

M A T H E M A T I K    I + II

2012/2013

**Ernst Kuwert**

Mathematisches Institut  
Universität Freiburg



# Inhaltsverzeichnis

<b>1</b>	<b>Zahlen und Vektoren</b>	<b>1</b>
1	Mengen und Abbildungen . . . . .	1
2	Zahlen . . . . .	4
3	Vollständige Induktion . . . . .	8
4	Geometrie im $\mathbb{R}^n$ . . . . .	12
5	Die komplexen Zahlen . . . . .	18
<b>2</b>	<b>Funktionen, Grenzwerte, Stetigkeit</b>	<b>21</b>
1	Reelle Funktionen . . . . .	21
2	Polynome und rationale Funktionen . . . . .	22
3	Kreisfunktionen . . . . .	26
4	Zahlenfolgen und Grenzwerte . . . . .	30
5	Grenzwerte und Stetigkeit von Funktionen . . . . .	39
<b>3</b>	<b>Differentialrechnung für Funktionen einer Variablen</b>	<b>45</b>
1	Die Ableitung: Definition und Regeln . . . . .	45
2	Mittelwertsatz und Anwendungen . . . . .	50
<b>4</b>	<b>Integralrechnung</b>	<b>55</b>
1	Das Riemannsches Integral . . . . .	55
2	Ableitung und Integral . . . . .	59
3	Darstellung von Funktionen durch Reihen . . . . .	67
4	Taylorentwicklung . . . . .	73
<b>5</b>	<b>Lineare Algebra</b>	<b>77</b>
1	Vektorräume . . . . .	77
2	Lineare Abbildungen und Matrizen . . . . .	84
3	Lineare Gleichungssysteme . . . . .	90
4	Die inverse Matrix . . . . .	92
5	Determinanten . . . . .	97
6	Lineare Abbildungen und Koordinatendarstellungen . . . . .	106
7	Eigenwerte und Normalformen . . . . .	115
<b>6</b>	<b>Differentialrechnung für Funktionen in mehreren Variablen</b>	<b>123</b>
1	Topologische Grundbegriffe . . . . .	123
2	Die partiellen Ableitungen . . . . .	127
3	Die Jacobimatrix . . . . .	131

4	Erste Anwendungen der Ableitung . . . . .	135
5	Parameterabhängige Integrale . . . . .	139
6	Kurvenintegrale . . . . .	144
7	Lokale Auflösung von Gleichungen . . . . .	151
<b>7</b>	<b>Mehrdimensionale Integration</b>	<b>157</b>
1	Prinzip von Cavalieri und Fubini . . . . .	157
2	Koordinatentransformationen . . . . .	162
3	Der Satz von Gauß . . . . .	168

# Kapitel 1

## Zahlen und Vektoren

### 1 Mengen und Abbildungen

Die Sprache der Mengenlehre ist grundlegend zur Formulierung und Kommunikation mathematischer Sachverhalte. Hier finden Sie einige Begriffe: Angabe von Mengen, Teil- und Obermenge, Vereinigung, Schnitt, Differenz, Komplement, leere Menge. Weiter erklären wir die Angabe von Funktionen, inklusive der Begriffe Bild, Urbild, Einschränkung, Verkettung, Graph einer Funktion. Wir behandeln auch die wichtigen Begriffe injektiv, surjektiv, bijektiv und Umkehrfunktion.

Um mathematisch zu kommunizieren, ist die Alltagssprache nur teilweise geeignet. Wir brauchen mathematische Vokabeln, zum Beispiel die Grundbegriffe Menge und Abbildung.

Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten zu einem Ganzen (Kurzfassung der Definition von Cantor, 1895).

Die Objekte nennen wir *Elemente der Menge*. Die Notation  $a \in M$  bedeutet, dass  $a$  ein Element der Menge  $M$  ist; andernfalls schreiben wir  $a \notin M$ . Die Entscheidung, ob irgendein Objekt  $a$  Element von  $M$  ist oder nicht, muss mithilfe der Beschreibung von  $M$  immer möglich sein. Beispiele:

- (a) Das griechische Alphabet ist die Menge

$$M = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, \xi, \omicron, \pi, \rho, \sigma, \tau, \upsilon, \phi, \psi, \chi, \psi, \omega\}$$

Die Elemente sind die einzelnen Buchstaben, und wir haben  $M$  durch *Aufzählen aller Elemente* angegeben. Die Menge  $M$  bleibt gleich, wenn wir die Reihenfolge in der Aufzählung ändern. Oft werden die Elemente nur unvollständig aufgezählt in der Erwartung, dass man sich den Rest denken kann:  $M = \{\alpha, \dots, \omega\}$ .

- (b) Betrachte die Menge  $N = \{1, 2, \dots, 29, 30\}$  sowie

$$M = \{p \in N \mid p \text{ ist Primzahl}\}.$$

Hier entsteht  $M$  durch Auswahl von Elementen der Menge  $N$  mittels einer Eigenschaft, nämlich Primzahl zu sein. Es ist  $M = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$ .

(c) Eine Variante der aufzählenden Beschreibung haben wir bei

$$M = \{n^2 | n = 1, 2, 3, \dots\}.$$

Hier wird  $M$  durch die Menge der natürlichen Zahlen “parametrisiert”.

Eine Menge  $M$  heißt Teilmenge der Menge  $N$ , wenn jedes Element von  $M$  auch Element von  $N$  ist (Notation:  $M \subset N$ ). Gibt es außerdem mindestens ein Element von  $N$ , das nicht zu  $M$  gehört, so ist  $M$  eine echte Teilmenge von  $N$ . Beispiel: unter den Studierenden der Mikrosystemtechnik bilden die männlichen eine echte Teilmenge, das heißt es gibt mindestens eine Studentin. Statt  $M \subset N$  schreiben wir manchmal auch  $N \supset M$ , also  $N$  ist Obermenge von  $M$ .

Zur Veranschaulichung stellen wir Mengen ab und zu als Gebiete in der Ebene dar (Venn-Diagramm). Für zwei Mengen  $A$  und  $B$  können wir dann folgende Mengen darstellen:

$$\begin{aligned} A \cup B &= \{x | x \in A \text{ oder } x \in B\} && \text{(Vereinigungsmenge)} \\ A \cap B &= \{x | x \in A \text{ und } x \in B\} && \text{(Schnittmenge)} \\ A \setminus B &= \{x | x \in A \text{ und } x \notin B\} && \text{(Differenzmenge)}. \end{aligned}$$

Wir führen das Symbol  $\emptyset$  ein für die Menge, die kein Element enthält, die *leere Menge*. Das ist praktisch zum Beispiel bei der Bildung von Schnittmengen, es gilt

$$A \cap B = \emptyset \quad \text{wenn } A, B \text{ kein gemeinsames Element haben.}$$

Ist  $X$  eine gegebene Grundmenge und  $A \subset X$ , so nennen wir  $A^C = X \setminus A$  das Komplement von  $A$ . Das Komplement hängt von der Grundmenge  $X$  ab: die Aussage *ich trinke alles außer Ganter* hat verschiedene Bedeutung, je nachdem ob sie sich auf die alle badischen Biere bezieht oder auf alle alkoholischen Getränke.

Oben haben wir die Worte *und* sowie *oder* benutzt, die Aussagen miteinander logisch verknüpfen. Unter einer *Aussage* verstehen wir einen Satz, bei dem wir sinnvoll nach dem Wahrheitswert fragen können, das heißt ist der Satz *wahr* oder *falsch*? Zum Beispiel ist *Der Mars besteht aus grünem Käse* eine Aussage, während *Geh nach Hause!* oder *Was gibts heute in der Mensa?* keine Aussagen sind. Für Aussagen  $P, Q$  setze

$$\begin{array}{lll} P \wedge Q & P \text{ und } Q & \text{wahr genau wenn } P \text{ und } Q \text{ beide wahr,} \\ P \vee Q & P \text{ oder } Q & \text{falsch genau wenn } P \text{ und } Q \text{ beide falsch,} \\ \neg P & \text{nicht } P & \text{wahr genau wenn } P \text{ falsch,} \\ P \Rightarrow Q & \text{aus } P \text{ folgt } Q & \text{falsch genau wenn } P \text{ wahr und } Q \text{ falsch.} \end{array}$$

Es handelt sich jeweils um Aussagen, denn rechts steht, wie die Frage nach wahr oder falsch zu entscheiden ist, abhängig von der Richtigkeit von  $P$  und  $Q$ . Beachten Sie: das *oder* ist nicht ausschließend, kein entweder/oder. Das beliebte Äquivalenzzeichen  $P \Leftrightarrow Q$  bedeutet einfach  $P \Rightarrow Q$  und  $Q \Rightarrow P$ . Weitere Symbole sind

$$\begin{array}{ll} \forall a \in A : E(a) & \text{für alle } a \in A \text{ gilt Eigenschaft } E(a) \\ \exists a \in A : E(a) & \text{es gibt ein } a \in A \text{ mit Eigenschaft } E(a). \end{array}$$

Aus Gründen der Lesbarkeit werden wir  $\wedge, \vee, \neg$  nicht verwenden, und  $\forall, \exists$  sowie  $\Leftrightarrow$  eher sparsam. Bei  $\Leftrightarrow$  müssen beide Implikationen  $\Rightarrow$  und  $\Leftarrow$  gecheckt werden. Das wird oft vergessen und ist auch unnötig, wenn nur eine Richtung gebraucht wird.

Wir kommen zum zweiten Begriff aus der Überschrift. Seien  $X, Y$  Mengen. Eine Abbildung von  $X$  nach  $Y$  (gleichbedeutend: eine Funktion auf  $X$  mit Werten in  $Y$ ) schreiben wir in der Form

$$f : X \rightarrow Y, x \mapsto f(x).$$

Jedem  $x \in X$  wird genau ein Bildpunkt (Funktionswert)  $f(x) \in Y$  zugeordnet. Die Menge  $X$  heißt Definitionsbereich von  $f$ . Mit dem Bild von  $f$  meinen wir die Menge der Bildpunkte (Menge der Funktionswerte)

$$f(X) = \{f(x) : x \in X\} \subset Y.$$

Das Urbild einer Menge  $B \subset Y$  unter  $f$  ist die Menge

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

Verkleinern des Definitionsbereichs ergibt eine Einschränkung von  $f$ . Für  $A \subset X$  ist

$$f|_A : A \rightarrow Y, (f|_A)(x) = f(x) \quad \text{für alle } x \in A.$$

*Beispiel:* sei  $X$  die Menge aller Waren in einem Supermarkt und  $Y$  die Menge aller möglichen Preise in Euro. Ordnen wir jeder Ware  $x \in X$  einen Preis  $f(x) \in Y$  zu, so haben wir eine Funktion  $f : X \rightarrow Y$ . Das Urbild der Menge  $\{0, 99\}$  ist die Menge der Waren, die 0,99 kosten. Betrachten wir statt aller Waren nur die Menge  $W$  der Wurstwaren, so ergibt sich die Einschränkung  $f|_W$ .

Eine naheliegende Abbildung ist  $\text{id}_X : X \rightarrow X, \text{id}_X(x) = x$ , die Identität oder identische Abbildung auf  $X$ . Die Verkettung (oder Hintereinanderschaltung) von Abbildungen  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  ist

$$g \circ f : X \xrightarrow{f} Y \xrightarrow{g} Z, (g \circ f)(x) = g(f(x)).$$

Eine aus der Schule bekannte Möglichkeit, sich Funktionen bzw. Abbildungen vorzustellen, bietet der Graph. Dazu bilden wir das kartesische Produkt

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

Es gilt  $(x, y) = (x', y')$  genau wenn  $x = x'$  und  $y = y'$ . Der Graph von  $f$  ist die Teilmenge

$$G_f = \{(x, f(x)) : x \in X\} \subset X \times Y.$$

Als nächstes drei Begriffe zum Abbildungsverhalten:  $f : X \rightarrow Y$  heißt

*injektiv* wenn gilt: aus  $f(x) = f(x')$  folgt  $x = x'$ ;

*surjektiv* wenn gilt: zu jedem  $y \in Y$  gibt es ein  $x \in X$  mit  $f(x) = y$  (also  $Y = f(X)$ );

*bijektiv* wenn  $f$  injektiv und surjektiv ist.

Ordnen wir jedem Kind seine Mutter zu, so erhalten wir eine Abbildung von der Menge  $K$  aller Kinder in die Menge  $F$  aller Frauen. Diese Abbildung ist nicht surjektiv, denn nicht jede Frau ist Mutter. Sie ist auch nicht injektiv, denn es gibt Mütter mit mehr als einem Kind.

Für die Gleichung  $f(x) = y$ , mit  $y \in Y$  gegeben, bedeuten die Begriffe Folgendes:  $f$  injektiv heißt, die Gleichung hat höchstens eine Lösung.  $f$  surjektiv heißt, es gibt mindestens eine Lösung. Ist  $f : X \rightarrow Y$  bijektiv, so hat die Gleichung genau eine Lösung. Indem wir jedem  $y \in Y$  die jeweilige Lösung zuordnen, erhalten wir eine Abbildung  $g : Y \rightarrow X$ ,  $y \mapsto g(y)$ , mit

$$f(g(y)) = y \text{ für alle } y \in Y, \text{ also } f \circ g = \text{id}_Y.$$

Wählen wir in der Gleichung als rechte Seite  $y := f(x)$ , so ist  $x$  die Lösung, das heißt

$$g(f(x)) = x, \text{ und damit } g \circ f = \text{id}_X.$$

Es ist leicht zu sehen, dass  $g : Y \rightarrow X$  ebenfalls bijektiv ist (Übungsaufgabe). Wir nennen  $g = f^{-1}$  die zu  $f$  inverse Abbildung oder Umkehrfunktion. Für den Graphen der Umkehrfunktion sehen wir, indem wir  $y = f(x)$  substituieren,

$$G_{f^{-1}} = \{(y, f^{-1}(y)) : y \in Y\} = \{(f(x), x) : x \in X\} \subset Y \times X.$$

Der Graph ergibt sich also durch "Spiegelung an der Winkelhalbierenden".

*Beispiel.* Bezeichne mit  $p_1, p_2$  die Projektionen von  $X \times Y$  auf die Faktoren, gewissermaßen die Koordinaten. Also

$$p_1 : X \times Y \rightarrow X, p_1((x, y)) = x, \quad p_2 : X \times Y \rightarrow Y, p_2((x, y)) = y.$$

Ist  $f : X \rightarrow Y$  eine Funktion, so ist die Abbildung

$$F : X \rightarrow G_f, F(x) = (x, f(x)),$$

bijektiv, die Umkehrabbildung ist  $p_1|_{G_f} : G_f \rightarrow X$ . Es folgt für  $f$  die Darstellung

$$f = p_2 \circ (p_1|_{G_f})^{-1}.$$

Die Funktion  $f$  ist also durch ihren Graph  $G_f$  bestimmt.

## 2 Zahlen

Wir wiederholen aus der Schule die Gesetze der Addition und Multiplikationen reeller Zahlen, und die Anordnungsgesetze mit Definition des Betrags einer reellen Zahl. Die reellen Zahlen werden durch die Existenz des Supremums charakterisiert. Die rationalen Zahlen bilden eine echte Teilmenge, die aber dicht in den reellen Zahlen liegt.

Aus der Schule kennen wir die Zahlen  $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$ :

$$\begin{array}{ll} \mathbb{N} = \{1, 2, 3, \dots\} & \text{natürliche Zahlen} \\ \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} & \text{ganze Zahlen} \\ \mathbb{Q} = \{\frac{p}{q} | q \in \mathbb{N}, p \in \mathbb{Z}\} & \text{rationale Zahlen} \end{array}$$

Die reellen Zahlen  $\mathbb{R}$  fassen wir als Punkte auf der Zahlengeraden auf. Anstatt eine abstrakte Konstruktion von  $\mathbb{R}$  durchzuführen, stellen wir im Folgenden die Regeln (Axiome) in  $\mathbb{R}$  zusammen. Die Gesetze der Addition und Multiplikation lauten wie folgt:

	+	•
<i>Assoziativgesetz:</i>	$(a + b) + c = a + (b + c)$	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$
<i>Kommutativgesetz:</i>	$a + b = b + a$	$a \cdot b = b \cdot a$
<i>Neutrales Element:</i>	$a + 0 = a$	$a \cdot 1 = a$
<i>Inverses Element:</i>	$a + (-a) = 0$	$a \cdot \frac{1}{a} = 1 \quad \text{falls } a \neq 0$
<i>Distributivgesetz:</i>	$a \cdot (b + c) = a \cdot b + a \cdot c.$	

Alle bekannten Rechenregeln wie zum Beispiel  $(-a)(-b) = ab$  können daraus hergeleitet werden. Wir erwähnen die *Nullteilerfreiheit*:

$$\text{Aus } a \cdot b = 0 \text{ folgt } a = 0 \text{ oder } b = 0.$$

Denn ist  $a \neq 0$ , so folgt

$$b = \left(\frac{1}{a} \cdot a\right) \cdot b = \frac{1}{a} \cdot (a \cdot b) = \frac{1}{a} \cdot 0 = 0.$$

Den Punkt bei der Multiplikation lassen wir meistens weg. Auch die Regeln der Bruchrechnung lassen sich aus den Körperaxiomen herleiten, wir listen sie hier auf:

**Satz 2.1 (Bruchrechnung)** Für  $a, b, c, d \in \mathbb{R}$  mit  $c, d \neq 0$  gilt:

- (1)  $\frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd}$  (*Gleichnamig machen von Brüchen*),
- (2)  $\frac{a}{c} \cdot \frac{b}{d} = \frac{ab}{cd}$  (*Multiplikation von Brüchen*),
- (3)  $\frac{a/c}{b/d} = \frac{ad}{bc}$ , falls zusätzlich  $b \neq 0$  (*Division von Brüchen*).

Als nächstes erklären wir die Anordnung von  $\mathbb{R}$ . Die Ungleichung  $a < b$  bedeutet auf der Zahlengeraden, dass  $a$  links von  $b$  liegt bzw.  $a - b$  links von Null. Offenbar gilt genau eine der Relationen  $a < b$ ,  $a > b$  oder  $a = b$ . Hier einige Regeln:

- Aus  $a, b > 0$  folgt  $a + b > 0$  und  $a \cdot b > 0$ .
- Aus  $a > b$ ,  $b > c$  folgt  $a > c$  (*Transitivität*).
- Aus  $a > b$  und  $c > d$  folgt  $a + c > b + d$  (*Addition von Ungleichungen*)
- Aus  $a > b$  folgt  $\begin{cases} ac > bc & \text{falls } c > 0 \\ ac < bc & \text{falls } c < 0 \end{cases}$  (*Multiplikation mit einer Zahl*).

Wir wollen diese und weitere Regeln nicht herleiten. Eine wichtige Konsequenz ist: *Quadrate in  $\mathbb{R}$  sind nichtnegativ*, denn

$$a^2 = \begin{cases} a \cdot a > 0 & \text{im Fall } a > 0, \\ (-a) \cdot (-a) > 0 & \text{im Fall } -a > 0. \end{cases}$$

**Definition 2.1** Der Betrag einer Zahl  $a \in \mathbb{R}$  ist

$$|a| = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0. \end{cases} \quad (2.1)$$

Auf der Zahlengeraden ist  $|a|$  der Abstand zum Nullpunkt. Die Definition des Betrags hat folgende Konsequenz: Wenn Sie ein Betragszeichen beseitigen möchten, müssen Sie eine Fallunterscheidung machen.

**Satz 2.2 (Rechnen mit Beträgen)** Für  $a, b \in \mathbb{R}$  gelten folgende Aussagen:

- (1)  $|-a| = |a|$  und  $a \leq |a|$ .
- (2)  $|a| \geq 0$ ; aus Gleichheit folgt  $a = 0$ .
- (3)  $|ab| = |a| \cdot |b|$ .
- (4)  $|a + b| \leq |a| + |b|$ .
- (5)  $|a - b| \geq ||a| - |b||$ .

BEWEIS: Aus Definition 2.1 folgt

$$|-a| = \begin{cases} -a & \text{falls } -a \geq 0 \\ -(-a) & \text{falls } -a < 0 \end{cases} = \begin{cases} -a & \text{falls } a \leq 0 \\ a & \text{falls } a > 0 \end{cases} = |a|.$$

Weiter folgt (2) aus

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

In (3) bleiben die linke und rechte Seite gleich, wenn wir  $a$  durch  $-a$  ersetzen, dasselbe gilt bezüglich  $b$ . Also können wir  $a, b \geq 0$  annehmen, und erhalten  $|ab| = ab = |a| \cdot |b|$  wie verlangt. Für (4) schätzen wir mit (1) wie folgt ab:

$$|a + b| = \pm(a + b) = \pm a + (\pm b) \leq |a| + |b|.$$

Schließlich gilt  $|a| = |a - b + b| \leq |a - b| + |b|$  nach (4), also  $|a - b| \geq |a| - |b|$ . Durch Vertauschen von  $a$  und  $b$  folgt (5).  $\square$

Alle bis jetzt zusammengestellten Regeln gelten auch für die rationalen Zahlen. Das folgende Beispiel zeigt aber, dass wir mit den rationalen Zahlen nicht auskommen werden. Wir sehen hier ein typisches Beispiel für einen indirekten Beweis.

**Beispiel 2.1** Es gibt kein  $x \in \mathbb{Q}$  mit  $x^2 = 2$ . Angenommen doch: dann können wir  $x > 0$  annehmen (ersetze  $x$  durch  $-x$ ), und weiter nach Kürzen  $x = p/q$  mit  $p, q \in \mathbb{N}$  nicht beide gerade. Aber dann folgt

$$\frac{p^2}{q^2} = 2 \Rightarrow p^2 = 2q^2 \Rightarrow p = 2p' \Rightarrow 4(p')^2 = 2q^2 \Rightarrow q = 2q'.$$

Also sind  $p, q$  doch beide gerade, ein Widerspruch. Wir haben benutzt, dass für  $p$  ungerade auch  $p^2$  ungerade ist. Die Quadrate von ungeraden Zahlen haben nur die Endziffern 1, 9, 5.

Der Unterschied zwischen  $\mathbb{Q}$  und  $\mathbb{R}$  ist die Vollständigkeit, das heißt in  $\mathbb{R}$  können wir beliebige Grenzprozesse durchführen. Zum Beispiel können wir Schritt für Schritt die Ziffern der Dezimalentwicklung von  $\sqrt{2} = 1,4142\dots$  bestimmen. Im ersten Schritt ist  $1^2 < 2$  und  $2^2 > 2$ , also nehmen wir die 1. Im zweiten Schritt ist  $1,4^2 = 1,96$  zu klein,  $1,5^2 = 2,25$  zu groß, also kommt die 4. Auf diese Weise erhalten wir eine Folge, die die irrationale Zahl  $\sqrt{2}$  approximiert. Wir wollen nun eine Version der Vollständigkeit von  $\mathbb{R}$  genauer formulieren, und führen gleichzeitig einige Begriffe ein.

Für  $a, b \in \mathbb{R}$  mit  $a < b$  definieren wir die Intervalle

$$\begin{aligned} (a, b) &= \{x \in \mathbb{R} : a < x < b\} && \text{offenes Intervall} \\ [a, b] &= \{x \in \mathbb{R} : a \leq x \leq b\} && \text{abgeschlossenes Intervall} \\ (a, b] &= \{x \in \mathbb{R} : a < x \leq b\} && \text{rechtsseitig offen, linksseitig abgeschlossen} \\ [a, b) &= \{x \in \mathbb{R} : a \leq x < b\} && \text{linksseitig offen, rechtsseitig abgeschlossen} \\ |I| &= b - a \text{ für ein Intervall } I && \text{Intervalllänge} \end{aligned}$$

Es ist praktisch,  $+\infty$  und  $-\infty$  als *offene* Intervallgrenzen zugelassen, zum Beispiel ist  $(-\infty, 1] = \{x \in \mathbb{R} : -\infty < x \leq 1\}$ . Beachten Sie:  $\pm\infty$  sind keine reellen Zahlen.

**Definition 2.2** Die Menge  $M \subset \mathbb{R}$  heißt

$$\begin{aligned} \text{nach oben beschränkt} &\Leftrightarrow \exists b \in \mathbb{R} \text{ mit } x \leq b \text{ für alle } x \in M, \\ \text{nach unten beschränkt} &\Leftrightarrow \exists a \in \mathbb{R} \text{ mit } x \geq a \text{ für alle } x \in M. \end{aligned}$$

Die Zahl  $b$  heißt dann *obere Schranke* ( $a$  *untere Schranke*). Weiter heißt  $M$  *beschränkt*, wenn  $M$  nach oben und unten beschränkt ist.

