  linear in  $\mathbf{v}$  ist. Nach Satz 7.4.11 gibt es eine Matrix  $\mathbf{I} \in \mathbb{R}^{3 \times 3}$  mit

$$\mathbf{L} = \mathbf{I} \cdot \mathbf{v} ,$$

diese Matrix heißt der **Trägheitstensor** des Körpers  $B$ . Falls  $\|\mathbf{v}\| = 1$ , heißt die Zahl

$$T = \mathbf{v}^T \cdot \mathbf{I} \cdot \mathbf{v} = \mathbf{v}^T \cdot \mathbf{L} \geq 0$$

das **axiale Trägheitsmoment** von  $B$  längs der Achse  $\mathbf{v}$ . Der Trägheitstensor ist symmetrisch, denn

$$\begin{aligned} \mathbf{w}^T \cdot \mathbf{I} \cdot \mathbf{v} &= \iiint_B \varrho(\mathbf{x}) \mathbf{w}^T \cdot (\mathbf{x} \times (\mathbf{v} \times \mathbf{x})) dV(\mathbf{x}) \\ &= \iiint_B \varrho(\mathbf{x}) (\mathbf{w} \times \mathbf{x})^T \cdot (\mathbf{v} \times \mathbf{x}) dV(\mathbf{x}) \\ &= \iiint_B \varrho(\mathbf{x}) (\mathbf{v} \times \mathbf{x})^T \cdot (\mathbf{w} \times \mathbf{x}) dV(\mathbf{x}) \\ &= \mathbf{v}^T \cdot \mathbf{I} \cdot \mathbf{w} . \end{aligned}$$

Für  $\mathbf{v} = \mathbf{w} \neq \mathbf{0}$  ist der Integrand positiv, also ist  $\mathbf{I}$  positiv definit, siehe 7.7.6.

Aus Satz 7.7.5 folgt, dass  $\mathbf{I}$  diagonalisierbar ist. Die Eigenvektoren von  $\mathbf{I}$  heißen die **Hauptachsen** des Körpers  $B$ . Dreht man  $B$  längs einer Hauptachse  $\mathbf{v} = \mathbf{b}_i$  von  $B$ , so zeigen Drehachse  $\mathbf{v}$  und Drehmoment  $\mathbf{L} = \mathbf{I} \cdot \mathbf{v}$  in die gleiche Richtung, und  $B$  kann sich ohne Einwirkung externer Kräfte entlang der Achse  $\mathbf{v}$  weiterdrehen. Bei einer Drehung um eine andere Achse würde sich der Drehimpuls ändern, somit kann sich  $B$  um eine solche Achse nur unter Krafteinwirkung drehen. Auf diese Weise erklärt sich sowohl der Name „Hauptachsen“ für die Eigenvektoren des Trägheitstensors, als auch der Name „Hauptachsentransformation“ für Satz 7.7.5.

### Beispiele:

- a) Das axiale Trägheitsmoment des Würfels  $-1 \leq x, y, z \leq 1$  um die  $z$ -Achse ist:

$$\begin{aligned} \int_{-1}^1 \int_{-1}^1 x^2 + y^2 \, dx \, dy \, dz &= 2 \int_{-1}^1 \left. \frac{1}{3}x^3 + y^2x \right|_{x=-1}^{x=1} dy \\ &= 2 \int_{-1}^1 \frac{2}{3} + 2y^2 \, dy = \frac{16}{3} \end{aligned}$$

- b) Für den Quader  $[-a, a] \times [-b, b] \times [-c, c]$  der konstanten Dichte  $\rho$  berechnet man analog

$$\begin{aligned} \mathbf{I} &= \frac{8}{3} abc \rho \begin{pmatrix} b^2 + c^2 & & \\ & a^2 + c^2 & \\ & & a^2 + b^2 \end{pmatrix} \\ &= \frac{M}{3} \begin{pmatrix} b^2 + c^2 & & \\ & a^2 + c^2 & \\ & & a^2 + b^2 \end{pmatrix}. \end{aligned}$$