**Beispiel 2.2** Die Menge  $[0, 1)$  ist nach oben beschränkt, eine obere Schranke ist zum Beispiel  $b = 2012$ . Es gibt in  $[0, 1)$  aber kein größtes Element, denn es gilt

$$x \in [0, 1) \Rightarrow \frac{x+1}{2} \in [0, 1), \text{ und } \frac{x+1}{2} > x.$$

Unter den oberen Schranken von  $[0, 1)$  gibt es aber eine kleinste, nämlich die Zahl 1.

Die folgende Eigenschaft ist nun für die reellen Zahlen charakteristisch.

**Vollständigkeitsaxiom** Jede nach oben beschränkte Menge  $M \subset \mathbb{R}$  hat eine kleinste obere Schranke.

Die Aussage, dass  $S \in \mathbb{R}$  kleinste obere Schranke von  $M$  ist, bedeutet zwei Dinge:

- (1)  $S$  ist eine obere Schranke von  $M$ , das heißt  $x \leq S$  für alle  $x \in M$ .
- (2)  $S$  ist kleinstmöglich, das heißt für alle  $S' < S$  gibt es ein  $x \in M$  mit  $x > S'$ .

Wir bezeichnen die kleinste obere Schranke mit  $S = \sup M$  (*Supremum von  $M$* ). Ist  $M$  nicht nach oben beschränkt, so definieren wir  $\sup M = +\infty$ . Analog ist das *Infimum*  $\inf M$ , die größte untere Schranke, definiert.

Wie liegen nun die rationalen Zahlen in  $\mathbb{R}$  drin? Die natürlichen Zahlen  $\mathbb{N}$  sind nicht nach oben beschränkt, denn andernfalls haben diese ein Supremum  $S \in \mathbb{R}$ . Dann ist  $S - 1$  keine obere Schranke, das heißt es gibt ein  $n \in \mathbb{N}$  mit  $n > S - 1$ , oder  $n + 1 > S$ . Damit ist  $S$  keine obere Schranke, Widerspruch. Es folgt weiter

Zu jedem  $x \in \mathbb{R}$  mit  $x > 0$  gibt es ein  $n \in \mathbb{N}$  mit  $\frac{1}{n} < x$ .

Denn sonst wäre  $\mathbb{N}$  durch  $\frac{1}{x}$  nach oben beschränkt. Dies bedeutet weiter, dass in jedem nichtleeren Intervall  $(a, b)$  rationale Zahlen liegen. Dazu wählen wir  $n \in \mathbb{N}$  mit  $1/n < b - a$ . Ein ganzzahliges Vielfaches von  $1/n$  muss dann im Intervall  $(a, b)$  liegen.

### 3 Vollständige Induktion

Wir behandeln das Prinzip der vollständigen Induktion. Erste Beispiele sind die rekursive Definition von Summe und Produkt, sowie Beweise der Formeln für die arithmetische und die geometrische Summe. Wir bestimmen die Anzahl von Permutationen und Kombinationen.

Das Beweisverfahren der vollständigen Induktion beruht auf folgendem Prinzip:

Sei  $A(n)$  eine Folge von Aussagen für  $n \in \mathbb{N}$ . Es gelte:

- (1)  $A(1)$  ist wahr.
- (2)  $A(n)$  ist wahr  $\Rightarrow A(n + 1)$  ist wahr.

Dann sind alle Aussagen  $A(n)$  wahr.

Statt bei  $n = 1$  kann die Induktion auch bei  $n_0 \in \mathbb{N}$  starten, die Aussage gilt dann für  $n \geq n_0$ . Das Verfahren kann auch zur Definition benutzt werden, wie hier beim Summen- und Produktzeichen:

$$\sum_{k=1}^{n+1} a_k = \left( \sum_{k=1}^n a_k \right) + a_{n+1}, \quad \sum_{k=1}^1 a_k = a_1.$$

$$\prod_{k=1}^{n+1} a_k = \left( \prod_{k=1}^n a_k \right) \cdot a_{n+1}, \quad \prod_{k=1}^1 a_k = a_1.$$

Die leere Summe wird gleich Null gesetzt, das leere Produkt Eins; sie kommen vor, wenn die obere Grenze der Summation kleiner ist als die untere Grenze. Ein Induktionsbeweis funktioniert immer in zwei Schritten: erst wird die Aussage  $A(n)$  für den Fall  $n = 1$  verifiziert (Induktionsanfang). Im zweiten Schritt wird *vorausgesetzt*, dass  $A(n)$  für ein  $n \in \mathbb{N}$  richtig ist (Induktionsannahme), und daraus  $A(n + 1)$  gefolgert (Induktionsschluss).

**Beispiel 3.1 (arithmetische Summe)** Wir zeigen die Summenformel

$$A(n) : \quad 1 + 2 + \dots + n = \sum_{k=1}^n k = \frac{n(n+1)}{2}.$$

Für  $n = 1$  ist sowohl die linke als auch die rechte Seite gleich Eins, also gilt der Induktionsanfang. Jetzt berechnen wir unter Verwendung von  $A(n)$

$$\sum_{k=1}^{n+1} k = \binom{n}{k=1} + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)((n+1)+1)}{2}.$$

Damit ist  $A(n)$  für alle  $n \in \mathbb{N}$  bewiesen. □

**Beispiel 3.2 (geometrische Summe)** Sei  $x \in \mathbb{R}$ ,  $x \neq 1$ . Dann gilt für alle  $n \in \mathbb{N}_0$

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

Wir zeigen das wieder durch vollständige Induktion, wobei wir bei  $n = 0$  beginnen:

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

Jetzt gelte die Formel für ein  $n \in \mathbb{N}$ . Dann folgt

$$\sum_{k=0}^{n+1} x^k = \left( \sum_{k=0}^n x^k \right) + x^{n+1} \stackrel{A(n)}{=} \frac{1 - x^{n+1}}{1 - x} + x^{n+1} = \frac{1 - x^{(n+1)+1}}{1 - x},$$

womit die Behauptung gezeigt ist. Falls Sie die Formel vergessen haben, können Sie sie herleiten mit dem Teleskopsummentrick:

$$1 - x^{n+1} = \sum_{k=0}^n x^k - \sum_{k=1}^{n+1} x^k = \sum_{k=0}^n (x^k - x^{k+1}) = \sum_{k=0}^n (1 - x)x^k = (1 - x) \sum_{k=0}^n x^k.$$

□

Ebenfalls mit dem Beweisverfahren der vollständigen Induktion zeigen wir folgende nützliche Ungleichung.

**Satz 3.1 (Bernoullische Ungleichung)** Für  $x \in \mathbb{R}$ ,  $x \geq -1$ , und  $n \in \mathbb{N}_0$  gilt

$$(1 + x)^n \geq 1 + nx.$$

BEWEIS: Wir führen Induktion über  $n \in \mathbb{N}_0$ . Für  $n = 0$  gilt nach Definition  $(1 + x)^0 = 1 = 1 + 0 \cdot x$ . Wegen  $1 + x \geq 0$  folgt weiter

$$\begin{aligned} (1 + x)^{n+1} &= (1 + x) \cdot (1 + x)^n \\ &\geq (1 + x) \cdot (1 + nx) \quad (\text{nach Induktionsannahme}) \\ &= 1 + (n + 1)x + nx^2 \\ &\geq 1 + (n + 1)x. \end{aligned}$$

□

Wir wollen als nächstes die Elemente gewisser Mengen zählen. Als erstes zeigen wir, dass man nicht vier Socken in drei verschiedene Schubladen packen kann, oder so.

**Satz 3.2 (Schubfachprinzip)** Ist  $f : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  injektiv mit  $m, n \in \mathbb{N}$ , so folgt  $m \leq n$ .

BEWEIS: Wir führen Induktion nach  $n \in \mathbb{N}$ . Für  $n = 1$  haben wir  $f : \{1, \dots, m\} \rightarrow \{1\}$  injektiv, das geht nur für  $m = 1$ . Sei nun eine injektive Abbildung  $f : \{1, \dots, m\} \rightarrow \{1, \dots, n+1\}$  gegeben; wir müssen zeigen dass  $m \leq n+1$  ist. Hat die Zahl  $n+1$  kein Urbild, so ist tatsächlich  $f : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  und nach Induktion gilt sogar  $m \leq n$ . Andernfalls hat  $n+1$  genau ein Urbild. Wenn wir dieses rausschmeißen und neu nummerieren, erhalten wir eine Abbildung

$$\tilde{f} : \{1, \dots, m-1\} \rightarrow \{1, \dots, n\} \quad \text{injektiv.}$$

Nach Induktion folgt  $m-1 \leq n$  bzw.  $m \leq n+1$  wie gewünscht.  $\square$

**Definition 3.1 (Zahl der Elemente)** Sei  $M$  eine Menge. Gibt es  $f : \{1, \dots, n\} \rightarrow M$  bijektiv für ein  $n \in \mathbb{N}$ , so heißt  $M$  endlich und  $n = \#M$  ist die Anzahl seiner Elemente. Die leere Menge ist ebenfalls endlich mit  $\#\emptyset = 0$ .

Dass die Zahl der Elemente eindeutig definiert ist, folgt aus dem Schubfachprinzip. Wir setzen

$$n! = \prod_{k=1}^n k = 1 \cdot 2 \cdot \dots \cdot n \quad (n\text{-Fakultät}).$$

Per Dekret ist  $0! = 1$  (vgl. Vereinbarung zum leeren Produkt). Der folgende Satz beantwortet die Frage nach der Anzahl der möglichen Anordnungen (oder Umordnungen oder Permutationen) von  $n$  Dingen.

**Satz 3.3 (Zahl der Permutationen)** Für  $n \in \mathbb{N}$  sei  $S_n$  die Menge der bijektiven Abbildungen  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ . Dann gilt  $\#S_n = n!$ .

BEWEIS: Wir zeigen die Behauptung durch Induktion, wobei der Induktionsanfang  $n = 1$  offensichtlich ist. Sei  $S_{n+1,k}$  die Menge der Permutationen von  $\{1, \dots, n+1\}$ , die die Nummer  $k$  auf  $n+1$  abbilden. Da die restlichen  $n$  Nummern beliebig bijektiv auf  $\{1, \dots, n\}$  abgebildet werden, hat  $S_{n+1,k}$  genau  $n!$  Elemente nach Induktion. Es folgt durch Summation

$$\#S_{n+1} = \sum_{k=1}^{n+1} \#S_{n+1,k} = (n+1) \cdot n! = (n+1)!.$$

$\square$

**Definition 3.2 (Binomialkoeffizienten)** Für  $\alpha \in \mathbb{R}$  und  $k \in \mathbb{N}$  setzen wir

$$\binom{\alpha}{k} = \frac{\alpha \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - k + 1)}{1 \cdot 2 \cdot \dots \cdot k}, \quad \text{sowie} \quad \binom{\alpha}{0} = 1.$$

**Lemma 3.1 (Additionstheorem für Binomialkoeffizienten)** Für  $\alpha \in \mathbb{R}$  und  $k \in \mathbb{N}$  erfüllen die Binomialkoeffizienten die Formel

$$\binom{\alpha + 1}{k} = \binom{\alpha}{k} + \binom{\alpha}{k-1}.$$

BEWEIS: Für  $k = 1$  ist leicht zu sehen, dass die Formel richtig ist. Für  $k \geq 2$  berechnen wir

$$\begin{aligned}
 \binom{\alpha}{k} + \binom{\alpha}{k-1} &= \frac{\alpha \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - k + 1)}{1 \cdot 2 \cdot \dots \cdot k} + \frac{\alpha \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - k + 2)}{1 \cdot 2 \cdot \dots \cdot (k - 1)} \\
 &= \frac{\alpha \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - k + 2) \cdot (\alpha - k + 1 + k)}{1 \cdot 2 \cdot \dots \cdot k} \\
 &= \frac{(\alpha + 1) \cdot \alpha \cdot \dots \cdot ((\alpha + 1) - k + 1)}{1 \cdot 2 \cdot \dots \cdot k} \\
 &= \binom{\alpha + 1}{k}.
 \end{aligned}$$

□

Im Fall  $\alpha = n \in \mathbb{N}_0$  erlaubt Lemma 3.1 die rekursive Berechnung der Binomialkoeffizienten  $\binom{n}{k}$  nach dem Dreiecksschema von Blaise Pascal (1623-1662).

n=0				1				
n=1				1	1			
n=2			1	2	1			
n=3		1	3	3	1			
n=4		1	4	6	4	1		
n=5	1	5	10	10	5	1		
n=6	1	6	15	20	15	6	1	

Ebenfalls für  $\alpha = n \in \mathbb{N}_0$  folgt durch Erweitern der Binomialkoeffizienten mit  $(n - k)!$  die alternative Darstellung

$$\binom{n}{k} = \frac{n!}{k!(n - k)!} \text{ für } n \in \mathbb{N}_0, k \in \{0, 1, \dots, n\}, \tag{3.2}$$

und daraus weiter die am Diagramm ersichtliche Symmetrieeigenschaft

$$\binom{n}{k} = \binom{n}{n - k} \text{ für } n \in \mathbb{N}_0, k \in \{0, 1, \dots, n\}. \tag{3.3}$$

**Satz 3.4 (Zahl der Kombinationen)** Sei  $n \in \mathbb{N}_0$  und  $k \in \{0, 1, \dots, n\}$ . Dann ist die Anzahl der  $k$ -elementigen Teilmengen von  $\{1, \dots, n\}$  gleich  $\binom{n}{k}$ .

BEWEIS: Es gilt  $\binom{n}{0} = 1$  nach Definition, und die einzige null-elementige Teilmenge von  $\{0, 1, \dots, n\}$  ist die leere Menge. Also stimmt die Aussage für alle  $n = 0, 1, \dots$  und  $k = 0$ . Die  $k$ -elementigen Teilmengen von  $\{1, \dots, n + 1\}$  mit  $k \geq 1$  zerfallen in zwei Klassen:

**Klasse 1:** Die Menge enthält die Nummer  $n + 1$  nicht.

**Klasse 2:** Die Menge enthält die Nummer  $n + 1$ .

Klasse 1 besteht genau aus den  $k$ -elementigen Teilmengen von  $\{1, \dots, n\}$ , Klasse 2 ergibt sich durch Hinzufügen der Nummer  $n + 1$  zu jeder der  $(k - 1)$ -elementigen Teilmengen von  $\{1, \dots, n\}$ . Nach Induktion folgt für die Gesamtzahl der Elemente mit Lemma 3.1

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

womit der Satz bewiesen ist. □

**Satz 3.5 (Binomische Formel)** Für  $a, b \in \mathbb{R}$  und  $n \in \mathbb{N}$  gilt

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}. \quad (3.4)$$

BEWEIS: Ausmultiplizieren des  $n$ -fachen Produkts  $(a+b)^n = (a+b) \cdot (a+b) \cdot \dots \cdot (a+b)$  mit dem Distributivgesetz und Ordnen nach dem Kommutativgesetz liefert Terme der Form  $a^k b^{n-k}$  für  $k \in \{0, 1, \dots, n\}$ . Die Häufigkeit eines solchen Terms ist gleich der Anzahl der Möglichkeiten, aus den  $n$  Klammern  $k$  Klammern auszusuchen, in denen  $a$  als Faktor genommen wird; in den restlichen Klammern muss dann der Faktor  $b$  gewählt werden. Nach Satz 3.4 kommt  $a^k b^{n-k}$  also genau  $\binom{n}{k}$  mal vor. □

## 4 Geometrie im $\mathbb{R}^n$

Wir führen den Raum  $\mathbb{R}^n$  der Vektoren  $\vec{x} = (x_1, \dots, x_n)$  ein, zusammen mit der Vektoraddition und der Skalarmultiplikation. Durch Wahl eines kartesischen Koordinatensystems werden die Punkte einer Geraden mit  $\mathbb{R}$ , Punkte einer Ebene mit  $\mathbb{R}^2$  und Punkte im Raum mit  $\mathbb{R}^3$  identifiziert. Für die Ebene berechnen wir die Transformation der Koordinaten bei Drehung des Koordinatensystems. Wir führen weiter den Euklidischen Betrag bzw. Abstand ein, sowie das Skalarprodukt und den Winkel. Auf  $\mathbb{R}^3$  behandeln wir schließlich das Vektorprodukt.

Im folgenden sei  $n \in \mathbb{N}$ , zum Beispiel  $n = 2$  oder  $n = 3$ . Wir bezeichnen mit  $\mathbb{R}^n$  das  $n$ -fache kartesische Produkt  $\mathbb{R} \times \dots \times \mathbb{R}$ , also die Menge aller  $n$ -Tupel

$$\mathbb{R}^n = \left\{ \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R} \right\}.$$

Wir nennen die Elemente von  $\mathbb{R}^n$  auch *Vektoren*. Zwei Vektoren  $\vec{x}, \vec{y} \in \mathbb{R}^n$  sind genau dann gleich, wenn  $x_i = y_i$  für alle  $i = 1, \dots, n$ , wenn also alle *Koordinaten* gleich sind. Wir können Vektoren addieren und mit reellen Zahlen multiplizieren (zum Beispiel strecken):

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \vec{x} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

In diesem Zusammenhang nennt man die Zahl  $\lambda$  auch einen Skalar, und spricht von Skalarmultiplikation. Die Standard-Einheitsvektoren sind

$$\vec{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad i \quad \text{für } i = 1, \dots, n.$$

Mit ihnen ergibt sich für jeden Vektor  $\vec{x} \in \mathbb{R}^n$  die Darstellung

$$\vec{x} = \sum_{i=1}^n x_i \vec{e}_i.$$

Die Vektoraddition und Skalarmultiplikation erfüllen folgende Regeln, die alle direkt aus den Definitionen folgen:

- (1)  $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$
- (2)  $\vec{x} + \vec{0} = \vec{x}$  für alle  $\vec{x}$ , mit  $\vec{0} = (0, \dots, 0)$  (Nullvektor)
- (3) Es gilt  $\vec{x} + (-\vec{x}) = \vec{0}$
- (4)  $\vec{x} + \vec{y} = \vec{y} + \vec{x}$
- (5)  $(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$
- (6)  $(\lambda\mu)\vec{x} = \lambda(\mu\vec{x})$
- (7)  $\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$
- (8)  $1 \cdot \vec{x} = \vec{x}$ .

Im drei-dimensionalen Raum kann jede Gerade mit  $\mathbb{R}^1 = \mathbb{R}$ , jede Ebene mit  $\mathbb{R}^2$  und der Raum selbst mit  $\mathbb{R}^3$  identifiziert werden, indem ein kartesisches Koordinatensystem gewählt wird. Wir wollen das erklären, indem wir den Begriff des Euklidischen Raums aus der Anschauung verwenden, das heißt der Euklidische Abstand zwischen zwei Punkten sei anschaulich gegeben. Ebenso setzen wir den Begriff der Orthogonalität für sich schneidende Geraden anschaulich voraus.

Auf einer Geraden  $g$  wählen wir einen Punkt  $O$ , den Ursprung (*origin*). Dann besteht  $g \setminus \{O\}$  aus zwei Halbgeraden. Wir erhalten eine Koordinate  $x$  auf  $g$ , indem wir auf einer der Halbgeraden für  $x$  den Abstand zu  $O$  wählen, auf der anderen Halbgeraden den negativen Abstand zu  $O$ . Die Koordinate  $x$  ist eindeutig bestimmt durch die Wahl des Ursprungs  $O$  und die Entscheidung, welche Halbgerade die positive ist.

In einer Ebene  $E$  wählen wir wieder einen Ursprung  $O$ , und weiter zwei zueinander orthogonale Geraden  $g, h$  durch  $O$ . Auf  $g, h$  bestimmen wir jeweils Koordinaten  $x, y$  wie oben beschrieben. Jeder Punkt  $p \in E$  kann jetzt auf  $g$  und  $h$  orthogonal projiziert werden, dies liefert für  $p$  die beiden Koordinaten  $(x, y)$ . Die Frage nach der Eindeutigkeit ist nun

schon schwieriger. Wir nehmen an, dass der Ursprung fest ist und betrachten ein zweites Paar  $g', h'$  zueinander senkrechter Geraden, genauer soll  $g', h'$  aus  $g, h$  durch Drehung um einen Winkel  $\alpha$  hervorgehen. Dann sieht man (Bild)

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

Wir kommen schließlich zur Wahl kartesischer Koordinaten im drei-dimensionalen Raum. Wieder wählen wir einen Ursprung  $O$  und nun drei zueinander senkrechte Geraden  $g, h, \ell$ . Seien  $x, y, z$  die Koordinaten auf den drei einzelnen Geraden wie oben erklärt. Jeder Punkt  $p$  im Raum wird auf die drei Geraden orthogonal projiziert und liefert so die Koordinaten  $(x, y, z)$ . Die Frage nach der Eindeutigkeit bzw. der Umrechnung in ein anderes Koordinatensystem ist etwas schwieriger, wir kommen später in der Vorlesung darauf zurück.

Indem wir  $\mathbb{R}^2$  mit einer Ebene bzw.  $\mathbb{R}^3$  mit dem Euklidischen Raum identifizieren (wie oben beschrieben), erhalten wir eine anschauliche Darstellung der Vektoraddition, Vektorsubtraktion und Skalarmultiplikation (Bild).

Die oben benutzten Begriffe Länge und Winkel wollen wir noch genauer erklären. Die Euklidische Länge (oder der Betrag) eines Vektors ist definiert durch

$$|\vec{x}| = \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad \text{für } \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Für  $n = 2$  ist  $|\vec{x}| = \sqrt{x_1^2 + x_2^2}$ , dies entspricht also dem Satz von Pythagoras (Bild). Für  $n = 3$  haben wir  $|\vec{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ . Dies lässt sich ebenfalls mit Pythagoras begründen, indem der Punkt  $(x_1, x_2, 0)$  betrachtet wird (Bild). Offenbar gilt für  $\vec{x} \in \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}$ ,

$$\begin{aligned} |\vec{x}| &\geq 0 \quad (\text{Gleichheit nur für } \vec{x} = \vec{0}) \\ |\lambda \vec{x}| &= |\lambda| |\vec{x}|. \end{aligned}$$

Der Euklidische Abstand ist

$$\text{dist}(\vec{x}, \vec{y}) = |\vec{x} - \vec{y}| \quad \text{für } \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Insbesondere ist der Betrag gleich dem Abstand zum Nullpunkt, i.e.  $|\vec{x}| = \text{dist}(\vec{x}, \vec{0})$ . Eine weitere nützliche Größe ist das Skalarprodukt zwischen Vektoren im  $\mathbb{R}^n$ :

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i \quad \text{für } \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

Statt mit Klammern wird das Skalarprodukt oft in der Form  $\vec{x} \cdot \vec{y}$  geschrieben, also mit einem Punkt. Wir bleiben aber bei den Klammern. Folgende Regeln ergeben sich direkt aus der Formel, wobei  $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$  und  $\lambda, \mu \in \mathbb{R}$ .

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= \langle \vec{y}, \vec{x} \rangle && (\text{Symmetrie}) \\ \langle \lambda \vec{x} + \mu \vec{y}, \vec{z} \rangle &= \lambda \langle \vec{x}, \vec{z} \rangle + \mu \langle \vec{y}, \vec{z} \rangle && (\text{Bilinearität}) \\ \langle \vec{x}, \vec{x} \rangle &= |\vec{x}|^2 \geq 0 && (\text{Positivität}). \end{aligned}$$

Um nun die Definition des Winkels zu motivieren, appellieren wir an das Schulwissen und betrachten die gleichförmige Kreisbewegung

$$\vec{c}: \mathbb{R} \rightarrow \mathbb{R}^2, \vec{c}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

Es gilt  $|\vec{c}(t)| = \sqrt{(\cos t)^2 + (\sin t)^2} = 1$  und  $\vec{c}(0) = (1, 0)$ . Der Punkt  $\vec{c}(t)$  durchläuft den Kreis entgegen dem Uhrzeigersinn mit konstanter Absolutgeschwindigkeit Eins:

$$\vec{c}'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \quad \text{und damit} \quad |\vec{c}'(t)| = \sqrt{(-\sin t)^2 + (\cos t)^2} = 1.$$

Der Einheitskreis hat die Gesamtlänge  $2\pi = 6,2831\dots$ . Als Maß für den Winkel nehmen wir die Länge des kürzeren Einheitskreisbogens zwischen  $c(t_1)$  und  $c(t_2)$ , also

$$\angle(\vec{c}(t_1), \vec{c}(t_2)) = |t_1 - t_2| \quad \text{falls } |t_1 - t_2| \leq \pi.$$

Die Bedingung  $|t_1 - t_2| \leq \pi$  garantiert, dass der von  $\vec{c}(t)$  auf  $[t_1, t_2]$  durchlaufene Bogen höchstens ein Halbkreis ist. Erhalte mit dem Additionstheorem des Kosinus

$$\cos \angle(\vec{c}(t_1), \vec{c}(t_2)) = \cos |t_1 - t_2| = \cos t_1 \cos t_2 + \sin t_1 \sin t_2 = \langle \vec{c}(t_1), \vec{c}(t_2) \rangle.$$

**Definition 4.1 (Winkel zwischen Vektoren)** Seien  $\vec{x}, \vec{y} \in \mathbb{R}^n$  mit  $\vec{x}, \vec{y} \neq \vec{0}$ . Dann definieren wir den Winkel  $\phi = \angle(\vec{x}, \vec{y}) \in [0, \pi]$  durch die Gleichung

$$\cos \phi = \left\langle \frac{\vec{x}}{|\vec{x}|}, \frac{\vec{y}}{|\vec{y}|} \right\rangle.$$

*Spezialfall:*  $\vec{x}, \vec{y}$  sind zueinander senkrecht (orthogonal), genau wenn  $\langle \vec{x}, \vec{y} \rangle = 0$ .

Es ist noch zu begründen, dass die Gleichung in der Definition eine und nur eine Lösung  $\phi$  besitzt. Die Funktion Kosinus ist auf dem Intervall  $[0, \pi]$  streng monoton fallend mit  $\cos 0 = 1$  und  $\cos \pi = -1$  (Bild). Sie nimmt daher jeden Wert in  $[-1, 1]$  genau einmal an. Es bleibt also zu zeigen, dass die rechte Seite in der Definition des Winkels im Intervall  $[-1, 1]$  liegt. Dies ergibt sich aus folgender Ungleichung.

**Satz 4.1 (Ungleichung von Cauchy-Schwarz)** Für  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gilt

$$|\langle \vec{x}, \vec{y} \rangle| \leq |\vec{x}| |\vec{y}|.$$

*Gleichheit gilt genau dann, wenn  $\vec{x}, \vec{y}$  parallel sind.*

BEWEIS: Für Vektoren  $\vec{a}, \vec{b}$  mit  $|\vec{a}| = |\vec{b}| = 1$  schätzen wir ab:

$$1 \pm \langle \vec{a}, \vec{b} \rangle = \frac{1}{2} \left( |\vec{a}|^2 + |\vec{b}|^2 \pm 2\langle \vec{a}, \vec{b} \rangle \right) = \frac{1}{2} |\vec{a} \pm \vec{b}|^2 \geq 0.$$

Für  $\vec{x}, \vec{y} \neq \vec{0}$  beliebig wenden wir das an mit  $\vec{a} = \frac{\vec{x}}{|\vec{x}|}$ ,  $\vec{b} = \frac{\vec{y}}{|\vec{y}|}$ :

$$|\langle \vec{x}, \vec{y} \rangle| = |\vec{x}| |\vec{y}| |\langle \vec{a}, \vec{b} \rangle| \leq |\vec{x}| |\vec{y}|.$$

Bei Gleichheit gilt für die Einheitsvektoren

$$0 = \vec{a} \pm \vec{b} = \frac{\vec{x}}{|\vec{x}|} \pm \frac{\vec{y}}{|\vec{y}|},$$

das heißt  $\vec{x}$  und  $\vec{y}$  sind parallel. □

Die Definition des Winkels ergibt trivial den *Kosinussatz*: in dem von  $\vec{0}$ ,  $\vec{x}$ ,  $\vec{y}$  gebildeten Dreieck folgt nämlich (Bild)

$$|\vec{x} - \vec{y}|^2 = |\vec{x}|^2 + |\vec{y}|^2 - 2\langle \vec{x}, \vec{y} \rangle = |\vec{x}|^2 + |\vec{y}|^2 - 2|\vec{x}||\vec{y}|\cos\angle(\vec{x}, \vec{y}).$$

Eine Konsequenz der Cauchy-Schwarz Ungleichung ist folgende Eigenschaft des Betrags.

**Satz 4.2 (Dreiecksungleichung)** Für  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gilt die Ungleichung

$$|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|,$$

mit Gleichheit genau wenn  $\vec{x}, \vec{y}$  gleichsinnig parallel sind.

BEWEIS: Aus der Cauchy-Schwarz Ungleichung folgt

$$|\vec{x} + \vec{y}|^2 = |\vec{x}|^2 + 2\langle \vec{x}, \vec{y} \rangle + |\vec{y}|^2 \leq |\vec{x}|^2 + 2|\vec{x}||\vec{y}| + |\vec{y}|^2 = (|\vec{x}| + |\vec{y}|)^2.$$

Durch Wurzelziehen folgt  $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$  wie behauptet. Bei Gleichheit müssen  $\vec{x}$  und  $\vec{y}$  parallel sein nach Satz 4.1. Es muss aber auch  $\langle \vec{x}, \vec{y} \rangle \geq 0$  sein, also sind  $\vec{x}, \vec{y}$  gleichsinnig parallel. □

Statt auf die Koordinatenachsen können wir einen Vektor  $\vec{x}$  auch auf eine beliebige Ursprungsgerade  $g$  projizieren. Wird  $g$  durch den Einheitsvektor  $\vec{v} \in \mathbb{R}^n$  erzeugt, das heißt  $g = \{\lambda\vec{v} \mid \lambda \in \mathbb{R}\}$ , so lautet der Projektionspunkt

$$\vec{x}^g = \langle \vec{x}, \vec{v} \rangle \vec{v}.$$

Denn  $\vec{x}^g$  liegt auf der Geraden  $g$ , und  $\vec{x} - \vec{x}^g$  steht senkrecht auf  $g$ :

$$\langle \vec{x} - \vec{x}^g, \vec{v} \rangle = \langle \vec{x}, \vec{v} \rangle - \langle \langle \vec{x}, \vec{v} \rangle \vec{v}, \vec{v} \rangle = \langle \vec{x}, \vec{v} \rangle - \langle \vec{x}, \vec{v} \rangle \underbrace{\langle \vec{v}, \vec{v} \rangle}_{=1} = 0.$$

Die Zahl  $\cos\angle(\vec{x}, \vec{v})$  wird manchmal als Richtungskosinus von  $\vec{x}$  bezüglich  $\vec{v}$  bezeichnet.

Wir kommen nun zu einer speziellen Struktur des drei-dimensionalen Raums. Das Kreuzprodukt (oder Vektorprodukt) ist gegeben durch

$$\vec{x} \times \vec{y} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Folgende Regeln sind offensichtlich:

$$\begin{aligned} \vec{x} \times \vec{y} &= -\vec{y} \times \vec{x} \quad (\text{Schiefsymmetrie}) \\ (\lambda\vec{x} + \mu\vec{y}) \times \vec{z} &= \lambda\vec{x} \times \vec{z} + \mu\vec{y} \times \vec{z} \quad (\text{Bilinearität}). \end{aligned}$$

Insbesondere gilt  $\vec{x} \times \vec{x} = 0$ . Für die Standardvektoren sehen wir

$$\begin{aligned}\vec{e}_1 \times \vec{e}_2 &= \vec{e}_3 = -\vec{e}_2 \times \vec{e}_1 \\ \vec{e}_2 \times \vec{e}_3 &= \vec{e}_1 = -\vec{e}_3 \times \vec{e}_2 \\ \vec{e}_3 \times \vec{e}_1 &= \vec{e}_2 = -\vec{e}_1 \times \vec{e}_3\end{aligned}$$

Man nennt die Vertauschungen 123, 231, 312 zyklisch, 213, 321, 132 antizyklisch. Welcher Vektor ist nun  $\vec{x} \times \vec{y}$ ? Als erstes berechnen wir dazu seine Länge (nachprüfen)

$$\begin{aligned}|\vec{x} \times \vec{y}|^2 &= (x_1y_2 - x_2y_1)^2 + (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= |\vec{x}|^2|\vec{y}|^2 - \langle \vec{x}, \vec{y} \rangle^2.\end{aligned}$$

Mit  $\langle \vec{x}, \vec{y} \rangle = |\vec{x}||\vec{y}| \cos \angle(\vec{x}, \vec{y})$  (Definition des Winkel) ergibt sich

$$|\vec{x} \times \vec{y}| = |\vec{x}||\vec{y}| \sqrt{1 - \cos^2 \angle(\vec{x}, \vec{y})} = |\vec{x}||\vec{y}| \sin \angle(\vec{x}, \vec{y}).$$

Insbesondere:  $\vec{x} \times \vec{y}$  ist null, wenn  $\angle(\vec{x}, \vec{y}) \in \{0, \pi\}$ , also wenn  $\vec{x}, \vec{y}$  parallel sind (oder null). Sind  $\vec{x}, \vec{y}$  orthogonal, so folgt  $\vec{x} \times \vec{y} = |\vec{x}||\vec{y}|$ .

Ab jetzt seien  $\vec{x}, \vec{y}$  nicht parallel, also  $\vec{x} \times \vec{y} \neq \vec{0}$ . Wir berechnen

$$\langle \vec{x} \times \vec{y}, \vec{z} \rangle = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2. \quad (4.1)$$

Ist  $\vec{z} = \vec{x}$  oder  $\vec{z} = \vec{y}$ , so ist die rechte Seite Null. Also steht  $\vec{x} \times \vec{y}$  senkrecht auf die von  $\vec{x}, \vec{y}$  aufgespannte Ebene. Da die Länge schon bekannt ist, bleiben nur zwei Vektoren, die sich um den Faktor  $-1$  unterscheiden.

Die rechte Seite von Gleichung (4.1) wird mit  $\det(\vec{x}, \vec{y}, \vec{z})$  bezeichnet (Determinante). Um sich die Formel zu merken, kann man die Vektoren in ein Schema schreiben und jeweils die Produkte über die Diagonalen bilden. Dabei werden Diagonalen nach rechts unten positiv, Diagonalen nach links unten negativ gezählt (Regel von Sarrus).

$$\left( \begin{array}{ccc|cc} x_1 & x_2 & x_3 & x_1 & x_2 \\ y_1 & y_2 & y_3 & y_1 & y_2 \\ z_1 & z_2 & z_3 & z_1 & z_2 \end{array} \right).$$

Die Zahl  $\det(\vec{x}, \vec{y}, \vec{z})$  wird auch als Spatprodukt der drei Vektoren bezeichnet. Sie ist gleich dem Volumen des von den Vektoren gebildeten Spats (Parallelepipeds), bis auf das Vorzeichen. Das wollen wir aber hier nicht ausführen. Wir vereinbaren:

$$\vec{x}, \vec{y}, \vec{z} \text{ sind positiv orientiert, genau wenn } \det(\vec{x}, \vec{y}, \vec{z}) > 0.$$

Beachten Sie, dass es hier auf die Reihenfolge der Vektoren ankommt: werden zwei Vektoren vertauscht, so ändert die Determinante ihr Vorzeichen. Die Richtung von  $\vec{x} \times \vec{y} \neq \vec{0}$  ist dadurch bestimmt, dass die Vektoren  $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$  in dieser Reihenfolge positiv orientiert sind. Denn wählen wir in der Formel für die Determinante  $\vec{z} = \vec{x} \times \vec{y}$ , so ist

$$\det(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) = \langle \vec{x} \times \vec{y}, \vec{x} \times \vec{y} \rangle = |\vec{x} \times \vec{y}|^2 > 0.$$

Zum Schluss eine Bemerkung zu Anschauungsraum und  $\mathbb{R}^3$ . Wählen wir ein kartesisches Koordinatensystem, so werden jedem Vektor  $\vec{v}$  im Raum Koordinaten  $(x, y, z) \in \mathbb{R}^3$  zugeordnet. Wir bestimmen seine Länge dann mit der Formel  $|\vec{v}| = \sqrt{x^2 + y^2 + z^2}$ . Es ist wesentlich, dass wir bei Wahl eines anderen kartesischen Koordinatensystems denselben Wert erhalten. Im  $\mathbb{R}^2$ , wo wir die Umrechnung der Koordinaten kennen, sehen wir das explizit:

$$(x')^2 + (y')^2 = ((\cos \alpha)x + (\sin \alpha)y)^2 + ((-\sin \alpha)x + (\cos \alpha)y)^2 = x^2 + y^2.$$

Auch der Winkel zwischen zwei Vektoren im Raum ist unabhängig von der Wahl des kartesischen Koordinatensystems. Beim Vektorprodukt gibt es eine gewisse Einschränkung: wählen wir zum Beispiel die zueinander senkrechten Daumen, Zeigefinger und Mittelfinger der linken Hand in dieser Reihenfolge als Koordinatensystem, so entsprechen ihnen die Standard-Vektoren  $\vec{e}_1, \vec{e}_2$  und  $\vec{e}_3$ . Als Vektorprodukt der ersten beiden Richtungen ergibt sich also  $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$ . Das Vektorprodukt im Anschauungsraum ist aber durch die Rechte-Hand-Regel gegeben, es zeigt in die entgegengesetzte Richtung. Zur Berechnung des Vektorprodukts müssen wir als Koordinaten ein Rechtssystem wählen: nur dann sind die Orientierung durch die Rechte-Hand-Regel und die Orientierung des  $\mathbb{R}^3$  konsistent.

## 5 Die komplexen Zahlen

Wir beginnen mit einer anderen Schreibweise für Punkte des  $\mathbb{R}^2$ . Und zwar setzen wir  $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$  und  $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = i$ , und haben dann die allgemeine Darstellung

$$z = x + iy \quad \text{für alle } z = (x, y) \in \mathbb{R}^2.$$

In diesem Zusammenhang heißen die Punkte des  $\mathbb{R}^2$  komplexe Zahlen (Symbol  $\mathbb{C}$ ). Realteil und Imaginärteil einer komplexen Zahl sind gegeben durch

$$\operatorname{Re} z = x \quad \operatorname{Im} z = y \quad \text{für } z = x + iy.$$

Die Addition in  $\mathbb{C}$  ist die übliche komponentenweise Addition, also

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2).$$

Für die Multiplikation setzen wir  $i^2 = -1$  und multiplizieren aus:

$$(x_1 + iy_1)(x_2 + iy_2) = x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1).$$

Um dies zu veranschaulichen, schreiben wir die komplexen Zahlen in Polarkoordinaten: zu jedem  $z \in \mathbb{C}$  mit  $z \neq 0$  gibt es ein  $r > 0$  und ein  $\varphi \in \mathbb{R}$  mit

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

Dabei ist  $r = |z| = \sqrt{x^2 + y^2}$ , und  $\varphi \in \mathbb{R}$  ist der Winkel, den  $z$  mit der positiven  $x$ -Achse einschließt (Bild). Da die Funktionen  $\cos, \sin$  die Periode  $2\pi$  haben, ist der Winkel nur bis auf ganzzahlige Vielfache von  $2\pi$  eindeutig bestimmt. Er ist eindeutig, wenn wir  $\varphi \in [0, 2\pi)$  verlangen.

Für die komplexen Zahlen  $\mathbb{C}$  gelten dieselben Rechenregeln wie für die reellen Zahlen. Bezüglich der Addition ist  $0 = 0 + i0$  das neutrale Element, und  $-z = -x - iy$  das zu

$z = x + iy$  inverse Element. Kommutativgesetz und Assoziativgesetz der Addition sind klar, denn diese ist einfach die komponentenweise Addition.

Das neutrale Element der Multiplikation ist  $1 = 1 + i0$ , denn

$$z \cdot 1 = (x + iy)(1 + i0) = x + iy.$$

Wir nennen  $\bar{z} = x - iy$  die zu  $z = x + iy$  konjugiert komplexe Zahl. Wegen  $i^2 = -1$  gilt

$$z\bar{z} = (x + iy)(x - iy) = x^2 - i^2y^2 + i(yx - xy) = x^2 + y^2 = |z|^2.$$

Damit können wir das zu  $z \neq 0$  inverse Element angeben, und zwar ist

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \frac{x - iy}{x^2 + y^2}, \quad \text{denn } z \frac{\bar{z}}{|z|^2} = \frac{z\bar{z}}{|z|^2} = 1.$$

In Polarkoordinaten lautet nun die Multiplikation (Additionstheoreme für  $\cos, \sin$ )

$$\begin{aligned} z_1 z_2 &= r_1 r_2 (\cos \varphi_1 + i \sin \varphi_1)(\cos \varphi_2 + i \sin \varphi_2) \\ &= r_1 r_2 (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)) \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)). \end{aligned}$$

Das Kommutativgesetz ist damit offensichtlich, und durch Multiplikation mit  $z_3 = r_3(\cos \varphi_3 + i \sin \varphi_3)$  folgt weiter

$$(z_1 z_2) z_3 = r_1 r_2 r_3 (\cos(\varphi_1 + \varphi_2 + \varphi_3) + i \sin(\varphi_1 + \varphi_2 + \varphi_3)).$$

Für  $z_1(z_2 z_3)$  kommt dasselbe raus, also gilt das Assoziativgesetz. Schließlich rechnen wir das Distributivgesetz nach:

$$\begin{aligned} (a + ib)((x_1 + iy_1) + (x_2 + iy_2)) &= (a + ib)(x_1 + x_2 + i(y_1 + y_2)) \\ &= a(x_1 + x_2) - b(y_1 + y_2) + i(b(x_1 + x_2) + a(y_1 + y_2)) \\ &= ax_1 - by_1 + i(bx_1 + ay_1) + ax_2 - by_2 + i(bx_2 + ay_2) \\ &= (a + ib)(x_1 + iy_1) + (a + ib)(x_2 + iy_2). \end{aligned}$$

Wir notieren im Vorbeigehen folgende nützliche Formeln, die sich leicht verifizieren lassen. Für die zweite Gleichung in (1) und für (3) kann man die Polardarstellung von  $z_{1,2}$  verwenden.

$$(1) \quad \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2,$$

$$(2) \quad \text{Für } z = x + iy \text{ ist } \operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \text{ und } \operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$$

$$(3) \quad |z_1 z_2| = |z_1| |z_2|.$$

Was ist der Gewinn, den wir von den komplexen Zahlen haben? Allgemein gesagt, sind mit den komplexen Zahlen Gleichungen lösbar, die mit reellen Zahlen nicht lösbar sind. Zum Beispiel gibt es keine reelle Lösung von  $x^2 + q = 0$ , wenn  $q > 0$  ist. In  $\mathbb{C}$  haben wir dagegen die beiden Lösungen  $z_{1,2} = \pm i\sqrt{q}$ <sup>1</sup>. Jede Gleichung  $x^2 + px + q = 0$  mit  $p, q \in \mathbb{R}$  kann auf das

<sup>1</sup>Mit  $\sqrt{x}$  ist immer die nichtnegative Wurzel gemeint

Ziehen einer Wurzel reduziert werden, und hat damit ebenfalls Lösungen. Substituiere dazu  $x = y + a$  und berechne

$$0 = (y + a)^2 + p(y + a) + q = y^2 + (2a + p)y + a^2 + pa + q.$$

Durch Wahl von  $a = -p/2$  ergibt sich

$$y^2 = \frac{p^2}{4} - q.$$

Diese Gleichung hat die Lösungen

$$y = \begin{cases} \pm \sqrt{\frac{p^2}{4} - q} & \text{falls } \frac{p^2}{4} - q > 0 \\ \pm i \sqrt{-(\frac{p^2}{4} - q)} & \text{falls } \frac{p^2}{4} - q < 0 \\ 0 & \text{falls } \frac{p^2}{4} - q = 0. \end{cases}$$

Durch Rückeinsetzen  $x = y - p/2$  ergeben sich die jeweiligen Lösungen für  $x$ .

Komplexe Zahlen traten zuerst in der italienischen Renaissance in der ersten Hälfte des 16. Jahrhunderts auf, bei der Lösung von quadratischen und kubischen Gleichungen. Die Bezeichnung  $i = \sqrt{-1}$  stammt von Euler (1777), die Bezeichnung *komplexe Zahl* von Gauß (1831). Er hat auch die Interpretation als Punkte in der Ebene eingeführt.

# Kapitel 2

## Funktionen, Grenzwerte, Stetigkeit

### 1 Reelle Funktionen

Eine reelle Funktion einer Variablen ist eine Abbildung

$$f : D \rightarrow \mathbb{R}, x \mapsto f(x) \quad \text{wobei } D \subset \mathbb{R}.$$

Der Definitionsbereich  $D$  kann zum Beispiel ganz  $\mathbb{R}$  oder ein Intervall sein. Die Bezeichnung *reelle Funktion* bezieht sich darauf, dass die Werte der Funktion reelle Zahlen sind. Standardbeispiele von Funktionen sind:

$$\begin{array}{ll} \text{lineare Funktionen:} & f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = ax + b. \\ \text{quadratische Funktionen:} & f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = ax^2 + bx + c. \\ \text{Wurzelfunktion:} & f : [0, \infty) \rightarrow \mathbb{R}, f(x) = \sqrt{x}. \end{array}$$

Die Wurzelfunktion kann für  $x < 0$  nicht als reelle Funktion definiert werden. Die Rechenoperationen von  $\mathbb{R}$  lassen sich für Funktionen definieren, genauer setzen wir für  $f, g : D \rightarrow \mathbb{R}$

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

Sind  $f : D \rightarrow \mathbb{R}$  und  $g : E \rightarrow \mathbb{R}$  mit  $f(D) \subset E$ , so ist auch die Verkettung definiert durch

$$g \circ f : D \rightarrow \mathbb{R}, (g \circ f)(x) = g(f(x)).$$

In vielen Anwendungen interessiert man sich dafür, wo die Maxima und Minima der Funktion liegen (Extremwerte), und allgemeiner in welchen Bereichen die Funktion anwächst beziehungsweise fällt. Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt

$$\begin{array}{ll} \text{monoton wachsend:} & x_2 > x_1 \Rightarrow f(x_2) \geq f(x_1) \\ \text{streng monoton wachsend:} & x_2 > x_1 \Rightarrow f(x_2) > f(x_1) \\ \text{monoton fallend:} & x_2 > x_1 \Rightarrow f(x_2) \leq f(x_1) \\ \text{streng monoton fallend:} & x_2 > x_1 \Rightarrow f(x_2) < f(x_1). \end{array}$$

Zum Beispiel hat die quadratische Funktion  $f(x) = x^2 + px + q$  im Punkt  $x = -p/2$  ein Minimum. Auf dem Intervall  $[-p/2, \infty)$  ist sie streng monoton wachsend, auf  $(-\infty, -p/2]$

streng monoton fallend. Ein weiterer Begriff bezieht sich auf die Symmetrie einer Funktion:  $f : D \rightarrow \mathbb{R}$  heißt

$$\begin{aligned} \text{gerade:} & \quad f(-x) = f(x) \text{ für alle } x \in D, \\ \text{ungerade:} & \quad f(-x) = -f(x) \text{ für alle } x \in D. \end{aligned}$$

Für diese Eigenschaft muss natürlich  $D$  symmetrisch zum Ursprung liegen, zum Beispiel  $D = (-a, a)$ , sonst macht es keinen Sinn. Die quadratische Funktion  $f(x) = x^2 + px + q$  ist genau dann gerade, wenn  $p = 0$  ist, und niemals ungerade. Denn

$$\frac{f(x) - f(-x)}{2} = px \quad \frac{f(x) + f(-x)}{2} = x^2 + q.$$

## 2 Polynome und rationale Funktionen

**Definition 2.1** Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt (reelles) Polynom vom Grad  $n \in \mathbb{N}_0$ , wenn es  $a_0, a_1, \dots, a_n \in \mathbb{R}$  gibt mit  $a_n \neq 0$ , so dass

$$f(x) = a_0 + a_1x + \dots + a_nx^n \quad \text{für alle } x \in \mathbb{R}. \quad (2.1)$$

Die  $a_i \in \mathbb{R}$  heißen Koeffizienten des Polynoms, und  $a_n$  heißt Leitkoeffizient. Ein Polynom vom Grad Null ist nach Definition konstant, und nicht die Nullfunktion.