- c)  $B$  sei vom Ellipsoid  $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1$ , und dem Kegel  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$  berandet und liege im Halbraum  $z \geq 0$ . Man kann  $B$  darstellen als

$$\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} \leq \frac{z}{c} \leq \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$$

mit  $(x, y)$  aus  $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq \frac{1}{2}$ . (Schnittfläche von Ellipsoid und Kegel).  
Es gilt:

$$V(B) = \iint_{\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq \frac{1}{2}} \left( \int_{c\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}}}^{c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}} dz \right) dy dx .$$

Die Koordinatentransformation (Kombination von Polarkoordinaten und affinen Koordinaten)  $x = ra \cos \varphi, y = ra \sin \varphi, 0 \leq r \leq \frac{1}{\sqrt{2}}, 0 \leq \varphi \leq 2\pi$  liefert

$$\begin{aligned} V(B) &= \int_0^{2\pi} \int_0^{\frac{1}{\sqrt{2}}} \int_c^{c\sqrt{1-r^2}} dz abr dr d\varphi \\ &= 2\pi abc \int_0^{\frac{1}{\sqrt{2}}} r \left( \sqrt{1-r^2} \right) - r dr \\ &= 2\pi abc \left( -\frac{1}{3}(1-r^2)^{\frac{1}{3}} - \frac{r^3}{3} \right) \Big|_0^{\frac{1}{\sqrt{2}}} \\ &= \frac{\pi abc}{3} (2 - \sqrt{2}) . \end{aligned}$$

**9.5.6** Transformation von Volumenintegralen. Sei  $B \subseteq \mathbb{R}^3$  ein regulärer Bereich. Unter einer **Koordinatentransformation** auf  $B$  verstehen wir eine  $C^1$ -Abbildung  $(u, v, w)^T \rightarrow \mathbf{x}(u, v, w) = (x(u, v, w), y(u, v, w), z(u, v, w))^T$  eines regulären Bereiches  $U$  des  $(u, v, w)$  Raumes auf  $B$  mit den Eigenschaften:

- 1) Zu jedem Punkt  $(x, y, z)^T \in B$  existiert genau ein Punkt  $(u, v, w) \in U$  mit  $\mathbf{x}(u, v, w) = (x, y, z)^T$ .
- 2) Für alle  $(u, v, w) \in U$  sind die Vektoren  $\mathbf{x}_u(u, v, w), \mathbf{x}_v(u, v, w)$  und  $\mathbf{x}_w(u, v, w)$  linear unabhängig.

Aus  $(u, v, w) \neq (u', v', w')$  folgt  $\mathbf{x}(u, v, w) \neq \mathbf{x}(u', v', w')$ . Bei festem  $u = u_0$  ist  $(v, w) \rightarrow x(u_0, v, w)$  eine **Koordinatenfläche**; analog für die Koordinatenflächen  $v = v_0$ , bzw.  $w = w_0$ . Das Bild des Quaders  $u_0 \leq u \leq u_0 + \Delta u, v_0 \leq$

$v \leq v_0 + \Delta v, w_0 \leq w \leq w_0 + \Delta w$  unter der Koordinatentransformation  $\mathbf{x}$  wird durch einen Spat, der von den Vektoren  $\mathbf{x}_u(u_0, v_0, w_0)\Delta u, \mathbf{x}_v(u_0, v_0, w_0)\Delta v$  und  $\mathbf{x}_w(u_0, v_0, w_0)\Delta w$  aufgespannt wird, approximiert. Das Volumen des Spats ist

$$V = |\det(\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w)| \Delta u \Delta v \Delta w$$

mit der **Funktionaldeterminante**

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} := \det(\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w). \quad (5.11)$$

Durch Grenzübergang  $\Delta u^2 + \Delta v^2 + \Delta w^2$  zeigt man wie im zweidimensionalen Fall:

**9.5.7 Satz.** *Entsteht ein regulärer Bereich  $B \subseteq \mathbb{R}^3$  unter der Koordinatentransformation*

$$x = x(u, v, w), \quad y = y(u, v, w), \quad z = z(u, v, w)$$

*aus dem regulären Bereich  $U \subseteq \mathbb{R}^3$ , dann gilt für jedes auf  $B$  stetige Skalarfeld  $f$  die Transformationsformel*

$$\begin{aligned} & \iiint_B f(x, y, z) \, dx \, dy \, dz \\ &= \iiint_U f(x(u, v, w), y(u, v, w), z(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| \, du \, dv \, dw. \end{aligned} \quad (5.12)$$

*Die Formel (5.12) gilt auch für Transformationen, welche die obigen Bedingungen 1), 2) in 9.5.6 nur auf Bereichen  $U \setminus M$  erfüllen, wobei  $M$  eine Nullmenge ist.*

**9.5.8 Spezialfälle.**

a) *affine Koordinaten*

$$\mathbf{x} = \mathbf{A}\mathbf{u} + \mathbf{x}_0, \quad \mathbf{A} \in \mathbb{R}^{3 \times 3}, \quad \det \mathbf{A} \neq 0$$

$$dV = |\det \mathbf{A}| \, du \, dv \, dw$$

$$\iiint_B f(\mathbf{x}) \, dx \, dy \, dz = \iiint_U f(\mathbf{A}\mathbf{u} + \mathbf{x}_0) |\det \mathbf{A}| \, du \, dv \, dw$$

b) Zylinderkoordinaten

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

$$dV = r \, dr \, d\varphi \, dz$$

$$\iiint_B f(\mathbf{x}) \, dx \, dy \, dz = \iiint_U f(r \cos \varphi, r \sin \varphi, z) r \, dr \, d\varphi \, dz$$

c) Kugelkoordinaten

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

$$dV = r^2 \sin \psi \, dr \, d\psi \, d\varphi$$

$$\iiint_B f(\mathbf{x}) \, dx \, dy \, dz = \iiint_U \tilde{f}(r, \psi, \varphi) r^2 \sin \psi \, dr \, d\psi \, d\varphi$$

mit  $\tilde{f}(r, \psi, \varphi) = f(r \cos \psi \cos \varphi, r \sin \psi \sin \varphi, r \cos \psi)$ .

**Beispiel:** Das axiale Trägheitsmoment der Kugel  $K$  mit dem Radius  $R$  und Dichte  $\varrho$  um die  $z$ -Achse ist

$$\begin{aligned} \iiint_K (x^2 + y^2) \varrho \, dV &= \int_0^{2\pi} \int_0^\pi \int_0^R r^2 \sin^2 \psi \varrho r^2 \sin \psi \, dr \, d\psi \, d\varphi \\ &= 2\pi \frac{R^5 \varrho}{5} \int_0^\pi \sin^3 \psi \, d\psi = \frac{2}{5} M R^2, \end{aligned}$$

da die Masse gegeben ist durch  $M = \varrho V(K) = \frac{4\pi}{3} R^3 \varrho$ .

**9.5.9 Satz (Gauss).** Sei  $B$  ein regulärer räumlicher Bereich mit einer Oberfläche  $\partial B$ , die aus endlich vielen geschlossenen stückweise regulären orientierbaren Flächen besteht, die sich höchstens in Randpunkten treffen. Bezeichnet  $\mathbf{n}$  die aus allen regulären Oberflächenstücken aus  $B$  nach außen weisende Einheitsnormale, dann gilt für alle in einer Umgebung von  $B$  definierten  $C^1$ -Vektorfelder  $\mathbf{v}$  die Formel

$$\iiint_B \operatorname{div} \mathbf{v} \, dV = \iint_{\partial B} \mathbf{v} \cdot \mathbf{n} \, dO. \quad (5.13)$$

**9.5.10 Interpretation der Divergenz.** Nach dem Satz von Gauss ist der Fluß des Vektorfeldes  $\mathbf{v}$  durch die Oberfläche (von innen nach außen) gleich dem

*Volumenintegral der Divergenz.* Sei  $P \in B$  und  $B_r(P)$  eine Kugel mit dem Mittelpunkt  $P$  die ganz in  $B$  liegt, sei  $S_r$  die Kugeloberfläche. Dann gilt

$$\iint_{S_r} \mathbf{v} \cdot \mathbf{n} \, dO = \iiint_{B_r(P)} \operatorname{div} \mathbf{v} \, dV = \operatorname{div} \mathbf{v}(P^*) V(B_r).$$