**Lemma 2.1 (Abspalten von Linearfaktoren)** Hat ein Polynom  $f(x)$  vom Grad  $n$  eine Nullstelle  $\lambda \in \mathbb{R}$ , also  $f(\lambda) = 0$ , so gibt es ein Polynom  $g(x)$  vom Grad  $n - 1$  mit

$$f(x) = (x - \lambda)g(x) \quad \text{für alle } x \in \mathbb{R}.$$

BEWEIS: Sei  $f(x) = a_0 + a_1x + \dots + a_nx^n$  mit  $a_n \neq 0$ . Im Fall  $\lambda = 0$  folgt  $0 = f(0) = a_0$ , und

$$f(x) = x(a_1 + \dots + a_nx^{n-1}) = xg(x) \quad \text{für alle } x \in \mathbb{R}.$$

Sei nun  $f(\lambda) = 0$  für irgendein  $\lambda \in \mathbb{R}$ . Betrachte dann

$$\tilde{f}(x) = f(x + \lambda) = a_0 + a_1(x + \lambda) + \dots + a_n(x + \lambda)^n.$$

Durch Auflösen der Klammern und Ordnen nach Potenzen von  $x$  sehen wir, dass  $\tilde{f}(x)$  Polynom vom Grad  $n$  ist mit Leitkoeffizient  $a_n$ . Aber  $\tilde{f}(0) = f(\lambda) = 0$ , und wie gezeigt gibt es ein Polynom  $\tilde{g}$  vom Grad  $n - 1$  mit  $\tilde{f}(x) = x\tilde{g}(x)$ , beziehungsweise

$$f(x) = \tilde{f}(x - \lambda) = (x - \lambda)\tilde{g}(x - \lambda) \quad \text{für alle } x \in \mathbb{R}.$$

Nun ist  $\tilde{g}(x - \lambda)$  ein Polynom vom Grad  $n - 1$ , indem wir wieder Ausmultiplizieren und Umordnen. Damit ist die Behauptung bewiesen.  $\square$

**Lemma 2.2 (Zahl der Nullstellen)** Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  ein Polynom vom Grad  $n \in \mathbb{N}_0$ . Dann hat  $f$  höchstens  $n$  verschiedene Nullstellen.

BEWEIS: Wir führen Induktion über  $n \in \mathbb{N}_0$ . Für  $n = 0$  ist  $f(x) = a_0$  für alle  $x \in \mathbb{R}$ , wobei  $a_0 \neq 0$ , also hat  $f$  keine Nullstelle. Ist  $f$  Polynom vom Grad  $n \in \mathbb{N}$ , so hat entweder  $f$  keine Nullstelle oder es gilt nach Lemma 2.1  $f(x) = (x - \lambda)g(x)$  für alle  $x \in \mathbb{R}$ , mit einem Polynom  $g$  vom Grad  $n - 1$ . Nach Induktion hat  $g$  höchstens  $n - 1$  Nullstellen, also  $f$  höchstens  $n$  Nullstellen.  $\square$

**Lemma 2.3 (Koeffizientenvergleich)** Seien  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  Polynome vom Grad  $m$  bzw.  $n$ , das heißt es gilt mit  $a_m, b_n \neq 0$

$$f(x) = \sum_{i=0}^m a_i x^i \quad \text{und} \quad g(x) = \sum_{i=0}^n b_i x^i \quad \text{für alle } x \in \mathbb{R}.$$

Sind  $f(x)$  und  $g(x)$  an mehr als  $\max(m, n)$  Stellen gleich, so ist  $m = n$  und  $a_i = b_i$  für  $i = 0, \dots, n$ .

BEWEIS: Ist  $m \neq n$  oder  $a_i \neq b_i$  für ein  $i$ , so ist  $f - g$  Polynom vom Grad höchstens  $\max(m, n)$ , und hat nach Lemma 2.2 höchstens  $\max(m, n)$  Nullstellen.  $\square$

All das hilft uns nicht weiter, wenn ein Polynom einfach keine Nullstellen hat, zum Beispiel  $f(x) = x^2 + 1$ . Um die Sache wirklich zu verstehen, müssen wir ins Komplexe gehen. Ein komplexes Polynom vom Grad  $n$  hat die Form

$$f : \mathbb{C} \rightarrow \mathbb{C}, f(z) = a_0 + a_1 z + \dots + a_n z^n \quad \text{wobei } a_i \in \mathbb{C}, a_n \neq 0.$$

Das Abspalten von Linearfaktoren und der Koeffizientenvergleich gelten in  $\mathbb{C}$  ganz analog, weil nur die gemeinsamen Rechenregeln benutzt wurden. Insbesondere hat auch ein komplexes Polynom höchstens  $n$  Nullstellen. Im Unterschied zum Reellen gilt aber der

**Satz 2.1 (Fundamentalsatz der Algebra)** Jedes komplexe Polynom vom Grad  $n \geq 1$  hat mindestens eine Nullstelle  $\lambda \in \mathbb{C}$ .

Es ist offenbar möglich, dass sich ein Linearfaktor mehrfach von einem Polynom  $f(z)$  abspalten lässt. Wir nennen  $\lambda$  eine Nullstelle der Vielfachheit  $k \in \mathbb{N}$ , wenn  $f(z) = (z - \lambda)^k g(z)$  für ein Polynom  $g(z)$  mit  $g(\lambda) \neq 0$ .

**Folgerung 2.1 (Polynomfaktorisierung)** Ein Polynom  $f : \mathbb{C} \rightarrow \mathbb{C}$  vom Grad  $n \geq 1$  hat eine Zerlegung

$$f(z) = a_n \prod_{k=1}^K (z - \lambda_k)^{n_k} \quad \text{für alle } z \in \mathbb{C}.$$

Dabei sind  $\lambda_1, \dots, \lambda_K \in \mathbb{C}$  die Nullstellen mit den zugehörigen Vielfachheiten  $n_k \in \mathbb{N}$ , wobei  $n = n_1 + \dots + n_K$ , und  $a_n \in \mathbb{C} \setminus \{0\}$  ist der Leitkoeffizient von  $f(z)$ .

BEWEIS: Nach dem Fundamentalsatz der Algebra hat  $f(z)$  eine Nullstelle  $\lambda_1 \in \mathbb{C}$ . Durch Abspalten folgt  $f(z) = (z - \lambda_1) f_1(z)$ , wobei  $f_1(z)$  komplexes Polynom vom Grad  $n - 1$  ist. Nun hat  $f_1(z)$  wieder eine Nullstelle  $\lambda_2 \in \mathbb{C}$ , und so weiter. Der Prozess stoppt genau nach  $n$  Schritten, denn dann hat das Restpolynom den Grad Null, ist also konstant.  $\square$

Wir können nun auf  $\mathbb{R}$  zurückkommen. Jedes reelle Polynom  $p(x) = a_0 + a_1 x + \dots + a_n x^n$ , das heißt  $a_i \in \mathbb{R}$ , kann nämlich als komplexes Polynom aufgefasst werden, indem wir für  $x$  auch komplexe Zahlen  $z$  einsetzen. Damit wird die reelle Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  zu einer komplexen Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$  fortgesetzt. Das Endergebnis lautet so.

**Folgerung 2.2 (Polynomfaktorisierung in  $\mathbb{R}$ )** Jedes reelle Polynom  $f : \mathbb{R} \rightarrow \mathbb{R}$  vom Grad  $n \geq 1$  zerfällt in lineare und quadratische Faktoren. Genauer gilt

$$f(x) = a_n \prod_{i=1}^I (x - \lambda_i)^{\ell_i} \prod_{j=1}^J (x^2 - 2\alpha_j x + \alpha_j^2 + \beta_j^2)^{m_j} \quad \text{für alle } x \in \mathbb{R}.$$

Dabei sind die  $\lambda_i$  die Nullstellen in  $\mathbb{R}$  und die  $\alpha_j \pm i\beta_j$  die Nullstellen in  $\mathbb{C} \setminus \mathbb{R}$ , mit jeweiligen Vielfachheiten  $\ell_i$  bzw.  $m_j$ , also  $n = \ell_1 + \dots + \ell_I + 2(m_1 + \dots + m_J)$ .

BEWEIS: Wir können annehmen, dass  $f(x)$  keine reellen Nullstellen hat, sonst spalten wir die zugehörigen Linearfaktoren ab; diese liefern dann das erste Produkt. Weiter gilt: ist  $\lambda = \alpha + i\beta$  eine nicht-reelle Nullstelle, so auch  $\bar{\lambda} = \alpha - i\beta$ . Wegen  $a_i \in \mathbb{R}$  gilt nämlich

$$0 = \overline{f(\lambda)} = \overline{a_0 + a_1\lambda + \dots + a_n\lambda^n} = a_0 + a_1\bar{\lambda} + \dots + a_n\bar{\lambda}^n = f(\bar{\lambda}).$$

Wir können nun in  $\mathbb{C}$  nacheinander  $z - \lambda$  und  $z - \bar{\lambda}$  als Faktoren abspalten, und erhalten ein Polynom  $g(z)$  vom Grad  $n - 2$  mit

$$f(z) = (z - \lambda)(z - \bar{\lambda})g(z) = (z^2 - 2\alpha z + \alpha^2 + \beta^2)g(z).$$

Jetzt machen wir mit dem Restpolynom  $g(z)$  weiter und erhalten schließlich die gewünschte Zerlegung in quadratische Faktoren. Ein Detail fehlt: es ist zu begründen, dass  $g(z)$  wieder reelle Koeffizienten hat! Sei

$$g(z) = b_0 + b_1z + \dots + b_{n-2}z^{n-2}, \quad \text{mit zunächst } b_i \in \mathbb{C}.$$

Wir setzen in  $f(z)$  nun  $x \in \mathbb{R}$  als Variable ein und erhalten

$$\begin{aligned} 0 &= \operatorname{Im} f(x) \quad (\text{da } f \text{ reelle Koeffizienten hat}) \\ &= \operatorname{Im} \left( (x^2 - 2\alpha x + \alpha^2 + \beta^2) g(x) \right) \\ &= (x^2 - 2\alpha x + \alpha^2 + \beta^2) \operatorname{Im} g(x) \\ &= (x^2 - 2\alpha x + \alpha^2 + \beta^2) \sum_{i=0}^{n-2} (\operatorname{Im} b_i) x^i. \end{aligned}$$

Die linke Klammer ist für  $x \in \mathbb{R}$  niemals Null. Also verschwindet die Summe für alle  $x \in \mathbb{R}$ , und es folgt  $\operatorname{Im} b_i = 0$  für alle  $i = 0, \dots, n - 2$  aus Lemma 2.2.  $\square$

Die Existenz einer Nullstelle ist in  $\mathbb{C}$  durch Satz 3.1 gesichert. Dieser Satz liefert aber keinen Anhaltspunkt, wie die Nullstellen tatsächlich berechnet werden sollen. Mit Substitutionen und Ziehen von  $k$ -ten Wurzeln kommt man ab Grad  $n \geq 5$  im allgemeinen nicht zum Ziel (Abel 1825). Deshalb spielen numerische Lösungsverfahren eine große Rolle, durch die die Nullstelle approximativ bestimmt wird. Darauf kommen wir zurück, wenn wir die Analysis weiter entwickelt haben.

Eine rationale Funktion  $f$  ist definiert als Quotient zweier Polynome. Seien genauer  $p(x)$  und  $q(x)$  reelle Polynome vom Grad  $m$  bzw.  $n$ , und  $N_q$  sei die endliche Menge der Nullstellen von  $q$ . Dann ist

$$f : \mathbb{R} \setminus N_q \rightarrow \mathbb{R}, \quad f(x) = \frac{p(x)}{q(x)}.$$

Es ist praktisch, rationale Funktionen folgendermaßen zu zerlegen.

**Lemma 2.4** Seien  $p(x) = a_m x^m + \dots + a_0$  und  $q(x) = b_n x^n + \dots + b_0$  Polynome mit  $m \geq n$ . Dann hat gibt es eindeutig bestimmte Polynome  $g(x)$  und  $r(x)$ , eventuell  $r(x)$  die Nullfunktion, mit

$$f(x) = g(x) + \frac{r(x)}{q(x)} \quad \text{für alle } x \in \mathbb{R} \setminus N_q.$$

BEWEIS: Durch Division mit Rest für Polynome. Wir schreiben

$$\frac{p(x)}{q(x)} = \frac{a_n}{b_m} x^{m-n} + \frac{p_1(x)}{q(x)} \quad \text{mit } p_1(x) = p(x) - \frac{a_n}{b_m} x^{m-n} q(x).$$

Es gilt  $\text{grad } p_1 < m$ . Im Fall  $m = n$  gilt die Behauptung also mit  $r(x) = p_1(x)$ . Für  $m > n$  können wir per Induktion annehmen, dass

$$\frac{p_1(x)}{q(x)} = g_1(x) + \frac{r_1(x)}{q(x)} \quad \text{mit } \text{grad } r_1 < n,$$

eventuell  $r_1 \equiv 0$ . Es folgt die gewünschte Zerlegung

$$\frac{p(x)}{q(x)} = \frac{a_n}{b_m} x^{m-n} + g_1(x) + \frac{r_1(x)}{q(x)}.$$

Zur Eindeutigkeit: angenommen die Zerlegung gilt mit  $g_1, r_1$  und  $g_2, r_2$ . Dann folgt durch Subtraktion und Multiplikation mit  $q(x)$

$$(g_1(x) - g_2(x))q(x) + r_1(x) - r_2(x) = 0 \quad \text{für alle } x \in \mathbb{R} \setminus N_q.$$

Wäre  $g_1 - g_2$  nicht Null, so ist die linke Seite ein Polynom vom Grad mindestens  $n$ , aber mit unendlich vielen Nullstellen. Das kann nicht sein wegen Lemma 2.2. Analog sehen wir, dass die Polynome  $r_{1,2}$  gleich sind.  $\square$

**Beispiel 2.1** Hier ein Beispiel für die Polynomdivision:

$$\begin{array}{rclcl} x^4 - x^3 + x^2 - x + 1 & : & x^2 + 2x & = & x^2 & \text{Rest } -3x^3 + x^2 - x + 1 \\ -3x^3 + x^2 - x + 1 & : & x^2 + 2x & = & -3x & \text{Rest } 7x^2 - x + 1 \\ 7x^2 - x + 1 & : & x^2 + 2x & = & 7 & \text{Rest } -15x + 1 \end{array}$$

Also haben wir hier

$$\frac{x^4 - x^3 - x + 1}{x^2 + 2x} = x^2 - 3x + 7 - \frac{15x - 1}{x^2 + 2x}.$$

Die Ausnahmestellen  $\lambda \in N_q$  in der Definition von  $f(x)$  sind natürlich von Interesse. Sei  $\lambda$  eine  $m$ -fache Nullstelle von  $q(x)$  und eine  $k$ -fache Nullstelle von  $p(x)$ . Im Fall  $p(x) \neq 0$  ist  $k = 0$ . Es gibt nun Polynome  $p_1(x), q_1(x)$  mit  $p_1(\lambda), q_1(\lambda) \neq 0$  und

$$f(x) = \frac{p(x)}{q(x)} = \frac{(x - \lambda)^k p_1(x)}{(x - \lambda)^m q_1(x)} = (x - \lambda)^{k-m} \frac{p_1(x)}{q_1(x)} \quad \text{für } x \in \mathbb{R} \setminus N_q.$$

Abhängig von dem Exponenten  $k - m$  sind zwei Fälle zu unterscheiden:

- (1) Ist  $k \geq m$ , so kann  $f(x)$  auch im Punkt  $\lambda$  sinnvoll definiert werden, und zwar durch

$$f(x) = \begin{cases} \frac{p_1(x)}{q_1(x)} & \text{falls } k = m, \\ 0 & \text{falls } k > m. \end{cases}$$

Wir nennen  $\lambda$  eine *hebbare Singularität* von  $f(x)$ .

- (2) Ist  $k < m$ , so nennen wir  $\lambda$  eine *Polstelle* von  $f(x)$ . Das Verhalten von  $|f(x)|$  in der Polstelle wird im wesentlichen durch den Faktor  $|x - \lambda|^{k-m}$  bestimmt, der gegen Unendlich geht.

### 3 Kreisfunktionen

Die trigonometrischen Funktionen sind bereits mehrfach aufgetreten. Wir wollen sie hier mit ihren Eigenschaften nochmal ausführlich besprechen. Sei  $D_x : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Drehung um den Winkel  $x \in \mathbb{R}$ , gemessen in Bogenmaß. Dabei drehen wir im mathematisch positiven Sinn (das heißt entgegen dem Uhrzeigersinn), wenn  $x \geq 0$ , im mathematisch negativen Sinn für  $x < 0$ . Es gilt (vgl. Abschnitt I.4 und Bild)

$$D_x \vec{e}_1 = \begin{pmatrix} \cos x \\ \sin x \end{pmatrix} \quad D_x \vec{e}_2 = \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}.$$

Aus der Interpretation am Einheitskreis lassen sich die Eigenschaften von  $\cos, \sin : \mathbb{R} \rightarrow \mathbb{R}$  leicht ablesen. Als erstes haben wir

$$\cos^2 x + \sin^2 x = 1.$$

Beachte  $\cos^2 x = (\cos x)^2$ . Es folgt auch  $-1 \leq \cos x, \sin x \leq 1$ . Weiter

$$\begin{aligned} \cos(-x) &= \cos x & (\cos \text{ ist gerade}) \\ \sin(-x) &= -\sin x & (\sin \text{ ist ungerade}) \end{aligned}$$

Drehung um den Vollwinkel  $2\pi$  ist die Identität, daher sind die Funktionen  $2\pi$ -periodisch:

$$\cos(x + 2k\pi) = \cos x \quad \sin(x + 2k\pi) = \sin x \quad \text{für alle } k \in \mathbb{Z}.$$

Die Nullstellen der Funktionen sind

$$\begin{aligned} \sin x = 0 &\Leftrightarrow x \in \mathbb{Z}\pi \\ \cos x = 0 &\Leftrightarrow x \in (2\mathbb{Z} + 1)\frac{\pi}{2} \end{aligned}$$

Um die Additionstheoreme zu zeigen, brauchen wir, dass die Drehung  $D_x$  eine lineare Abbildung ist. Das bedeutet, dass sie sich mit der Vektoraddition und der Skalarmultiplikation verträgt, genauer behaupten wir für beliebige  $\vec{v}, \vec{w} \in \mathbb{R}^2$  und  $\lambda \in \mathbb{R}$ :

$$D_x(\vec{v} + \vec{w}) = D_x \vec{v} + D_x \vec{w} \quad D_x(\lambda \vec{v}) = \lambda D_x \vec{v}.$$

Für die erste Eigenschaft beachten wir, dass  $\vec{v} + \vec{w}$  der vierte Eckpunkt in dem von  $\vec{0}, \vec{v}, \vec{w}$  gebildeten Parallelogramm ist. Drehung mit  $D_x$  ergibt daher den vierten Eckpunkt in dem gedrehten Parallelogramm, das von  $\vec{0}, D_x \vec{v}, D_x \vec{w}$  gebildet wird, also den Punkt  $D_x \vec{v} + D_x \vec{w}$ . Für  $\lambda > 0$  hat der Vektor  $D_x(\lambda \vec{v})$  dieselbe Richtung wie  $\lambda D_x \vec{v}$ . Da die Drehung längentreu ist, sind die Vektoren gleich. Schließlich kann man für  $\lambda = -1$  argumentieren, dass  $D_x(-\vec{v}) = D_x D_\pi \vec{v} = D_\pi D_x \vec{v} = -D_x \vec{v}$ .

**Satz 3.1 (Additionstheoreme)** *Für alle  $x, y \in \mathbb{R}$  gilt*

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

BEWEIS: Wir berechnen Mit  $D_{x+y} = D_y \circ D_x$  berechnen wir

$$\begin{aligned}
 \begin{pmatrix} \cos(x+y) \\ \sin(x+y) \end{pmatrix} &= D_{x+y} \vec{e}_1 \\
 &= D_y(D_x \vec{e}_1) \quad (\text{da } D_{x+y} = D_y \circ D_x) \\
 &= D_y((\cos x) \vec{e}_1 + (\sin x) \vec{e}_2) \\
 &= (\cos x) D_y \vec{e}_1 + (\sin x) D_y \vec{e}_2 \quad (\text{wegen } D_y \text{ linear}) \\
 &= \cos x \begin{pmatrix} \cos y \\ \sin y \end{pmatrix} + \sin x \begin{pmatrix} -\sin y \\ \cos y \end{pmatrix} \\
 &= \begin{pmatrix} \cos x \cos y - \sin x \sin y \\ \sin x \cos y + \cos x \sin y \end{pmatrix}.
 \end{aligned}$$

□

Im Spezialfall  $y = \pm \frac{\pi}{2}$  ist

$$\sin\left(x + \frac{\pi}{2}\right) = \cos x \quad \text{bzw.} \quad \cos\left(x - \frac{\pi}{2}\right) = \sin x.$$

Es lassen sich viele weitere nützliche Formeln aus den Additionstheoremen herleiten (siehe Übungsaufgaben).

**Definition 3.1** Die Tangens- bzw. Cotangensfunktion ist definiert durch

$$\begin{aligned}
 \tan x &= \frac{\sin x}{\cos x} \quad \text{für } x \neq (2\mathbb{Z} + 1)\frac{\pi}{2} \\
 \cot x &= \frac{\cos x}{\sin x} \quad \text{für } x \neq \mathbb{Z}\pi.
 \end{aligned}$$

Für die Funktionen  $\tan$  und  $\cot$  ergeben sich folgende Eigenschaften:

$$\begin{aligned}
 \tan(-x) &= -\tan x \quad (\tan \text{ ist ungerade}) \\
 \cot(-x) &= -\cot x \quad (\cot \text{ ist ungerade}) \\
 \tan(x + \pi) &= \tan x \quad (\tan \text{ hat Periode } \pi) \\
 \cot(x + \pi) &= \cot x \quad (\cot \text{ hat Periode } \pi)
 \end{aligned}$$

Wir kommen an diesem Punkt auf die Polardarstellung komplexer Zahlen zurück. Jeder Zahl  $z \in \mathbb{C}$ ,  $z \neq 0$ , hat eine Darstellung

$$z = r(\cos \varphi + i \sin \varphi) \quad \text{mit } r > 0, \varphi \in \mathbb{R}.$$

Wir führen folgende Abkürzung ein:

$$e^{ix} = \cos x + i \sin x \quad \text{für } x \in \mathbb{R} \quad (\text{Eulersche Formel}).$$

Was soll diese Notation? Die aus der Schule bekannte Exponentialfunktion erfüllt folgendes Funktionalgesetz, das in der Schule im Rahmen der Potenzrechnung hergeleitet wird:

$$e^{\lambda(x+y)} = e^{\lambda x + \lambda y} = e^{\lambda x} e^{\lambda y} \quad \text{wobei } \lambda \in \mathbb{R}.$$

Genau dieses Gesetz gilt auch für  $e^{ix}$ , nur dass  $\lambda \in \mathbb{R}$  durch die komplexe Zahl  $i \in \mathbb{C}$  ersetzt ist, und zwar folgt es aus den Additionstheoremen:

$$\begin{aligned} e^{i(x+y)} &= \cos(x+y) + i \sin(x+y) \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y) \\ &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= e^{ix} e^{iy}. \end{aligned}$$

Umgekehrt lassen sich die Additionstheoreme aus der Regel  $e^{i(x+y)} = e^{ix} e^{iy}$  mühelos herleiten. Mit Schulwissen können wir zusätzlich die Ableitungen der Funktionen  $e^{\lambda x}$  ( $\lambda \in \mathbb{R}$ ) und  $e^{ix}$  bei  $x = 0$  vergleichen. Es gilt

$$\begin{aligned} (e^{\lambda x})'(0) &= \lambda e^{\lambda x}|_{x=0} = \lambda \\ (e^{ix})'(0) &= \cos'(0) + i \sin'(0) = i. \end{aligned}$$

Damit ist die als Notation eingeführte Eulersche Formel gut begründet. Die Polardarstellung hat die endgültige Form

$$z = r e^{i\varphi} \quad \text{für } z \in \mathbb{C} \setminus \{0\}.$$

**Satz 3.2 (de Moivre)** Für  $x, y \in \mathbb{R}$  gelten die Formeln

$$\begin{aligned} e^{i(x+y)} &= e^{ix} e^{iy} \quad (\text{Funktionalgleichung}) \\ \overline{e^{ix}} &= e^{-ix} = \frac{1}{e^{ix}} \\ e^{inx} &= (e^{ix})^n \quad \text{für } n \in \mathbb{N}. \end{aligned}$$

BEWEIS: Die Funktionalgleichung wurde oben hergeleitet. Die zweite Gleichung gilt wegen

$$\overline{e^{ix}} = \overline{\cos x + i \sin x} = \cos x - i \sin x = e^{-ix} \quad \text{und} \quad e^{ix} e^{-ix} = e^{i0} = 1.$$

Die dritte Formel ist klar für  $n = 1$ , und per Induktion ergibt sich

$$e^{i(n+1)x} = e^{inx+ix} = e^{inx} e^{ix} = (e^{ix})^n e^{ix} = (e^{ix})^{n+1}.$$

□

Zur Berechnung des Winkels zwischen Vektoren sowie der Polardarstellung komplexer Zahlen wird die Umkehrfunktion des Kosinus gebraucht. Dazu müssen wir ein Intervall auswählen, auf dem der Kosinus injektiv ist, üblicherweise

$$\cos : [0, \pi] \rightarrow [-1, 1].$$

Die Funktion  $\cos$  ist auf  $[0, \pi]$  streng monoton fallend und damit injektiv:

$$0 \leq x < y \leq \pi \quad \Rightarrow \quad \cos x > \cos y.$$

Das ist anschaulich am Einheitskreis klar, rigoros rechnen wir

$$e^{iy} - e^{ix} = e^{i\frac{y+x}{2}} \left( e^{i\frac{y-x}{2}} - e^{-i\frac{y-x}{2}} \right) = 2ie^{i\frac{y+x}{2}} \sin \frac{y-x}{2}.$$

Bilden wir die Realteile, so folgt wegen  $\sin t > 0$  für  $t \in (0, \pi)$

$$\cos y - \cos x = -2 \underbrace{\sin \frac{y+x}{2}}_{\in(0,\pi)} \underbrace{\sin \frac{y-x}{2}}_{\in(0,\pi)} < 0.$$

Wir definieren nun die Umkehrfunktion

$$\arccos : [-1, 1] \rightarrow [0, \pi], \quad \cos(\arccos t) = t.$$

Natürlich gilt auch  $\arccos(\cos x) = x$  für  $x \in [0, \pi]$ . Damit können wir die Polarkoordinaten von  $z = x + iy \neq 0$  angeben:

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

Die trigonometrischen Funktionen sind von Bedeutung für die Beschreibung von Schwingungen und periodischen Prozessen.  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt periodisch mit Periode  $T > 0$ , falls

$$f(t + T) = f(t) \quad \text{für alle } t \in \mathbb{R}.$$

Ist  $T > 0$  Periode von  $f$ , so auch alle Zahlen  $kT$  mit  $k \in \mathbb{Z} \setminus \{0\}$ . Wenn man von *der* Periode einer Funktion spricht, so meint man die kleinstmögliche Periode. Eine harmonische Schwingung, abhängig von der Zeit  $t > 0$ , ist von der Form

$$x : \mathbb{R} \rightarrow \mathbb{R}, \quad x(t) = A \cos(\omega t + \alpha).$$

Die Schwingungsdauer, also die Periode der Schwingung, ist

$$T = \frac{2\pi}{\omega}.$$

Weitere Größen sind die Frequenz  $\nu = \frac{1}{T}$ , die Amplitude  $A > 0$  und der Phasenwinkel  $\alpha \in \mathbb{R}$ . Die Zahl  $\omega$  bezeichnet man auch als Kreisfrequenz.

Die Überlagerung (Superposition) zweier Schwingungen  $x_1(t)$ ,  $x_2(t)$  ist durch  $x(t) = x_1(t) + x_2(t)$  gegeben. Im allgemeinen ist eine solche Überlagerung nicht periodisch. Ausnahme ist der Fall

$$\frac{T_2}{T_1} \in \mathbb{Q}, \quad \text{also } \frac{T_2}{T_1} = \frac{n_1}{n_2} \quad \text{mit } n_1, n_2 \in \mathbb{N}.$$

Dann haben die Funktionen die gemeinsame Periode  $n_1 T_1 = n_2 T_2$ . Für die Frequenzen folgt  $\nu_1/\nu_2 = n_1/n_2$  beziehungsweise  $\omega_1/\omega_2 = n_1/n_2$ .

**Satz 3.3** *Seien  $x_{1,2}(t)$  harmonische Schwingungen mit derselben Schwingungsdauer, also*

$$x_1(t) = A_1 \cos(\omega t + \alpha_1) \quad x_2(t) = A_2 \cos(\omega t + \alpha_2).$$

*Dann ist  $x(t) = x_1(t) + x_2(t)$  wieder eine harmonische Schwingung, genauer gilt*

$$x(t) = A \cos(\omega t + \alpha) \quad \text{mit } Ae^{i\alpha} = A_1 e^{i\alpha_1} + A_2 e^{i\alpha_2}.$$

BEWEIS: Es ist praktisch, ins Komplexe zu gehen. Definiere

$$z_1(t) = A_1 e^{i(\omega t + \alpha_1)} \quad z_2(t) = A_2 e^{i(\omega t + \alpha_2)}$$

Wir berechnen

$$z_1(t) + z_2(t) = e^{i\omega t} (A_1 e^{i\alpha_1} + A_2 e^{i\alpha_2}) = A e^{i\alpha} e^{i\omega t} = A e^{i(\omega t + \alpha)}.$$

Durch Bilden des Realteils folgt

$$x_1(t) + x_2(t) = \operatorname{Re}(z_1(t) + z_2(t)) = A \cos(\omega t + \alpha).$$

Damit ist die Behauptung schon verifiziert.  $\square$

Aus Zeitgründen müssen wir für die Anwendungen auf Beugungsmuster und Schwebungen auf das Buch von Meyberg-Vachenaur verweisen.

## 4 Zahlenfolgen und Grenzwerte

Eine Folge reeller Zahlen  $a_1, a_2, a_3, \dots$  ist streng genommen eine Abbildung von  $\mathbb{N}$  nach  $\mathbb{R}$ . Jedem  $n \in \mathbb{N}$  wird das  $n$ -te Folgenglied  $a_n \in \mathbb{R}$  zugeordnet. Um eine Folge zu definieren, kann man entweder die ersten soundsoviel Folgenglieder angeben, oder ein Bildungsgesetz oder eine Rekursionsvorschrift. Zum Beispiel für die Folge der Quadratzahlen:

*erste Folgenglieder:*  $a_n = 1, 4, 9, 16, \dots$

*Bildungsgesetz:*  $a_n = n^2$  für  $n \in \mathbb{N}$

*Rekursionsvorschrift:*  $a_{n+1} = (\sqrt{a_n} + 1)^2$  und  $a_1 = 1$ .

Bei manchen Folgen ist es sinnvoll, die Nummerierung bei  $n = 0$  zu beginnen statt bei  $n = 1$ . Weitere Beispiele von Folgen sind:

- |     |                                   |   |
|-----|-----------------------------------|---|
| (1) | $a_n = a$                         | konstante Folge $a, a, a, \dots$                    |
| (2) | $a_n = n$                         | Folge der natürlichen Zahlen $1, 2, 3, \dots$       |
| (3) | $a_n = a_0 + nd, n = 0, 1, \dots$ | arithmetische Folge $a_0, a_0 + d, a_0 + 2d, \dots$ |
| (4) | $a_n = a_0 q^n, n = 0, 1, \dots$  | geometrische Folge $a_0, a_0 q, a_0 q^2, \dots$     |
| (5) | $a_n = n a_{n-1}, a_0 = 1$        | $1, 1, 2, 6, 24, \dots$ bzw. $a_n = n!$             |

**Definition 4.1 (Konvergenz)** Die Folge  $a_n$  konvergiert gegen  $a \in \mathbb{R}$ , falls gilt:

Zu jedem  $\varepsilon > 0$  gibt es ein  $N \in \mathbb{R}$ , so dass für alle  $n > N$  gilt:  $|a_n - a| < \varepsilon$ .

$a$  heißt Grenzwert der Folge. Wir schreiben  $\lim_{n \rightarrow \infty} a_n = a$  oder  $a_n \rightarrow a$  für  $n \rightarrow \infty$ . Die Folge  $a_n$  heißt konvergent, wenn sie gegen irgendein  $a \in \mathbb{R}$  konvergiert. Divergent bedeutet nicht konvergent.

Die Zahl  $\varepsilon > 0$  gibt vor, wie groß der Fehler zwischen  $a_n$  und  $a$  höchstens sein soll. In der Regel werden die ersten Folgenglieder das nicht leisten. Es soll aber – so die Definition der Konvergenz – eine Schranke  $N$  geben, so dass für alle  $n > N$  die verlangte Genauigkeit erfüllt wird. Für ein kleineres  $\varepsilon > 0$  müssen wir  $N$  typischerweise vergrößern, um die  $\varepsilon$ -Genauigkeit

zu erreichen, das heißt  $N$  hängt von  $\varepsilon > 0$  ab. Mit den Quantoren  $\forall$  (für alle),  $\exists$  (existiert) und  $\Rightarrow$  (daraus folgt) lässt sich die Definition der Konvergenz auch wie folgt fassen:

$$\forall \varepsilon > 0 \exists N \in \mathbb{R} : \left( n > N \Rightarrow |a_n - a| < \varepsilon \right).$$

**Beispiel 4.1 (Harmonische Folge)** Die Folge  $a_n = 1/n$  konvergiert gegen  $a = 0$ . Denn zu gegebenem  $\varepsilon > 0$  wählen wir  $N = 1/\varepsilon$ . Es folgt für alle  $n > N$

$$|a_n - a| = |1/n - 0| = 1/n < 1/N = \varepsilon.$$

□

**Beispiel 4.2 (Konstante Folge)** Ist  $a_n = a$  für alle  $n \in \mathbb{N}$ , so folgt  $\lim_{n \rightarrow \infty} a_n = a$ . Denn für  $\varepsilon > 0$  gilt  $|a_n - a| = 0 < \varepsilon$  für alle  $n > 0$ , also können wir  $N = 0$  wählen. □

**Beispiel 4.3 (Geometrische Folge)** Sei  $q \in \mathbb{R}$  mit  $|q| < 1$ . Dann gilt  $\lim_{n \rightarrow \infty} q^n = 0$ . Um das zu zeigen, können wir  $q \neq 0$  voraussetzen und haben dann  $1/|q| > 1$ , also gilt  $1/|q| = 1+x$  für ein  $x > 0$ . Es folgt mit der Bernoulli-Ungleichung, Satz 3.1,

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \leq \frac{1}{1+nx} \leq \frac{1}{nx} < \varepsilon$$

für alle  $n > 1/(\varepsilon x)$ . Wir können also  $N = 1/(\varepsilon x)$  wählen. □

**Beispiel 4.4 (Plusminusfolge)** Die Folge  $a_n = (-1)^n$  ist nicht konvergent. Denn es gilt für jedes  $a \in \mathbb{R}$

$$2 = |a_n - a_{n+1}| \leq |a_n - a| + |a_{n+1} - a| \text{ für alle } n \in \mathbb{N}.$$

Die rechte Seite müsste aber für große  $n$  klein sein.

Bei der Wahl von  $N$  kommt es nicht drauf an, dass die Schranke kleinstmöglich ist. Dies ist anders, wenn ein Grenzwert numerisch berechnet werden soll, weil dann die Geschwindigkeit der Konvergenz ein Thema ist. Für den Nachweis der Konvergenz an sich reicht es völlig, irgendeine Schranke zu finden. Ist  $N < N'$  und gilt  $|a_n - a| < \varepsilon$  für  $n > N$ , so erst recht für  $n > N'$ . Wir können also  $N$  stets vergrößern. Zum Beispiel können wir statt  $N$  den nächsten Folgenindex  $n_0 \in (N, N+1]$  wählen. Der Grenzwert wird anschaulicher, indem wir folgende Teilmengen von  $\mathbb{R}$  einführen.

**Definition 4.2 ( $\varepsilon$ -Umgebung)** Die  $\varepsilon$ -Umgebung von  $a \in \mathbb{R}$  ist die Menge

$$U_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\} = \{x \in \mathbb{R} : a - \varepsilon < x < a + \varepsilon\}.$$

Eine Folge  $a_n$  konvergiert genau dann gegen  $a \in \mathbb{R}$ , wenn die Folgenglieder ab einer gewissen Nummer in der  $\varepsilon$ -Umgebung von  $a$  liegen, egal wie klein  $\varepsilon > 0$  gewählt ist.

Manchmal ist diese Beschreibung der Konvergenz praktischer. Zum Beispiel verwenden wir sie, um die Eindeutigkeit des Grenzwerts zu zeigen.

**Satz 4.1 (Eindeutigkeit des Grenzwerts)** Ist die Folge  $a_n$  konvergent, so ist ihr Grenzwert eindeutig bestimmt.

BEWEIS: Angenommen die Folge  $a_n$  hat zwei Grenzwerte  $a \neq a'$ . Wir wählen  $\varepsilon = |a - a'|/2 > 0$  und haben für  $n$  groß (Bild)

$$a_n \in U_\varepsilon(a) \cap U_\varepsilon(a') = \emptyset,$$

ein Widerspruch. □

**Definition 4.3** Eine reelle Folge  $a_n$  heißt beschränkt, wenn es ein  $K \geq 0$  gibt mit

$$|a_n| \leq K \quad \text{für alle } n.$$

Genauer kann noch wie folgt differenziert werden:

$$\begin{aligned} a_n \text{ nach unten beschränkt} &\Leftrightarrow \exists K_1 \in \mathbb{R} \text{ mit } a_n \geq K_1 \text{ für alle } n \in \mathbb{N} \\ a_n \text{ nach oben beschränkt} &\Leftrightarrow \exists K_2 \in \mathbb{R} \text{ mit } a_n \leq K_2 \text{ für alle } n \in \mathbb{N}. \end{aligned}$$

Die Folge  $a_n$  ist genau dann beschränkt, wenn sie nach oben und unten beschränkt ist. Denn aus  $|a_n| \leq K$  folgt  $-K \leq a_n \leq K$ . Umgekehrt folgt aus  $K_1 \leq a_n \leq K_2$ , dass

$$a_n \leq K_2 \leq |K_2| \quad \text{und} \quad -a_n \leq -K_1 \leq |K_1|,$$

also  $|a_n| \leq \max(|K_1|, |K_2|)$ .

**Beispiel 4.5** Die Folge  $a_n = n$  ist nach unten beschränkt, denn es ist zum Beispiel  $a_n \geq 0$  für alle  $n$ . Sie ist aber nicht nach oben beschränkt.

Überlegen Sie, welche der obigen Folgen (nach oben bzw. unten) beschränkt sind!

**Satz 4.2 (konvergent  $\Rightarrow$  beschränkt)** Konvergente Folgen sind beschränkt.

BEWEIS: Sei  $a_n \rightarrow a$  mit  $a \in \mathbb{R}$ . Wähle  $N \in \mathbb{N}$  mit  $|a_n - a| < 1$  für alle  $n > N$ . Dann folgt aus der Dreiecksungleichung  $|a_n| \leq |a| + 1$  für  $n > N$ , also

$$|a_n| \leq \max(|a_1|, \dots, |a_N|, |a| + 1) \text{ für alle } n \in \mathbb{N}.$$

□

Nächstes Ziel:

**Satz 4.3 (Rechenregeln für Grenzwerte)** Es gelte  $a_n \rightarrow a, b_n \rightarrow b$  mit  $n \rightarrow \infty$ .

- a)  $\lim_{n \rightarrow \infty} (\lambda a_n + \mu b_n) = \lambda a + \mu b$  für alle  $\lambda, \mu \in \mathbb{R}$
- b)  $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b$ .
- c)  $\lim_{n \rightarrow \infty} a_n/b_n = a/b$ , falls  $b \neq 0$ .

BEWEIS: Wir beginnen mit dem Beweis von b). Nach Satz 4.2 gibt es ein  $K > 0$  mit  $|a_n| \leq K$  für alle  $n \in \mathbb{N}$ , und außerdem mit  $|b| \leq K$ . Dann gilt für alle  $n \in \mathbb{N}$

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \\ &\leq |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| \\ &\leq K(|a_n - a| + |b_n - b|). \end{aligned}$$

Zu  $\varepsilon > 0$  gibt es nun ein  $N \in \mathbb{R}$  mit  $|a_n - a| < \varepsilon/(2K)$  sowie  $|b_n - a| < \varepsilon/(2K)$  für  $n > N$ . Also folgt für  $n > N$

$$|a_n b_n - ab| < K \left( \frac{\varepsilon}{2K} + \frac{\varepsilon}{2K} \right) = \varepsilon.$$

Für a) reicht es wegen b), den Fall  $\lambda = \mu = 1$  zu betrachten. Zu  $\varepsilon > 0$  gibt es ein  $N \in \mathbb{R}$  mit  $|a_n - a| < \varepsilon/2$  und  $|b_n - b| < \varepsilon/2$  für  $n > N$ . Es folgt für  $n > N$

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Für c) behandeln wir erst den Fall  $a_n = b = 1$ . Zu  $\varepsilon > 0$  wähle  $N \in \mathbb{R}$  mit

$$|b_n - 1| \leq \frac{1}{2} \min(\varepsilon, 1) \quad \text{für } n > N.$$

Dann folgt  $|b_n| = |1 - (1 - b_n)| \geq 1 - |1 - b_n| \geq \frac{1}{2}$ , und weiter

$$\left| \frac{1}{b_n} - 1 \right| = \frac{|1 - b_n|}{|b_n|} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

Für  $a_n$  und  $b \neq 0$  beliebig schließen wir mit  $b'_n = b_n/b \rightarrow 1$

$$\frac{a_n}{b_n} = \frac{a_n}{b} \cdot \frac{1}{b'_n} \rightarrow \frac{a}{b} \cdot 1 = \frac{a}{b}.$$

□

Hier zwei Anwendungen der Rechenregeln für Grenzwerte.

**Beispiel 4.6 (geometrische Reihe)** Für  $-1 < q < 1$  betrachten wir die Folge

$$a_n = 1 + q + \dots + q^n = \sum_{k=0}^n q^k.$$

Dann ergibt sich aus Beispiel 3.2, Beispiel 4.3 und Satz 4.3

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sum_{k=0}^n q^k = \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}.$$

Wir schreiben hierfür auch  $\sum_{k=0}^{\infty} q^k = 1/(1 - q)$ . Folgen, deren Folgenglieder Summen sind, heißen Reihen. Sie spielen eine große Rolle in der Analysis und werden in Kürze ausführlicher untersucht.

**Beispiel 4.7 (Grenzwerte rationaler Funktionen)** Betrachte die Folge

$$x_n = \frac{a_k n^k + a_{k-1} n^{k-1} + \dots + a_0}{b_\ell n^\ell + b_{\ell-1} n^{\ell-1} + \dots + b_0} \quad \text{für } n \in \mathbb{N},$$

wobei  $k, \ell \in \mathbb{N}_0$ ,  $a_i, b_j \in \mathbb{R}$  mit  $a_k, b_\ell \neq 0$ . Durch Ausklammern folgt

$$x_n = n^{k-\ell} \frac{a_k + a_{k-1} n^{-1} + \dots + a_0 n^{-k}}{b_\ell + b_{\ell-1} n^{-1} + \dots + b_0 n^{-\ell}} \rightarrow \begin{cases} a_k/b_\ell & \text{falls } k = \ell, \\ 0 & \text{falls } k < \ell \\ \pm\infty & \text{falls } k > \ell, \text{ sign } \frac{a_k}{b_\ell} = \pm 1. \end{cases}$$

Siehe Definition 4.4 für den Begriff der (uneigentlichen) Konvergenz gegen  $\pm\infty$ .

**Satz 4.4 (Grenzwerte und Ungleichungen)** Seien  $a_n$  und  $b_n$  konvergent, mit Grenzwerten  $\lim_{n \rightarrow \infty} a_n = a$  und  $\lim_{n \rightarrow \infty} b_n = b$ . Dann gelten folgende Aussagen:

- a) Ist  $a_n \leq b_n$  für alle  $n$ , so folgt  $a \leq b$ .
- b) Gilt  $c \leq a_n \leq d$  für alle  $n$  mit  $c, d \in \mathbb{R}$ , so folgt  $c \leq a \leq d$ .
- c) Ist  $a_n \leq c_n \leq b_n$  und gilt  $a = b$ , so konvergiert auch die Folge  $c_n$  gegen  $a = b$ .

BEWEIS: Da  $a_n \rightarrow a$  und  $b_n \rightarrow b$ , gibt es zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{R}$  mit  $a_n > a - \varepsilon$  und  $b_n < b + \varepsilon$  für alle  $n > N$ . In a) folgt

$$a - \varepsilon < a_n \leq b_n < b + \varepsilon \quad \text{für } n > N,$$

also  $a < b + 2\varepsilon$  für alle  $\varepsilon > 0$ , das heißt  $a \leq b$ . Aussage b) folgt unmittelbar aus a), indem wir  $c, d$  als konstante Folgen auffassen. Unter den Voraussetzungen in c) gilt für  $n > N$  die Ungleichungskette

$$a - \varepsilon < a_n \leq c_n \leq b_n < b + \varepsilon = a + \varepsilon,$$

also  $\lim_{n \rightarrow \infty} c_n = a$  nach Definition des Grenzwerts. □

**Achtung:** aus  $a_n < b_n$  folgt *nicht*  $a < b$ , sondern nur  $a \leq b$ . Die Striktheit von Ungleichungen geht beim Übergang zu Grenzwerten im allgemeinen verloren. Zum Beispiel gilt  $1/n > 0$  für alle  $n \in \mathbb{N}$ , aber  $\lim_{n \rightarrow \infty} 1/n = 0$ .

**Beispiel 4.8 ( $n$ -te Wurzel)** Sei  $a > 0$ . Wir bezeichnen mit  $a^{1/n}$  oder  $\sqrt[n]{a}$  die positive Lösung der Gleichung  $x^n = a$ . Es gibt nur eine, denn für  $x, y > 0$  mit  $x > y$  gilt auch  $x^n > y^n$ . Zur Konstruktion der Lösung kann das Intervallhalbierungsverfahren benutzt werden (vgl. Kapitel I, Abschnitt 2). Wir behaupten nun

$$\lim_{n \rightarrow \infty} a^{1/n} = 1.$$

Im Fall  $a \geq 1$  ist auch  $a^{1/n} \geq 1$ , also  $a^{1/n} = 1 + x_n$  mit  $x_n \geq 0$ . Die Bernoulli-Ungleichung, Satz 3.1, liefert

$$a = (1 + x_n)^n \geq 1 + nx_n \quad \Rightarrow \quad 0 \leq x_n \leq \frac{a - 1}{n} \rightarrow 0.$$

Mit Satz 4.4 c) folgt  $x_n \rightarrow 0$  bzw.  $a^{1/n} \rightarrow 1$ . Für  $0 < a < 1$  folgt wegen  $a^{-1} > 1$

$$a^{1/n} = \frac{1}{(a^{-1})^{1/n}} \rightarrow 1 \quad \text{nach Satz 4.3 c).}$$

**Definition 4.4 (Uneigentliche Konvergenz)** Die Folge  $a_n$  konvergiert *uneigentlich* (oder *divergiert bestimmt*) gegen  $+\infty$ , falls gilt:

Zu jedem  $K > 0$  gibt es ein  $N \in \mathbb{R}$ , so dass  $a_n > K$  für alle  $n > N$ .

Wir schreiben  $\lim_{n \rightarrow \infty} a_n = +\infty$  oder  $a_n \rightarrow +\infty$  mit  $n \rightarrow \infty$ . Uneigentliche Konvergenz gegen  $-\infty$  ist analog definiert.

**Beispiel 4.9** Für  $q > 1$  gilt  $\lim_{n \rightarrow \infty} q^n = +\infty$ . Denn zu gegebenem  $K > 0$  gibt es nach Beispiel 4.3 ein  $N \in \mathbb{R}$  mit  $(1/q)^n < 1/K$  für  $n > N$ , also  $q^n > K$  für  $n > N$ . Insgesamt haben wir für das Verhalten der Folge  $q^n$  mit  $n \rightarrow \infty$  folgende Tabelle:

$$\begin{aligned} q > 1 &\Rightarrow \lim_{n \rightarrow \infty} q^n = +\infty, \\ q = 1 &\Rightarrow \lim_{n \rightarrow \infty} q^n = 1, \\ -1 < q < 1 &\Rightarrow \lim_{n \rightarrow \infty} q^n = 0, \\ q \leq -1 &\Rightarrow \text{nicht konvergent.} \end{aligned}$$

Der Fall  $-1 < q < 1$  wurde in Beispiel 4.3 behandelt. Für  $q \leq -1$  vergleiche Beispiel 4.4.