Der Grenzübergang  $r \rightarrow 0$  liefert

$$\operatorname{div} \mathbf{v}(P) = \lim_{r \rightarrow 0} \frac{1}{V(B_r)} \iint_{S_r} \mathbf{v} \cdot \mathbf{n} \, dO,$$

d.h. die Divergenz ist Fluß pro Volumeneinheit. Anders gesagt ist die Divergenz eines Vektorfeldes in einem Punkt die Quelldichte. Ist sie ungleich Null hat das Vektorfeld dort eine Quelle ( $\operatorname{div} \mathbf{v} > 0$ ) oder Senke ( $\operatorname{div} \mathbf{v} < 0$ ). Ein Vektorfeld heißt **quellfrei** wenn überall  $\operatorname{div} \mathbf{v} = 0$  gilt.

**Beispiel:** Der Fluß des Feldes  $\mathbf{v} = (xy^2, x^2y, y)^T$  durch die Oberfläche des Zylinders  $B = \{(x, y, z); x^2 + y^2 \leq 1, -1 \leq z \leq 1\}$  ist

$$\begin{aligned} \iint_S \mathbf{v} \cdot \mathbf{n} \, dO &= \iiint_B \operatorname{div} \mathbf{v} \, dV \\ &= \iint_{B_1(0)} \int_{-1}^1 x^2 + y^2 \, dz \, dx \, dy \\ &= 2 \int_0^{2\pi} \int_0^1 r^2 r \, dr \, d\varphi = \pi. \end{aligned}$$

Da  $\operatorname{div} \mathbf{v} = x^2 + y^2 \neq 0$  für alle  $(x, y) \neq (0, 0)$  ist jeder Punkt außerhalb der  $z$ -Achse ein Quellpunkt.

**9.5.11 Massenerhaltung.** Sei  $\mathbf{v}$  das Geschwindigkeitsfeld einer Flüssigkeit und  $\rho$  ihre Dichte. Sie  $B$  ein reguläres Testvolumen mit der Oberfläche  $S$ . Durch das Oberflächenelement  $dO$  tritt pro Zeiteinheit das Flüssigkeitsvolumen  $\mathbf{v} \cdot \mathbf{n} \, dO$  hindurch, also ist

$$\iint_S \rho \mathbf{v} \cdot \mathbf{n} \, dO$$

die durch  $S$  aus  $B$  herausströmende Masse. Wird in  $B$  keine Masse erzeugt oder vernichtet ist die zeitliche Veränderung der Masse gegeben durch (Massenfluß aus  $B$  heraus)

$$-\frac{d}{dt} \iiint_B \rho \, dV.$$

Also gilt unter geeigneten Voraussetzungen an  $B, S, \mathbf{v}$  und  $\varrho$

$$\iiint_B \frac{\partial}{\partial t} \varrho dV + \iint_S \varrho \mathbf{v} \cdot \mathbf{n} dO = 0$$

und mit Satz 9.5.9

$$\iiint_B \frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) dV = 0.$$

Dies gilt für beliebige Testvolumen  $B$ , also auch punktweise, d.h. wir haben

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) = 0,$$

das sogenannte **Massenerhaltungsgesetz**.

- Mit ähnlichen Überlegungen können die **Wärmeleitungsgleichung**

$$\frac{\partial \theta}{\partial t} - k \Delta \theta = 0,$$

mit der Temperatur  $\theta$  und dem Wärmeleitungskoeffizienten  $k$ , sowie die **Maxwell Gleichungen**

$$\begin{aligned} \operatorname{rot} \mathbf{E} + \frac{\partial}{\partial t} \mathbf{B} &= 0 \\ \operatorname{rot} \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \\ \operatorname{div} \mathbf{D} &= \varrho \\ \operatorname{div} \mathbf{B} &= 0 \end{aligned}$$

hergeleitet werden.