**Beispiel 4.10 (Harmonische Reihe)** Die Folge  $a_n = \sum_{k=1}^n 1/k$  ist bestimmt divergent gegen  $+\infty$ . Dies zeigen wir, indem wir wie folgt Klammern setzen:

$$\underbrace{\left(\frac{1}{1}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{2} + \frac{1}{3}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{8} + \frac{1}{9} + \dots + \frac{1}{15}\right)}_{\geq 1/2} + \dots$$

Die Summe der  $1/k$  mit  $2^m \leq k < 2^{m+1}$  ist größer als  $2^m \cdot 2^{-(m+1)} = 1/2$ .

**Satz 4.5 (Konvergenz von Kehrwerten)** Für eine Folge  $a_n$  gilt:

- (1) Aus  $a_n \rightarrow +\infty$  (bzw.  $a_n \rightarrow -\infty$ ) folgt  $1/a_n \rightarrow 0$ .
- (2) Aus  $a_n \rightarrow 0$  und  $a_n > 0$  (bzw.  $a_n < 0$ ) folgt  $1/a_n \rightarrow +\infty$  (bzw.  $1/a_n \rightarrow -\infty$ ).

BEWEIS: Übungsaufgabe. □

Das Ziel der Analysis ist es, neue Objekte – Zahlen, Funktionen, Operationen – durch Grenzprozesse zu konstruieren. Unsere Definition des Grenzwerts setzt voraus, dass wir den Grenzwert  $a$  der Folge bereits kennen, damit können wir noch nichts Neues definieren. Hier kommt das Vollständigkeitsaxiom ins Spiel. Wir brauchen eine leicht abgeänderte Fassung.

**Satz 4.6 (Konvergenz monotoner Folgen)** Sei  $a_n$  nach oben beschränkt und monoton wachsend, also  $a_1 \leq a_2 \leq \dots$ . Dann ist die Folge  $a_n$  konvergent.

BEWEIS: Setze  $a = \sup\{a_n : n \in \mathbb{N}\}$ . Es gilt  $a \in \mathbb{R}$ , weil die Folge nach oben beschränkt ist. Nach Definition des Supremums gibt es zu  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  mit  $a_N > a - \varepsilon$ , also gilt

$$a - \varepsilon < a_N \leq a_n \leq a \quad \text{für } n \geq N.$$

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

**Satz 4.7 (Definition der Exponentialfunktion)** Die Exponentialfunktion  $\exp : \mathbb{R} \rightarrow \mathbb{R}$  ist definiert als der Grenzwert

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} := \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{x^k}{k!}.$$

BEWEIS: Sei  $x \in \mathbb{R}$  fest gegeben. Wir müssen zeigen, dass die Folge der Summen

$$\exp_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$$

konvergiert. Wähle  $m \in \mathbb{N}$  mit  $m + 1 \geq 2|x|$ . Für die Summanden  $k \geq m + 1$  folgt

$$\frac{|x|^k}{k!} \leq \frac{|x|^{m+1}}{(m+1)!} \frac{|x|^{k-(m+1)}}{(m+1)^{k-(m+1)}} \leq \frac{|x|^{m+1}}{(m+1)!} \frac{1}{2^{k-(m+1)}}. \quad (4.2)$$

Für  $n \geq m + 1$  folgt mit Dreiecksungleichung und Abschätzung der geometrischen Summe

$$\left| \sum_{k=m+1}^n \frac{x^k}{k!} \right| \leq \sum_{k=m+1}^n \frac{|x|^k}{k!} \leq \frac{|x|^{m+1}}{(m+1)!} \left( 1 + \frac{1}{2} + \dots + \frac{1}{2^{n-(m+1)}} \right) \leq \frac{2|x|^{m+1}}{(m+1)!}. \quad (4.3)$$

Es folgt, immer noch für  $m + 1 \geq 2|x|$ ,

$$|\exp_n(x) - \exp_m(x)| \leq \sum_{k=m+1}^n \frac{|x|^k}{k!} \leq \frac{2|x|^{m+1}}{(m+1)!} \quad \text{für alle } n \geq m + 1. \quad (4.4)$$

Die rechte Seite ist eine Konstante  $K = K(x, m)$ , also unabhängig von  $n$ . Da  $\exp_m(x)$  ebenfalls nicht von  $n$  abhängt, ist die Folge  $\exp_n(x)$  beschränkt. Für  $x \geq 0$  ist sie monoton wachsend, also konvergent nach Satz 4.6. Für  $x < 0$  kann man in gerade und ungerade  $k$  aufspalten, und dann wieder die Monotonie verwenden: es gilt  $E_n(x) = E_n^+(x) + E_n^-(x)$  mit

$$E_n^+(x) = \sum_{k \leq n, k \text{ gerade}} \frac{x^k}{k!} \quad E_n^-(x) = \sum_{k \leq n, k \text{ ungerade}} \frac{x^k}{k!}.$$

Die  $E_n^\pm(x)$  sind monoton wachsend bzw. fallend, konvergieren also wieder nach Satz 4.6.  $\square$

Aus (4.4) erhalten wir noch mit  $n \rightarrow \infty$  für  $m + 1 \geq 2|x|$  die Abschätzung

$$|\exp(x) - \exp_m(x)| \leq \frac{2|x|^{m+1}}{(m+1)!} \quad \text{für } m + 1 \geq 2|x|. \quad (4.5)$$

Mit  $m \rightarrow \infty$  geht die rechte Seite gegen Null nach (4.2).

Die Exponentialfunktion beschreibt das natürliche Wachstum. Wir erläutern das am (weniger natürlichen) Beispiel der Zinseszinsrechnung. Wird ein Euro für ein Jahr mit einem Zinssatz  $x \in \mathbb{R}$  angelegt, so beträgt die Ausszahlung  $a_1(x) = 1 + x$ . Die Idee des Zinseszinses ist es, den Zeitraum in kürzere Abschnitte zu unterteilen und den Zins anteilig pro Abschnitt anzurechnen mit dem Effekt, dass der schon angerechnete Teil des Zinses seinerseits Zinsen produziert. Zum Beispiel ergibt das bei monatlicher Verzinsung nach einem Monat  $1 + \frac{x}{12}$ , nach zwei Monaten  $(1 + \frac{x}{12})(1 + \frac{x}{12}) = (1 + \frac{x}{12})^2$ , und nach zwölf Monaten  $a_{12}(x) = (1 + \frac{x}{12})^{12}$ . Allgemein ergibt sich nach einem Jahr bei Unterteilung in  $n$  Zeiteinheiten

$$a_n(x) = \left( 1 + \frac{x}{n} \right)^n \quad \text{für } x \in \mathbb{R}, n \in \mathbb{N}. \quad (4.6)$$

Es stellt sich ganz natürlich die Frage nach einer kontinuierlichen Verzinsung, also nach dem Grenzwert  $n \rightarrow \infty$ .

**Satz 4.8 (natürliches Wachstum)** Für alle  $x \in \mathbb{R}$  gilt

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

BEWEIS: Mit der binomischen Formel, siehe Satz 3.5, folgt

$$a_n(x) = \sum_{k=0}^n \binom{n}{k} \frac{x^k}{n^k} = \sum_{k=0}^n \underbrace{\frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n}}_{=:c(n,k)} \frac{x^k}{k!}.$$

Es gilt  $c(n, k) \rightarrow 1$  mit  $n \rightarrow \infty$ , also folgt für festes  $m$

$$\lim_{n \rightarrow \infty} \left( \sum_{k=0}^m c(n, k) \frac{x^k}{k!} \right) = \sum_{k=0}^m \frac{x^k}{k!} = \exp_m(x).$$

Zu  $\varepsilon > 0$  wähle  $m \in \mathbb{N}$ , so dass gilt:

$$m+1 \geq 2|x| \quad \text{und} \quad \frac{4|x|^{(m+1)}}{(m+1)!} < \frac{\varepsilon}{2}.$$

Für  $n \geq m+1$  folgt wegen  $0 < c(n, k) < 1$  mit (4.3) und (4.5)

$$\begin{aligned} |a_n(x) - \exp(x)| &= \left| \sum_{k=0}^m c(n, k) \frac{x^k}{k!} - \exp_m(x) + \sum_{k=m+1}^n c(n, k) \frac{x^k}{k!} + \exp_m(x) - \exp(x) \right| \\ &\leq \left| \sum_{k=0}^m c(n, k) \frac{x^k}{k!} - \exp_m(x) \right| + \frac{2|x|^{m+1}}{(m+1)!} + \frac{2|x|^{m+1}}{(m+1)!} \\ &\leq \left| \sum_{k=0}^m c(n, k) \frac{x^k}{k!} - \exp_m(x) \right| + \frac{\varepsilon}{2}. \end{aligned}$$

Jetzt wähle  $n$  so groß, dass der erste Term kleiner  $\varepsilon/2$  ist. □

**Definition 4.5 (Eulersche Zahl)** Die Eulersche Zahl ist

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \approx 2,71828 \dots$$

Jährliche Verzinsung von 1 Euro mit Zinssatz Eins ergibt nach einem Jahr 2 Euro, kontinuierliche Verzinsung dagegen  $e \approx 2,71828 \dots$  Euro. Betrachten wir allgemeiner eine Laufzeit von  $x$  Jahren, wieder mit Zinssatz Eins, so ergibt sich bei kontinuierlicher Verzinsung gerade der Grenzwert

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

Legt man dieses Geld weiter für  $y$  Jahre an, wieder mit Zinssatz Eins, so ist der Kontostand dann  $\exp(x)\exp(y)$ . Andererseits hätte man das Geld genausogut direkt für  $x+y$  Jahre anlegen können, dann bekommt man  $\exp(x+y)$ . Es sollte also gelten

$$\exp(x+y) = \exp(x)\exp(y) \quad \text{für } x, y \in \mathbb{R}.$$

Diese Funktionalgleichung brauchen wir auch im Komplexen, und verallgemeinern dazu die Definition der Exponentialfunktion aus Satz 4.7 wie folgt:

$$\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} \quad \text{für } z \in \mathbb{C}.$$

Die Konvergenz (mit gleichen Abschätzungen) folgt wie im reellen Fall, siehe Satz 4.7.

**Satz 4.9 (Funktionalgleichung der Exponentialfunktion)** *Es gilt*

$$\exp(z + w) = \exp(z) \exp(w) \quad \text{für alle } z, w \in \mathbb{C}.$$

BEWEIS: Die Binomische Formel, siehe Satz 3.5, ergibt

$$\frac{(z + w)^n}{n!} = \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} z^k w^{n-k} = \sum_{k+\ell=n} \frac{z^k w^\ell}{k! \ell!}.$$

Wir schätzen nun wie folgt ab:

$$\begin{aligned} |\exp_{2n}(z) \exp_{2n}(w) - \exp_{2n}(z + w)| &= \left| \sum_{k,\ell \leq 2n} \frac{z^k w^\ell}{k! \ell!} - \sum_{k+\ell \leq 2n} \frac{z^k w^\ell}{k! \ell!} \right| \\ &\leq \sum_{k,\ell \leq 2n, \max(k,\ell) > n} \frac{|z|^k |w|^\ell}{k! \ell!} \\ &= \exp_{2n}(|z|) \exp_{2n}(|w|) - \exp_n(|z|) \exp_n(|w|). \end{aligned}$$

Die rechte Seite geht mit  $n \rightarrow \infty$  gegen Null nach Satz 4.7, und die Behauptung folgt.  $\square$

Wir stellen jetzt den Anschluss her an die Exponentialfunktion aus der Schule, und zeigen

$$\exp(r) = e^r \quad \text{für alle } r \in \mathbb{Q}. \quad (4.7)$$

Durch Induktion erhalten wir sofort  $\exp(nx) = \exp(x)^n$  für  $n \in \mathbb{N}$ . Weiter gilt  $\exp(x) \exp(-x) = \exp(0) = 1$ . Daraus folgt für  $k \in \mathbb{Z}^-$

$$\exp(kx) = \exp(-(-kx)) = \frac{1}{\exp(-kx)} = \frac{1}{\exp(x)^{-k}} = \exp(x)^k.$$

Schließlich gilt für  $r = p/q$  mit  $p \in \mathbb{Z}$ ,  $q \in \mathbb{N}$

$$(\exp(rx))^q = \exp(q \cdot rx) = \exp(px) = \exp(x)^p,$$

also

$$\exp(rx) = \exp(x)^{\frac{p}{q}} = \exp(x)^r \quad \text{für alle } r \in \mathbb{Q}.$$

Mit  $x = 1$  folgt insbesondere  $\exp(r) = e^r$  für  $r \in \mathbb{Q}$ . Für rationale  $x$  kann  $\exp(x)$  als Potenz definiert werden. Für irrationale  $x$  ist das nicht möglich. Die Notation  $e^x$  statt  $\exp(x)$  ist aber durchaus üblich, einfach weil sie sehr suggestiv ist.

Hier eine weitere Anwendung des Konvergenzkriteriums der Monotonie und Beschränktheit, Satz 4.6.

**Beispiel 4.11 (Dezimalbrüche)** Jede Dezimalbruchfolge  $a_n = k_0, k_1 k_2 \dots k_n$  mit  $k_0 \in \mathbb{Z}$  und  $k_j \in \{0, 1, \dots, 9\}$  konvergiert gegen eine gewisse reelle Zahl. Denn die Folge  $a_n$  ist monoton wachsend und es gilt (geometrische Reihe)

$$a_n \leq k_0 + \sum_{j=1}^n 9 \cdot 10^{-j} \leq k_0 + 1,$$

das heißt  $a_n$  ist nach oben beschränkt. Für eine gegebene Zahl  $a \in \mathbb{R}$  kann man die Ziffern induktiv bestimmen durch  $k_0 = \max\{k \in \mathbb{Z} : k \leq a\}$  und

$$k_n = \max\{k \in \mathbb{Z} : a_{n-1} + k \cdot 10^{-n} \leq a\}.$$

Es kann vorkommen, dass zwei Dezimalbrüche dieselbe reelle Zahl liefern – wann?

Wie gesagt ist der Vorteil des Kriteriums der Monotonie und Beschränktheit, dass die Konvergenz ohne a priori Kenntnis des Grenzwerts gezeigt werden kann. Das nachfolgende Kriterium von Augustin Louis Cauchy (1789–1857) ist von derselben Form, braucht aber nicht die Monotonie. Die Idee besteht darin, die Glieder der Folge nicht mit dem unbekanntem Grenzwert, sondern *untereinander* zu vergleichen. Aus Zeitgründen verzichten wir auf den Beweis.

**Satz 4.10 (Konvergenz von Cauchyfolgen)** Sei  $a_n$  eine Cauchyfolge, das heißt :

$$\text{Zu jedem } \varepsilon > 0 \text{ gibt es ein } N \in \mathbb{R}, \text{ so dass } |a_n - a_m| < \varepsilon \text{ für alle } n, m > N.$$

Dann gibt es ein  $a \in \mathbb{R}$  mit  $a_n \rightarrow a$  mit  $n \rightarrow \infty$ .

Beim Nachweis dieser Eigenschaft reicht es aus, die Zahlen  $n, m > 0$  mit  $n < m$  zu betrachten, denn die Definition ist symmetrisch in  $n$  und  $m$  und für  $n = m$  ist nichts zu tun.

## 5 Grenzwerte und Stetigkeit von Funktionen

Wir übertragen jetzt das Konzept des Grenzwerts auf Funktionen.

**Definition 5.1 (Grenzwert für Funktionen)** Die Funktion  $f : D \rightarrow \mathbb{R}$ , wobei  $D \subset \mathbb{R}$ , konvergiert für  $x \rightarrow x_0$  gegen  $a \in \mathbb{R}$ , falls gilt:

$$f(x_n) \rightarrow a \quad \text{für jede Folge } x_n \in D, x_n \neq x_0, \text{ mit } x_n \rightarrow x_0.$$

Hier sind einige Bemerkungen angesagt:

- (1) Für die Existenz und den Wert des Grenzwerts ist es egal, ob  $f(x)$  im Punkt  $x_0$  definiert ist bzw. welchen Funktionswert die Funktion dort hat.
- (2) Der Begriff ist nur sinnvoll, wenn es überhaupt eine solche Folge  $x_n$  gibt. Das ist genau dann der Fall, wenn  $D \cap U_\varepsilon(x_0) \setminus \{x_0\} \neq \emptyset$  für alle  $\varepsilon > 0$ .
- (3) Beim rechtsseitigen (linksseitigen) Grenzwert betrachtet man nur Folgen  $x_n \rightarrow x_0$  mit  $x_n > 0$  (bzw.  $x_n < 0$ ). *Notation:*  $\lim_{x \searrow x_0} f(x)$  bzw.  $\lim_{x \nearrow x_0} f(x)$ .
- (4) Die Definition gilt sinngemäß für die Grenzwerte  $\lim_{x \rightarrow \infty} f(x)$  bzw.  $\lim_{x \rightarrow -\infty} f(x)$ .

**Beispiel 5.1** Die Funktion  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ ,  $f(x) = \sin \frac{1}{x}$ , hat in  $x_0 = 0$  keinen rechtsseitigen Grenzwert. Denn es gilt für  $n \in \mathbb{N}$

$$f\left(\frac{1}{n\pi}\right) = \sin n\pi = 0 \quad \text{aber} \quad f\left(\frac{1}{2n\pi + \pi/2}\right) = \sin(2n\pi + \pi/2) = 1.$$

Für die Funktion  $g(x) = x \sin \frac{1}{x}$  ist dagegen  $\lim_{x \rightarrow 0} g(x) = 0$ , denn es gilt

$$|g(x_n)| \leq |x_n| \rightarrow 0 \quad \text{für } x_n \rightarrow 0, x_n \neq 0.$$

**Beispiel 5.2** Die Signumfunktion

$$\text{sign} : \mathbb{R} \rightarrow \mathbb{R}, \text{sign}(x) = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0 \\ 0 & \text{für } x = 0 \end{cases}$$

hat die einseitigen Grenzwerte  $\lim_{x \searrow 0} \text{sign}(x) = +1$  und  $\lim_{x \nearrow 0} \text{sign}(x) = -1$ , während der Grenzwert  $\lim_{x \rightarrow 0} \text{sign}(x)$  nicht existiert.

Folgende Regeln ergeben sich aus den Aussagen für Folgen, siehe Satz 4.3 und Satz 4.4.

**Satz 5.1 (Rechenregeln für Grenzwerte)** *Es gelten folgende Aussagen:*

(1) *Aus  $f(x) \rightarrow a$ ,  $g(x) \rightarrow b$  für  $x \rightarrow x_0$ , wobei  $a, b \in \mathbb{R}$  und  $x_0 \in \mathbb{R} \cup \{\pm\infty\}$ , folgt*

$$\begin{aligned} \alpha f(x) + \beta g(x) &\rightarrow \alpha a + \beta b \quad (\alpha, \beta \in \mathbb{R}), \\ f(x)g(x) &\rightarrow ab, \\ f(x)/g(x) &\rightarrow a/b, \quad \text{falls } b \neq 0. \end{aligned}$$

(2) *Sei  $f(x) \leq g(x) \leq h(x)$  nahe bei  $x_0$ . Falls  $f(x), h(x) \rightarrow a$  mit  $x \rightarrow x_0$ , so folgt auch  $\lim_{x \rightarrow x_0} g(x) = a$ .*

**Definition 5.2 (Stetigkeit)** *Die Funktion  $f : D \rightarrow \mathbb{R}$  heißt stetig in  $x_0 \in D$ , falls gilt:*

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

Unsere Definition beruft sich auf den Konvergenzbegriff für Folgen. Viele Bücher verwenden eine Formulierung, die nicht auf Folgen zurückgreift, die sogenannte  $\varepsilon$ - $\delta$ -Definition der Stetigkeit: *für alle  $\varepsilon > 0$  gibt es ein  $\delta > 0$ , so dass gilt:*

$$x \in D, |x - x_0| < \delta \quad \Rightarrow \quad |f(x) - f(x_0)| < \varepsilon.$$

Die beiden Formulierungen sind aber äquivalent, und unsere Definition der Konvergenz für Folgen war ja nach demselben Muster gebaut. Die Regeln für Grenzwerte implizieren direkt folgende Regeln zur Bildung stetiger Funktionen.

**Satz 5.2 (Stetigkeitsregeln)** *Seien  $f, g : D \rightarrow \mathbb{R}$  stetig in  $x_0 \in D$ . Dann gilt:*

(1) *Für beliebige  $\alpha, \beta \in \mathbb{R}$  ist die Funktion  $\alpha f + \beta g$  stetig in  $x_0$ .*

(2) Die Funktion  $fg$  ist stetig in  $x_0$ .

(3) Ist  $g(x_0) \neq 0$ , so ist die Funktion  $f/g : D \cap U_\delta(x_0) \rightarrow \mathbb{R}$  für  $\delta > 0$  hinreichend klein definiert und stetig in  $x_0$ .

In (3) muss man auf eine Umgebung  $U_\delta(x_0)$  gehen, da sonst  $g(x)$  Nullstellen haben kann.

**Beispiel 5.3** Konstante Funktionen  $f(x) = c$  sind stetig auf  $\mathbb{R}$ , denn für sie ist

$$x_n \rightarrow x_0 \quad \Rightarrow \quad f(x_n) = c = f(x_0).$$

**Beispiel 5.4** Die Funktion  $f(x) = x$  ist stetig auf  $\mathbb{R}$ , denn es gilt

$$x_n \rightarrow x_0 \quad \Rightarrow \quad f(x_n) = x_n \rightarrow x_0 = f(x_0).$$

**Beispiel 5.5** Betrachte für Polynome  $p(x)$  und  $q(x)$  die rationale Funktion

$$f : \mathbb{R} \setminus N_q \rightarrow \mathbb{R}, \quad f(x) = \frac{p(x)}{q(x)} \quad \text{wobei } N_q = \{x \in \mathbb{R} \mid q(x) = 0\}.$$

Mit den vorangehenden Beispielen und Satz 5.2 folgt, dass  $f(x)$  stetig ist auf  $\mathbb{R} \setminus N_q$ . Sei nun  $\lambda \in N_q$   $m$ -fache Nullstelle von  $q(x)$  und  $k$ -fache Nullstelle von  $p(x)$  (mit  $k = 0$  im Fall  $p(x) \neq 0$ ). Dann gibt es Polynome  $p_1(x)$  und  $q_1(x)$  mit  $p_1(\lambda), q_1(\lambda) \neq 0$ , so dass gilt:

$$f(x) = (x - \lambda)^{k-m} \frac{p_1(x)}{q_1(x)} =: f_1(x) \quad \text{für alle } x \in \mathbb{R} \setminus N_q.$$

Im Fall  $k \geq m$  ist  $f_1 : \mathbb{R} \setminus N_{q_1} \rightarrow \mathbb{R}$  stetig im Punkt  $\lambda$  mit Funktionswert

$$f_1(\lambda) := \begin{cases} \frac{p_1(\lambda)}{q_1(\lambda)} & \text{falls } k = m, \\ 0 & \text{falls } k > m. \end{cases}$$

Damit ist  $f_1$  stetige Fortsetzung von  $f$  auf  $\mathbb{R} \setminus N_q \cup \{\lambda\}$ . Dies erklärt die Bezeichnung *hebbare Singularität* aus Kapitel 2.2. Im Fall  $k < m$  kann es keine stetige Fortsetzung geben, da  $f(x_n) \rightarrow \pm\infty$  für jede Folge  $x_n \rightarrow x_0$ .

**Beispiel 5.6** Die charakteristische Funktion von  $\mathbb{Q}$  (oder Dirichlet-Funktion)

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{sonst.} \end{cases}$$

ist nirgends stetig, denn  $\mathbb{Q}$  und  $\mathbb{R} \setminus \mathbb{Q}$  sind beide dicht in  $\mathbb{R}$  (vgl. Kapitel 1.2). Ist zum Beispiel  $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ , so gibt es eine Folge  $x_n \in \mathbb{Q}$  mit  $x_n \rightarrow x_0$ , also  $\lim_{n \rightarrow \infty} \chi_{\mathbb{Q}}(x_n) = 1 \neq 0 = \chi_{\mathbb{Q}}(x_0)$ .

**Satz 5.3 (Verkettung stetiger Funktionen)** Seien  $f : D \rightarrow \mathbb{R}$ ,  $g : E \rightarrow \mathbb{R}$  mit  $f(D) \subset E \subset \mathbb{R}$ . Ist  $f$  stetig in  $x_0$  und  $g$  stetig in  $y_0 = f(x_0)$ , so ist  $g \circ f : D \rightarrow \mathbb{R}$  stetig in  $x_0$ .

BEWEIS: Ist  $x_n \in D$  eine beliebige Folge mit  $\lim_{n \rightarrow \infty} x_n = x_0$ , so folgt  $f(x_n) \rightarrow f(x_0)$  aus der Stetigkeit von  $f$  in  $x_0$ , und weiter  $g(f(x_n)) \rightarrow g(f(x_0))$  wegen der Stetigkeit von  $g$  in  $y_0 = f(x_0)$ .  $\square$

**Beispiel 5.7** Die Betragsfunktion ist stetig auf  $\mathbb{R}$ , denn es gilt

$$x_n \rightarrow x_0 \quad \Rightarrow \quad ||x_n| - |x_0|| \leq |x_n - x_0| \rightarrow 0.$$

Ist  $f : D \rightarrow \mathbb{R}$  stetig, so auch  $|f| : D \rightarrow \mathbb{R}$ .

**Satz 5.4 (Intervallschachtelungsprinzip)** Seien  $I_n = [a_n, b_n]$  Intervalle mit  $I_1 \supset I_2 \supset \dots$  und  $b_n - a_n \rightarrow 0$  mit  $n \rightarrow \infty$ . Dann gibt es genau ein  $x \in \mathbb{R}$  mit  $x \in I_n$  für alle  $n \in \mathbb{N}$ , und zwar gilt  $x = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$ .

BEWEIS: Übungsaufgabe. □

**Satz 5.5 (Zwischenwertsatz)** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Dann gibt es zu jedem  $y_0$  zwischen  $f(a)$  und  $f(b)$  ein  $x_0 \in [a, b]$  mit  $f(x_0) = y_0$ .

*Bemerkung.* Die Gleichung  $f(x) = y_0$  kann mehrere Lösungen in  $[a, b]$  besitzen, das heißt  $x_0$  ist im allgemeinen nicht eindeutig bestimmt.

BEWEIS: Wir können annehmen, dass  $y_0 = 0$ , sonst betrachte  $f(x) - y_0$ . Setze  $[a_0, b_0] = [a, b]$ , und konstruiere eine Intervallschachtelung  $[a_n, b_n]$  so dass  $f(a_n)$  und  $f(b_n)$  nicht dasselbe Vorzeichen haben, also  $f(a_n)f(b_n) \leq 0$ . Nun hat  $f(\frac{a_n+b_n}{2})$  höchstens mit einer der Zahlen  $f(a_n)$  und  $f(b_n)$  gleiches Vorzeichen, also können wir als Folgeintervall eines der Intervalle  $[a_n, \frac{a_n+b_n}{2}]$  oder  $[\frac{a_n+b_n}{2}, b_n]$  wählen. Der durch die Intervallschachtelung definierte Punkt  $x \in [a, b]$  ist eine Nullstelle, denn  $f(x)^2 = \lim_{n \rightarrow \infty} f(a_n)f(b_n) \leq 0$ . □

**Satz 5.6 (Monotonie und Umkehrfunktion)** Sei  $f : I = [a, b] \rightarrow \mathbb{R}$  streng monoton wachsend und stetig. Dann gilt:

(1)  $f(I) = [f(a), f(b)]$ .

(2) Die Umkehrfunktion  $g : [f(a), f(b)] \rightarrow \mathbb{R}$  ist streng monoton wachsend und stetig.

BEWEIS: Aus der Monotonie folgt  $f(I) \subset [f(a), f(b)]$ , Gleichheit liefert der Zwischenwertsatz. Wäre  $g$  nicht streng monoton wachsend, so gibt es  $y_1, y_2 \in f(I)$  mit  $y_1 < y_2$ , aber  $g(y_2) \leq g(y_1)$ . Aus der Monotonie von  $f$  folgt aber

$$y_2 = f(g(y_2)) \leq f(g(y_1)) = y_1, \quad \text{Widerspruch.}$$

Wir zeigen die linksseitige Stetigkeit von  $g$  in einem Punkt  $y_0 \in (f(a), f(b))$ , also  $y_0 = f(x_0)$  mit  $x_0 \in (a, b)$ . Da  $f$  streng monoton ist, gilt  $f(x_0 - \varepsilon) < y_0$  für alle  $\varepsilon > 0$  mit  $x_0 - \varepsilon \geq a$ . Für jede Folge  $y_n \rightarrow y_0$ ,  $y_n < y_0$ , gibt es dann ein  $N \in \mathbb{N}$  mit

$$f(x_0 - \varepsilon) < y_n < y_0 \quad \text{für } n > N.$$

Da  $g$  streng monoton, folgt weiter

$$g(y_0) - \varepsilon = x_0 - \varepsilon < g(y_n) < g(y_0) \quad \text{für } n > N.$$

Also gilt  $\lim_{y \nearrow y_0} g(y) = g(y_0)$ . Die rechtsseitige Stetigkeit folgt analog. □

Der Satz gilt sinngemäß auch auf offenen oder halboffenen Intervallen.

**Beispiel 5.8 (Definition des natürlichen Logarithmus)**  $\exp : (-\infty, \infty) \rightarrow (0, \infty)$  ist streng monoton wachsend, stetig und bijektiv, und es gilt

$$\lim_{x \rightarrow -\infty} \exp(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} \exp(x) = \infty. \quad (5.1)$$

Die Umkehrfunktion  $\ln : (0, \infty) \rightarrow (-\infty, \infty)$  heißt (natürlicher) Logarithmus. Die Funktion ist ebenfalls streng monoton wachsend, stetig und bijektiv, und es gilt

$$\lim_{y \searrow 0} \ln(y) = -\infty \quad \text{und} \quad \lim_{y \rightarrow \infty} \ln(y) = \infty. \quad (5.2)$$

Weiter ist  $\ln(1) = 0$  und  $\ln(e) = 1$ , und  $\ln$  erfüllt die Funktionalgleichung

$$\ln(y_1 y_2) = \ln(y_1) + \ln(y_2) \quad \text{für alle } y_1, y_2 > 0. \quad (5.3)$$

**Definition 5.3 (Potenz mit reellen Exponenten)** Für  $a > 0$ ,  $x \in \mathbb{R}$  definieren wir

$$a^x = \exp(x \ln(a)).$$



## Kapitel 3

# Differentialrechnung für Funktionen einer Variablen

### 1 Die Ableitung: Definition und Regeln

Im diesem Abschnitt betrachten wir reellwertige Funktionen einer Variablen, die auf einem offenen Intervall  $I \subset \mathbb{R}$  definiert sind.

**Definition 1.1 (Ableitung)** Die Funktion  $f : I \rightarrow \mathbb{R}$  hat im Punkt  $x_0 \in I$  die Ableitung  $a \in \mathbb{R}$  (Notation:  $f'(x_0) = a$  oder  $\frac{df}{dx}(x_0) = a$ ), falls gilt:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = a. \quad (1.1)$$

Wir nennen  $f$  differenzierbar in  $x_0$ , falls es ein  $a \in \mathbb{R}^n$  mit (1.1) gibt, falls also der in (1.1) betrachtete Grenzwert existiert.

Eine alternative Formulierung ergibt sich durch die Substitution  $x = x_0 + h$ :

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

Leibniz interessierte sich für die Definition der Ableitung im Zusammenhang mit dem Problem, die Tangente an eine ebene Kurve in einem gegebenen Punkt zu definieren. Nehmen wir dazu an, dass die Kurve als Graph einer Funktion  $f : I \rightarrow \mathbb{R}$  gegeben ist, und dass die Tangente im Punkt  $(x_0, f(x_0))$  gesucht ist. Der Differenzenquotient

$$\frac{f(x) - f(x_0)}{x - x_0} \quad (x_0, x \in I, x \neq x_0)$$

ist geometrisch die Steigung der Sekante durch die Punkte  $(x_0, f(x_0))$  und  $(x, f(x))$ . Die Existenz der Ableitung bedeutet, dass die Sekantensteigungen für  $x \rightarrow x_0$  gegen den Wert  $f'(x_0)$  konvergieren. Die Tangente wird nun definiert als die Gerade, die durch den Punkt  $(x_0, f(x_0))$  geht und die Steigung  $f'(x_0)$  hat. Daraus ergibt sich ihre Gleichung

$$y = f(x_0) + f'(x_0)(x - x_0) \quad \text{für alle } x \in \mathbb{R}.$$

Newton entwickelte den Differentialkalkül (Englisch: Calculus) unter anderem um die Kepler'schen Gesetze für die Planetenbewegung zu begründen, genauer konnte er diese Gesetze alle

aus dem Gravitationsgesetz ableiten. Dazu wird die Bewegung eines Planeten durch seinen Ortsvektor

$$\vec{f}: I \rightarrow \mathbb{R}^3, \quad \vec{f}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

beschrieben, also durch dessen Koordinaten zur Zeit  $t \in I$  bezüglich eines Euklidischen Koordinatensystems. Erstes Ziel ist dann die Definition der Momentangeschwindigkeit als Vektor in  $\mathbb{R}^3$ . Die vektorielle Durchschnittsgeschwindigkeit auf dem Zeitintervall  $[t_0, t]$  ist der Quotient von Weg und Zeit, also gleich

$$\frac{\vec{f}(t) - \vec{f}(t_0)}{t - t_0} \in \mathbb{R}^3.$$

Die Momentangeschwindigkeit  $\vec{v}(t_0)$  zum Zeitpunkt  $t = t_0$  ist deshalb als vektorielle Ableitung zu definieren, wobei Newton einen Punkt statt eines Strichs benutzt hat:

$$\vec{v}(t_0) = \dot{\vec{f}}(t_0) = \begin{pmatrix} x'(t_0) \\ y'(t_0) \\ z'(t_0) \end{pmatrix} \in \mathbb{R}^3.$$

**Definition 1.2 (Ableitungsfunktion)** Die Funktion  $f: I \rightarrow \mathbb{R}$  heißt differenzierbar, falls  $f$  in jedem  $x_0 \in I$  differenzierbar ist. Die hierdurch gegebene Funktion

$$f': I \rightarrow \mathbb{R}, \quad x_0 \mapsto f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

heißt Ableitungsfunktion oder schlicht Ableitung von  $f$ .

**Beispiel 1.1** Für eine konstante Funktion  $f(x) = c$  gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0 \quad \Rightarrow \quad f'(x_0) = 0 \text{ bzw. } f' = 0.$$

**Beispiel 1.2** Für  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x$ , gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1 \quad \text{für alle } x \neq x_0,$$

also folgt  $f'(x_0) = 1$  bzw.  $f' = 1$ .

**Beispiel 1.3** Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = |x|$ , ist nicht differenzierbar in  $x_0 = 0$ :

$$\lim_{x \searrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \searrow 0} \frac{x}{x} = 1 \quad \text{und} \quad \lim_{x \nearrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \nearrow 0} \frac{-x}{x} = -1.$$

Die rechts- und linksseitige Ableitung existieren in  $x_0 = 0$ , sie sind aber verschieden.

**Beispiel 1.4** Für die Funktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}$  gilt  $\exp' = \exp$ . Wir verwenden dazu die Abschätzung (4.5) im Fall  $m = 1$ :

$$|\exp(x) - (1 + x)| \leq |x|^2 \quad \text{für } |x| \leq 1.$$

Es folgt

$$\left| \frac{\exp(x) - \exp(0)}{x - 0} - 1 \right| = \frac{|\exp(x) - (1 + x)|}{|x|} \leq |x| \rightarrow 0 \quad \text{mit } x \rightarrow 0.$$

Also ist  $\exp'(0) = 1$ . Für  $x \neq 0$  schließen wir weiter mit der Funktionalgleichung

$$\frac{\exp(x+h) - \exp(x)}{h} = \exp(x) \frac{\exp(h) - \exp(0)}{h} \rightarrow \exp(x) \quad \text{mit } h \rightarrow 0.$$

Aus Differenzierbarkeit folgt Stetigkeit (aber nicht umgekehrt):

**Satz 1.1** *Ist  $f : I \rightarrow \mathbb{R}$  differenzierbar in  $x_0 \in I$ , so ist  $f$  auch stetig in  $x_0$ .*

BEWEIS: Es gilt mit  $x \rightarrow x_0$

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

□

**Satz 1.2 (Differentiationsregeln)** *Seien  $f, g : I \rightarrow \mathbb{R}$  differenzierbar in  $x_0 \in I$ . Dann sind auch die Funktionen  $\alpha f + \beta g$  ( $\alpha, \beta \in \mathbb{R}$ ),  $fg$  und  $f/g$  (im Fall  $g(x_0) \neq 0$ ) in  $x_0$  differenzierbar mit folgenden Ableitungen:*

(1) *Linearität:*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

(2) *Produktregel:*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$$

(3) *Quotientenregel:*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$$

BEWEIS: Wir müssen jeweils für  $x \neq x_0$  die Differenzenquotienten bilden und zeigen, dass diese mit  $x \rightarrow x_0$  gegen das gewünschte konvergieren. Für (1) haben wir

$$\begin{aligned} \frac{(\alpha f(x) + \beta g(x)) - (\alpha f(x_0) + \beta g(x_0))}{x - x_0} &= \alpha \frac{f(x) - f(x_0)}{x - x_0} + \beta \frac{g(x) - g(x_0)}{x - x_0} \\ &\rightarrow \alpha f'(x_0) + \beta g'(x_0). \end{aligned}$$

Die Produktregel folgt durch „Mischen der Terme“:

$$\begin{aligned} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \\ &\rightarrow f'(x_0)g(x_0) + f(x_0)g'(x_0), \end{aligned}$$

wobei die Stetigkeit von  $g$  in  $x_0$  benutzt wurde (Satz 1.1). Für die Quotientenregel reicht es, die Funktion  $1/g$  zu betrachten, also  $f \equiv 1$ .

$$\frac{1}{x - x_0} \left( \frac{1}{g(x)} - \frac{1}{g(x_0)} \right) = -\frac{1}{g(x)g(x_0)} \frac{g(x) - g(x_0)}{x - x_0} \rightarrow -\frac{g'(x_0)}{g(x_0)^2}.$$

□

**Beispiel 1.5** Für  $f_n(x) = x^n$  folgt aus Beispiel 1.2, also  $f'_1 = 1$ , und der Produktregel

$$f'_n(x) = (f_1 f_{n-1})'(x) = f'_1(x) f_{n-1}(x) + f_1(x) f'_{n-1}(x) = x^{n-1} + x f'_{n-1}(x),$$

und damit per Induktion  $f'_n(x) = nx^{n-1}$ . Allgemeiner ergibt sich mit Satz 2.1(1) für Polynome  $p(x) = \sum_{k=0}^n a_k x^k$  die Formel

$$p'(x) = \sum_{k=1}^n k a_k x^{k-1} = \sum_{j=0}^{n-1} (j+1) a_{j+1} x^j.$$

**Beispiel 1.6** Für  $f(x) = x^{-n}$ ,  $n \in \mathbb{N}$ , gilt  $f'(x) = -nx^{-n-1}$  nach der Quotientenregel:

$$f'(x) = -\frac{nx^{n-1}}{(x^n)^2} = -nx^{-n-1}.$$

**Satz 1.3 (Kettenregel)** Seien  $f : I \rightarrow \mathbb{R}$ ,  $g : J \rightarrow \mathbb{R}$  mit  $f(I) \subset J$ . Ist  $f$  in  $x_0$  sowie  $g$  in  $y_0 = f(x_0)$  differenzierbar, so ist auch  $g \circ f : I \rightarrow \mathbb{R}$  in  $x_0$  differenzierbar und

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0).$$

BEWEIS: Wir betrachten für  $x \neq x_0$  den Differenzenquotienten:

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

Hier haben wir benutzt, dass  $f(x) \rightarrow f(x_0)$  mit  $x \rightarrow x_0$  nach Satz 1.1. Ein technisches Problem gibt es, wenn  $f(x) = f(x_0)$  für  $x$  nahe  $x_0$ , aber das wollen wir hier nicht behandeln.  $\square$

**Satz 1.4 (Ableitung der Umkehrfunktion)** Die Funktion  $f : (a, b) \rightarrow \mathbb{R}$  sei streng monoton und stetig. Ist  $f'(x_0) \neq 0$ , so ist die Umkehrfunktion  $g = f^{-1}$  differenzierbar in  $y_0 = f(x_0)$  mit Ableitung

$$g'(y_0) = \frac{1}{f'(g(y_0))}.$$

BEWEIS:  $g$  existiert und ist stetig nach Satz 5.6. Wir berechnen für  $y \rightarrow y_0$ , also  $g(y) \rightarrow g(y_0)$ ,

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g(y) - g(y_0)}{f(g(y)) - f(g(y_0))} = \frac{1}{\frac{f(g(y)) - f(g(y_0))}{g(y) - g(y_0)}} \rightarrow \frac{1}{f'(g(y_0))}.$$

$\square$

Die Formel für die Ableitung folgt auch aus der Kettenregel:

$$f(g(y)) = y \quad \Rightarrow \quad f'(g(y_0)) g'(y_0) = 1.$$

Wir sehen: damit die Umkehrfunktion im Punkt  $y_0 = f(x_0)$  differenzierbar ist, muss  $f'(x_0) \neq 0$  gelten. Zum Beispiel kann die Umkehrfunktion von  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$ , im Nullpunkt nicht differenzierbar sein.

Die beiden vorangegangenen Regeln sind in der von Leibniz eingeführten Notation besonders suggestiv. Leibniz schreibt Funktionen in der Form  $y = y(x)$  und bezeichnet die Ableitung mit dem Symbol  $\frac{dy}{dx}$ , das auch als *Differentialquotient* bezeichnet wird. Formal ergeben sich die Kettenregel und die Regel für die Ableitung der Umkehrfunktion dann aus der Bruchrechnung:

$$\begin{aligned} y = y(x), z = z(y) &\Rightarrow \frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}, \\ y = y(x), x = x(y) &\Rightarrow \frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}. \end{aligned}$$

Bei der Anwendung dieser saloppen Notation ist jedoch darauf zu achten, an welchen Stellen die jeweiligen Funktionen auszuwerten sind.

**Beispiel 1.7** Die Funktion  $\ln : (0, \infty) \rightarrow \mathbb{R}$  ist differenzierbar mit Ableitung

$$\ln'(y) = \frac{1}{y}.$$

Dies folgt aus Beispiel 1.4 und Satz 1.4, genauer ist

$$\ln'(y) = \frac{1}{\exp'(\ln y)} = \frac{1}{\exp(\ln y)} = \frac{1}{y}.$$

**Beispiel 1.8** Die Potenzfunktion  $f : (0, \infty) \rightarrow \mathbb{R}$ ,  $f(x) = x^\alpha$  mit  $\alpha \in \mathbb{R}$  ist Verkettung der Funktionen  $\exp : \mathbb{R} \rightarrow \mathbb{R}$  und  $h : \mathbb{R} \rightarrow \mathbb{R}$ ,  $h(x) = \alpha \ln x$ . Mit der Kettenregel berechnen wir

$$f'(x) = \exp(\alpha \ln x) \frac{\alpha}{x} = \alpha \exp(\alpha \ln x) \exp(-\ln x) = \alpha \exp((\alpha - 1) \ln x) = \alpha x^{\alpha-1}.$$

**Beispiel 1.9** Die Exponentialfunktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = a^x$  mit  $a > 0$ , ist Verkettung von  $\exp$  und  $h(x) = (\ln a)x$ , deshalb folgt

$$f'(x) = \exp((\ln a)x) \ln a = (\ln a)a^x.$$

Wir wollen nun die Ableitungen der Funktionen  $\cos x$  und  $\sin x$  in Angriff nehmen.

**Beispiel 1.10 (Ableitung von Kosinus/Sinus)** Für alle  $t \in \mathbb{R}$  gilt

$$\cos'(t) = -\sin t, \quad \sin'(t) = \cos t \quad \text{bzw.} \quad \frac{d}{dt} e^{it} = i e^{it}.$$

Wir berechnen dazu erst die Ableitungen im Punkt  $t = 0$ . Für  $t \in (0, \frac{\pi}{2})$  liegt  $e^{it} = (\cos t, \sin t)$  auf dem rechten oberen Viertelkreis. Da die gerade Verbindung der Punkte  $(1, 0)$  und  $(\cos t, \sin t)$  kürzer ist als der Kreisbogen, gilt

$$\sin t \leq \sqrt{(\sin t)^2 + (1 - \cos t)^2} = |e^{it} - 1| \leq t. \quad (1.2)$$

Mit  $\cos t = \cos^2 \frac{t}{2} - \sin^2 \frac{t}{2}$  folgt aus (1.2)

$$1 - \cos t = 2 \sin^2 \frac{t}{2} \leq 2 \left(\frac{t}{2}\right)^2 = \frac{t^2}{2}. \quad (1.3)$$

Mit Satz 5.1(2) folgt bereits die Ableitung des Kosinus: es ist  $\cos 0 = 1$  und

$$0 \geq \frac{\cos t - 1}{t - 0} \geq -\frac{t}{2} \rightarrow 0 \quad \text{mit } t \searrow 0.$$

Für den Sinus brauchen wir ein weiteres geometrisches Argument: der Schnittpunkt des Strahls  $\lambda(\cos t, \sin t)$  mit der Geraden  $x = 1$  ist  $p = (1, \tan t)$ . Das zu  $t$  gehörige Kreis-segment ist dann ganz enthalten im Dreieck mit Eckpunkten  $(0, 0)$ ,  $(1, 0)$  und  $(1, \tan t)$ . Flächenvergleich ergibt

$$\frac{t}{2\pi} \cdot \pi \leq \frac{1}{2} \tan t \quad \text{bzw.} \quad t \leq \tan t. \quad (1.4)$$

Nun folgt

$$1 \geq \frac{\sin t}{t} = \frac{\tan t}{t} \cos t \geq \cos t \rightarrow 1 \quad \text{mit } t \searrow 0.$$

Der Grenzwert  $\lim_{t \searrow 0} \cos t = \cos 0 = 1$  folgt dabei aus Satz 1.1. Die Grenzwerte für  $t \nearrow 0$  folgen mit  $\sin(-t) = -\sin t$  und  $\cos(-t) = \cos t$ . Insgesamt haben wir  $\cos'(0) = 0$  und  $\sin'(0) = 1$  bewiesen.

Für beliebige  $t$  verwenden wir nun die Additionstheoreme: es gilt

$$\begin{aligned} \frac{\cos(t+h) - \cos t}{h} &= \cos t \frac{\cos h - 1}{h} - \sin t \frac{\sin h}{h} \rightarrow -\sin t \\ \frac{\sin(t+h) - \sin t}{h} &= \sin t \frac{\cos h - 1}{h} + \cos t \frac{\sin h}{h} \rightarrow \cos t. \end{aligned}$$

Schließlich folgt

$$\frac{d}{dt} e^{it} = \frac{d}{dt} (\cos t + i \sin t) = -\sin t + i \cos t = i(\cos t + i \sin t) = i e^{it}.$$

Der Vorteil des gegebenen Arguments liegt in der Anschaulichkeit. Allerdings liegt darin auch ein Nachteil: unsere Definition von  $\cos t$  und  $\sin t$  stützt sich immer noch auf die Anschauung, da wir die Bogenlänge nicht definiert haben. Dies macht es schwierig, mit den Funktionen effektiv umzugehen. Die Differentialgleichungen bedeuten einen großen Gewinn, weil wir sie einsetzen können, um die trigonometrischen Funktionen weiter zu studieren.

**Beispiel 1.11** Die Differenzierbarkeit der Arcusfunktionen auf dem offenen Intervall  $(-1, 1)$  folgt aus Beispiel 1.10 und Satz 1.4. Beachtet man  $\cos^2 + \sin^2 = 1$  sowie  $\arccos x \in (0, \pi)$  bzw.  $\arcsin x \in (-\pi/2, \pi/2)$ , so erhält man

$$\begin{aligned} \arccos'(x) &= \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sin(\arccos x)} = -\frac{1}{\sqrt{1-x^2}} \\ \arcsin'(x) &= \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}}. \end{aligned}$$

## 2 Mittelwertsatz und Anwendungen

Die Funktion  $f : I \rightarrow \mathbb{R}$  hat in  $x_0 \in I$  ein Minimum (bzw. ein Maximum), wenn gilt:

$$f(x_0) \leq f(x) \quad \text{für alle } x \in I \quad (\text{bzw. } f(x_0) \geq f(x) \text{ für alle } x \in I).$$

Man nennt dann  $x_0$  eine Minimalstelle bzw. Maximalstelle. Der folgende Satz garantiert die Existenz solcher Stellen unter geeigneten Voraussetzungen.

**Satz 2.1 (Extremalstellen)** Sei  $f : I = [a, b] \rightarrow \mathbb{R}$  stetig. Dann gibt es  $x_0, x_1 \in I$  mit

$$f(x_0) = \inf_{x \in I} f(x) \quad \text{und} \quad f(x_1) = \sup_{x \in I} f(x).$$

Insbesondere ist  $f$  beschränkt.

BEWEIS: Für  $I = I' \cup I''$  sieht man leicht  $\inf_I f = \min(\inf_{I'} f, \inf_{I''} f)$ . Setze  $\lambda = \inf_I f \in [-\infty, \infty)$  und bestimme durch fortgesetzte Halbierung  $I_k = [a_k, b_k]$  mit  $I_0 = I$  und

$$\inf_{I_k} f = \lambda \quad \text{für } k = 0, 1, \dots$$

Sei  $x \in I_k$  für alle  $k$ . Wäre  $f(x) > \lambda$ , so gibt es ein  $\lambda' > \lambda$  und ein  $\delta > 0$  mit

$$f(y) \geq \lambda' \quad \text{für alle } y \in (x - \delta, x + \delta) \cap I.$$

Dies folgt aus der Stetigkeit von  $f$ . Da  $I_k \subset (x - \delta, x + \delta) \cap I$  für  $k$  hinreichend groß, folgt  $\inf_{I_k} f \geq \lambda' > \lambda$ , ein Widerspruch. Also gilt  $f(x) = \lambda$ , was zu zeigen war.  $\square$

**Satz 2.2 (notwendige Bedingung für Extrema I)** Die Funktion  $f : (a, b) \rightarrow \mathbb{R}$  habe in  $x_0 \in (a, b)$  ein Extremum. Ist  $f$  in  $x_0$  differenzierbar, so gilt  $f'(x_0) = 0$ .

BEWEIS: Im Fall eines Minimums in  $x_0$  haben wir

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geq 0 & \text{für } x > x_0, \\ \leq 0 & \text{für } x < x_0. \end{cases}$$

Mit  $x \searrow x_0$  folgt  $f'(x_0) \geq 0$ , mit  $x \nearrow x_0$  folgt  $f'(x_0) \leq 0$ .  $\square$

Die Funktion  $f(x) = x^3$  erfüllt  $f'(0) = 0$ , aber in  $x = 0$  liegt kein Extremum vor. Die Bedingung  $f'(x_0) = 0$  ist notwendig für eine Extremalstelle einer differenzierbaren Funktion, aber sie ist nicht hinreichend.

**Satz 2.3 (Mittelwertsatz der Differentialrechnung)** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig und differenzierbar auf  $(a, b)$ . Dann gibt es ein  $\xi \in (a, b)$  mit

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

BEWEIS: Wir zeigen die Behauptung zuerst im Fall  $f(a) = f(b) = 0$  (Satz von Rolle). Wir brauchen dann ein  $\xi \in (a, b)$  mit  $f'(\xi) = 0$ . Nach Satz 2.1 gibt es  $\xi_1, \xi_2 \in [a, b]$  mit mit

$$f(\xi_1) = \inf_{x \in [a, b]} f(x) \quad \text{und} \quad f(\xi_2) = \sup_{x \in [a, b]} f(x).$$

Ist  $\xi_1 \in (a, b)$ , so folgt  $f'(\xi_1) = 0$  nach Satz 2.2 und wir können  $\xi = \xi_1$  wählen. Analog, wenn  $\xi_2 \in (a, b)$ . Im verbleibenden Fall  $\xi_1, \xi_2 \in \{a, b\}$  folgt  $\inf f = \sup f = 0$  bzw.  $f(x) = 0$  für alle  $x \in [a, b]$ , und damit auch  $f'(x) = 0$  für alle  $x$ . Seien nun  $f(a), f(b)$  beliebig. Definiere  $h : [a, b] \rightarrow \mathbb{R}$  durch Abziehen der Sekante:

$$h(x) = f(x) - \left( f(a) + \frac{f(b) - f(a)}{b - a} (x - a) \right).$$

Es gilt  $h(a) = h(b) = 0$ . Also existiert ein  $\xi \in (a, b)$  mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

$\square$

**Folgerung 2.1 (Monotoniekriterium)** Sei  $f$  differenzierbar auf  $(a, b)$ , stetig auf  $[a, b]$ . Dann gelten folgende Aussagen:

$$\begin{aligned} f'(x) \geq 0 \text{ f\"ur alle } x \in (a, b) &\Rightarrow f \text{ ist wachsend auf } [a, b] \\ f'(x) \leq 0 \text{ f\"ur alle } x \in (a, b) &\Rightarrow f \text{ ist fallend auf } [a, b] \\ f'(x) = 0 \text{ f\"ur alle } x \in (a, b) &\Rightarrow f \text{ ist konstant.} \end{aligned}$$

Bei strikter Ungleichung folgt strenge Monotonie auf  $[a, b]$ .

BEWEIS: Sei  $a \leq x_1 < x_2 \leq b$ . Nach dem Mittelwertsatz gibt es ein  $\xi \in (x_1, x_2)$ , so dass gilt:

$$f(x_2) - f(x_1) = f'(\xi) \underbrace{(x_2 - x_1)}_{>0} \begin{cases} \geq 0 & \text{wenn } f'(\xi) \geq 0 \\ > 0 & \text{wenn } f'(\xi) > 0 \\ \leq 0 & \text{wenn } f'(\xi) \leq 0 \\ < 0 & \text{wenn } f'(\xi) < 0 \\ = 0 & \text{wenn } f'(\xi) = 0 \end{cases}.$$

□

Wir kommen jetzt zu höheren Ableitungen.

**Definition 2.1** Die  $k$ -te Ableitung von  $f : (a, b) \rightarrow \mathbb{R}$  in  $x_0$  ist induktiv definiert durch

$$f^{(k)}(x_0) = (f^{(k-1)})'(x_0).$$

Damit  $f^{(k)}(x_0)$  definiert ist, müssen also die Ableitungen bis Ordnung  $k-1$  in einer Umgebung von  $x_0$  definiert sein, und  $f^{(k-1)}$  muss in  $x_0$  differenzierbar sein.

Natürlich schreiben wir  $f'$  und  $f''$  statt  $f^{(1)}$  bzw.  $f^{(2)}$ . Mit der zweiten Ableitung gewinnen wir genauere Informationen über lokale Extrema.

**Satz 2.4 (notwendige Bedingung für Extrema II)** Sei  $f : (a, b) \rightarrow \mathbb{R}$  differenzierbar, mit einem Minimum in  $x_0$ . Falls  $f''(x_0)$  existiert, so gilt

$$f'(x_0) = 0, \quad f''(x_0) \geq 0.$$

Für Maxima gilt analog  $f'(x_0) = 0, f''(x_0) \leq 0$ .

BEWEIS:  $f'(x_0) = 0$  wurde in Satz 2.2 bewiesen. Es folgt

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

Wäre  $f''(x_0) < 0$ , so wäre  $f'(x) > 0$  links von  $x_0$  und  $f'(x) < 0$  rechts von  $x_0$ , jeweils nahe bei  $x_0$ . Nach Folgerung 2.1 ist  $f$  dann streng monoton wachsend links von  $x_0$  und streng monoton fallend rechts von  $x_0$ , hat also in  $x_0$  ein striktes, lokales Maximum im Widerspruch zur Annahme. □

Die Funktion  $f(x) = x^4$  zeigt, dass in einem Minimum  $f''(x_0) = 0$  gelten kann. Wir haben nun notwendige Bedingungen für Extrema, aber was ist mit hinreichenden Bedingungen? Offensichtlich kann man aus Eigenschaften im Punkt  $x_0$  höchstens lokale Konsequenzen ziehen, unser Interesse gilt aber globalen Extremaleigenschaften. Dafür spielt folgender Begriff eine Rolle.

**Definition 2.2** Sei  $I \subset \mathbb{R}$  ein Intervall. Eine zweimal differenzierbare Funktion  $f : I \rightarrow \mathbb{R}$  heißt konvex (bzw. konkav), wenn  $f'' \geq 0$  auf  $I$  (bzw.  $f'' \leq 0$  auf  $I$ ).

Fährt man auf dem Graph von  $f$  in Richtung der positiven  $x$ -Achse, so bedeutet Konvexität eine Linkskurve, Konkavität eine Rechtskurve.

**Satz 2.5 (Konvexität)** Ist  $f : (a, b) \rightarrow \mathbb{R}$  zweimal differenzierbar, so sind äquivalent:

(1)  $f$  ist konvex.

(2) Der Graph von  $f$  liegt oberhalb jeder Tangente:

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) \quad \text{für alle } x_0, x \in (a, b).$$

BEWEIS: Für  $g(x) = f(x) - (f(x_0) + f'(x_0)(x - x_0))$  gilt  $g''(x) = f''(x)$ . Sei nun  $f$  und damit  $g$  konvex. Dann ist  $g'$  monoton wachsend nach Folgerung 2.1. Wegen  $g'(x_0) = 0$  folgt

$$g'(x) \begin{cases} \leq 0 & \text{für } x < x_0 \\ \geq 0 & \text{für } x > x_0. \end{cases}$$

Wieder mit Folgerung 2.1 ist  $g$  fallend für  $x < x_0$  und wachsend für  $x > x_0$ . Da  $g(x_0) = 0$ , folgt  $g(x) \geq 0$  für alle  $x \in (a, b)$ , das heißt es gilt (2). Umgekehrt folgt aus (2), dass  $g$  ein Minimum bei  $x_0$  hat, und Satz 2.4 liefert

$$0 \leq g''(x_0) = f''(x_0).$$

Da  $x_0 \in (a, b)$  beliebig, folgt  $f$  konvex. □

**Beispiel 2.1 (Youngsche Ungleichung)** Für  $x, y > 0$  gilt die Ungleichung

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \text{falls } p, q \in (1, \infty) \text{ mit } \frac{1}{p} + \frac{1}{q} = 1.$$

Dazu betrachten wir für festes  $y > 0$  die Funktion

$$f : (0, \infty) \rightarrow \mathbb{R}, f(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy.$$

Es gilt  $f'(x) = x^{p-1} - y$  und  $f''(x) = (p-1)x^{p-2} \geq 0$ , das heißt  $f$  ist konvex. Aber  $f'(x_0) = 0$  für  $x_0 = y^{q/p}$  (beachte  $q = \frac{p}{p-1}$ ), also folgt mit Satz 2.5

$$\frac{x^p}{p} + \frac{y^q}{q} - xy = f(x) \geq f(x_0) = \frac{y^q}{p} + \frac{y^q}{q} - y^{q/p}y = 0.$$

**Definition 2.3** Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $k \in \mathbb{N}_0$ . Wir bezeichnen mit  $C^k(I)$  die Menge der  $k$ -mal stetig differenzierbaren Funktionen  $f : I \rightarrow \mathbb{R}$ , das heißt

$$C^k(I) = \{f : I \rightarrow \mathbb{R} : f^{(i)} : I \rightarrow \mathbb{R} \text{ sind definiert und stetig für } i = 0, 1, \dots, k\}.$$

Weiter sei  $C^\infty(I)$  die Menge der unendlich oft differenzierbaren Funktionen, also

$$C^\infty(I) = \bigcap_{k \geq 0} C^k(I).$$

Der Umgang mit  $C^\infty$ -Funktionen ist besonders angenehm, weil die Klasse im Gegensatz zu  $C^k(I)$  unter der Bildung von Ableitungen abgeschlossen ist. Es ist klar, dass Polynome unendlich oft differenzierbar sind, ebenso die Exponentialfunktion sowie Kosinus und Sinus. Man kann auch eine Funktion  $f \in C^\infty(\mathbb{R})$  basteln, die für  $x \in (-1, 1)$  positiv ist und sonst Null.



# Kapitel 4

## Integralrechnung

### 1 Das Riemannsche Integral

Das Integral einer nichtnegativen Funktion  $f : I = [a, b] \rightarrow \mathbb{R}$  ist anschaulich der Flächeninhalt des Gebiets  $\{(x, y) : x \in I, 0 < y < f(x)\}$ . Falls  $f$  das Vorzeichen wechselt, sind die Gebiete unterhalb der  $x$ -Achse negativ zu zählen. Allerdings haben wir den Flächeninhalt von Teilmengen des  $\mathbb{R}^2$  noch gar nicht definiert. Deshalb approximieren wir das Integral durch Rechtecksummen.

Eine Zerlegung  $Z$  eines Intervalls  $[a, b]$  ist gegeben durch Punkte  $a = x_0 \leq x_1 \leq \dots \leq x_N = b$ . Wir setzen  $I_k = [x_{k-1}, x_k]$  und  $\Delta x_k = x_k - x_{k-1}$  für  $k = 1, \dots, N$ . Die Feinheit von  $Z$  ist

$$\Delta(Z) = \max_{1 \leq k \leq N} \Delta x_k. \quad (1.1)$$

**Definition 1.1 (Riemannsche Summe)** Sei  $f : I = [a, b] \rightarrow \mathbb{R}$ . Eine Riemannsche Summe zur Zerlegung  $Z$  ist

$$S_Z(f) = \sum_{k=1}^N f(\xi_k) \Delta x_k \in \mathbb{R}.$$

Die Punkte  $\xi_k \in I_k$  heißen Stützstellen.

Die Riemannsche Summe ist ein Näherungswert für das Integral. Eine konkrete Wahl der Zerlegung und der Stützstellen, zum Beispiel die äquidistante Zerlegung mit Intervallmittelpunkten als Stützstellen, führt auf ein numerisches Verfahren zur Approximation des Integrals. Wir wollen aber beliebige Zerlegungen zulassen.

**Satz 1.1 (Integral stetiger Funktionen)** Sei  $f : I = [a, b] \rightarrow \mathbb{R}$  stetig. Dann gibt es ein  $S \in \mathbb{R}$ , das Integral von  $f$ , so dass  $\lim_{n \rightarrow \infty} S_{Z_n}(f) = S$  für jede Folge von Zerlegungen  $Z_n$  mit  $\Delta(Z_n) \rightarrow 0$ :

$$\lim_{n \rightarrow \infty} S_{Z_n}(f) = S =: \int_a^b f(x) dx \quad \text{falls } \Delta(Z_n) \rightarrow 0.$$

Die Integrationsvariable ist analog zu einem Summationsindex, sie kann beliebig umbenannt werden. In Anwendungen ist es aber manchmal hilfreich, den Namen richtig zu wählen. Wir wollen auf den Beweis des wichtigen Satzes zunächst verzichten, da er etwas technisch ist.

**Beispiel 1.1** Die konstante Funktion  $f : I = [a, b] \rightarrow \mathbb{R}$ ,  $f(x) = c$ , ist integrierbar mit

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

Denn für jede Zerlegung  $Z$  und jede Wahl der  $\xi_k \in I_k$  gilt

$$S_Z(f) = \sum_{k=1}^N c \Delta x_k = c \sum_{k=1}^N (x_k - x_{k-1}) = c(b - a).$$

**Beispiel 1.2** Für  $f : [a, b] \rightarrow \mathbb{R}$ ,  $f(x) = x$  wähle die Unterteilungspunkte  $x_k = a + k \frac{b-a}{N}$ ,  $k = 0, 1, \dots, N$ , und die Stützstellen  $\xi_k = x_k$  mit  $k \geq 1$ . Dann folgt

$$\begin{aligned} S_Z(f) &= \sum_{k=1}^N \left( a + k \frac{b-a}{N} \right) \frac{b-a}{N} \\ &= a(b-a) + (b-a)^2 \frac{N(N+1)}{2N^2} \\ &\rightarrow \frac{1}{2}(b^2 - a^2) \quad \text{mit } N \rightarrow \infty. \end{aligned}$$

Es folgt mit Satz 1.1

$$\int_a^b x dx = \frac{1}{2}(b^2 - a^2).$$

**Beispiel 1.3** Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  heißt stückweise stetig, wenn es eine Zerlegung  $a = x_0 < x_1 < \dots < x_N = b$  gibt, so dass  $f$  auf jedem Intervall  $[x_{k-1}, x_k]$  stetig ist nach eventueller Abänderung in den Endpunkten  $x_{k-1}, x_k$ . Dies ist gleichbedeutend damit, dass die links- und rechtsseitigen Grenzwerte in den Punkten  $x_k$  existieren. Es ist plausibel, dass Satz 1.1 auch für stückweise stetige Funktionen gilt, das heißt alle Folgen von Riemannschen Summen  $S_{Z_n}(f)$  mit  $\Delta(Z_n) \rightarrow 0$  haben einen gemeinsamen Grenzwert. Damit ist das Integral definiert, und es gilt

$$\int_a^b f(x) dx = \sum_{k=1}^N \int_{x_{k-1}}^{x_k} f(x) dx. \quad (1.2)$$

Der einfachste Fall sind die (Riemannschen) Treppenfunktionen: für diese ist  $f$  auf  $(x_{k-1}, x_k)$  konstant gleich  $c_k \in \mathbb{R}$ , und es gilt

$$\int_a^b f(x) dx = \sum_{k=1}^N c_k (x_k - x_{k-1}).$$

Im allgemeinen ist die Berechnung des Integrals als Grenzwert von Riemannschen Summen unpraktisch. Es ist viel effektiver, den Hauptsatz der Differential- und Integralrechnung einzusetzen. Bevor wir diese Methode behandeln, sammeln wir einige Eigenschaften des Integrals.

**Satz 1.2 (Linearität des Integrals)** Für  $f, g : I = [a, b] \rightarrow \mathbb{R}$  stetig und  $\alpha, \beta \in \mathbb{R}$  gilt

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

BEWEIS: Für jede Zerlegung  $Z$  mit Stützstellen  $\xi_k$  gilt

$$\begin{aligned} S_Z(\alpha f + \beta g) &= \sum_{k=1}^N (\alpha f(\xi_k) + \beta g(\xi_k)) \Delta x_k \\ &= \alpha \sum_{k=1}^N f(\xi_k) \Delta x_k + \beta \sum_{k=1}^N g(\xi_k) \Delta x_k \\ &= \alpha S_Z(f) + \beta S_Z(g). \end{aligned}$$

Wählen wir eine Folge  $Z_n$  mit  $\Delta(Z_n) \rightarrow 0$ , so folgt die Behauptung mit Satz 1.1.  $\square$

**Satz 1.3 (Monotonie des Integrals)** Seien  $f, g : I = [a, b] \rightarrow \mathbb{R}$  stetig. Dann folgt:

$$f(x) \leq g(x) \text{ für alle } x \in I \quad \Rightarrow \quad \int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

BEWEIS: Für jede Zerlegung  $Z$  mit Stützstellen  $\xi_k$  gilt

$$S_Z(f) = \sum_{k=1}^N f(\xi_k) \Delta x_k \leq \sum_{k=1}^N g(\xi_k) \Delta x_k = S_Z(g).$$

Wähle wieder eine Folge  $Z_n$  mit  $\Delta(Z_n) \rightarrow 0$  mit  $n \rightarrow \infty$ .  $\square$

**Satz 1.4** Für  $f : I = [a, b] \rightarrow \mathbb{R}$  stetig gelten die Ungleichungen

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

BEWEIS: Für jede Zerlegung  $Z$  mit Stützstellen  $\xi_k$  gilt mit der Dreiecksungleichung

$$\left| \sum_{k=1}^n f(\xi_k) \Delta x_k \right| \leq \sum_{k=1}^N |f(\xi_k)| \Delta x_k \leq \sup_I |f| \sum_{k=1}^N \Delta x_k = (b - a) \sup_I |f|.$$

Mit anderen Worten

$$|S_Z(f)| \leq S_Z(|f|) \leq (b - a) \sup_I |f|.$$

Die Abschätzungen folgen nun durch Wahl von  $Z_n$  wie oben.  $\square$

Aus der Monotonie des Integrals folgt mit dem Zwischenwertsatz der

**Satz 1.5 (Mittelwertsatz der Integralrechnung)** Seien  $f, \varphi : [a, b] \rightarrow \mathbb{R}$  stetig und es sei  $\varphi \geq 0$ . Dann gibt es ein  $\xi \in [a, b]$  mit

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

Im Spezialfall  $\varphi = 1$  folgt  $\int_a^b f(x) dx = f(\xi)(b - a)$ .

BEWEIS: Wir können annehmen, dass  $\int_a^b \varphi(x) dx = 1$ . Setze  $m = \min_{x \in I} f(x)$  und  $M = \max_{x \in I} f(x)$ . Dann gilt  $m\varphi \leq f\varphi \leq M\varphi$ , also

$$m = \int_a^b m\varphi(x) dx \leq \int_a^b f(x)\varphi(x) dx \leq \int_a^b M\varphi(x) dx = M.$$

Nach dem Zwischenwertsatz gibt es ein  $\xi \in [a, b]$  mit  $f(\xi) = \int_a^b f(x)\varphi(x) dx$ . □

Wir bringen jetzt der Vollständigkeit halber den Beweis von Satz 1.1, dieser kann aber problemlos übersprungen werden. Sei eine stetige Funktion  $f : I = [a, b] \rightarrow \mathbb{R}$  gegeben. Für eine Zerlegung  $Z$  mit Punkten  $a = x_0 < x_1 < \dots < x_N = b$  definieren wir die Ober- und Untersumme durch

$$\underline{S}_Z(f) = \sum_{k=1}^N (\min_{I_k} f) \Delta x_k \leq \sum_{k=1}^N (\max_{I_k} f) \Delta x_k = \overline{S}_Z(f).$$

Für Zerlegungen  $Z, Z'$  bezeichnen wir mit  $Z \cup Z'$  die Zerlegung, die durch Vereinigung und Anordnung aller Unterteilungspunkte entsteht. Es gilt dann

$$\overline{S}_{Z \cup Z'}(f) \leq \overline{S}_Z(f) \quad \text{und} \quad \underline{S}_{Z \cup Z'}(f) \geq \underline{S}_Z(f). \quad (1.3)$$

Denn nehmen wir zu  $Z$  einen Punkt  $\xi \in (x_{k-1}, x_k)$  hinzu, so folgt mit  $I_k = I'_k \cup I''_k$

$$\begin{aligned} \overline{S}_{Z \cup \{\xi\}}(f) - \overline{S}_Z(f) &= (\max_{I'_k} f) |I'_k| + (\max_{I''_k} f) |I''_k| - (\max_{I_k} f) |I_k| \leq 0 \\ \underline{S}_{Z \cup \{\xi\}}(f) - \underline{S}_Z(f) &= (\min_{I'_k} f) |I'_k| + (\min_{I''_k} f) |I''_k| - (\min_{I_k} f) |I_k| \geq 0. \end{aligned}$$

Um zu messen, wie stark  $f$  lokal schwankt, definieren wir die Oszillationsfunktion

$$\text{osc}_f(\delta) = \sup_{x, x' \in I, |x-x'| \leq \delta} |f(x) - f(x')|.$$

Für  $\Delta(Z) \leq \delta$  gilt dann die Abschätzung

$$\overline{S}_Z(f) - \underline{S}_Z(f) = \sum_{k=1}^N (\max_{I_k} f - \min_{I_k} f) \Delta x_k \leq \text{osc}_f(\delta) (b - a). \quad (1.4)$$

Wähle nun eine Folge  $Z_n$  mit  $Z_1 \subset Z_2 \subset \dots$  und  $\delta_n = \Delta(Z_n) \rightarrow 0$ . Dann ist  $\overline{S}_{Z_n}(f)$  monoton fallend,  $\underline{S}_{Z_n}(f)$  monoton wachsend, und beide Folgen sind beschränkt:

$$\underline{S}_{Z_1}(f) \leq \underline{S}_{Z_n}(f) \leq \overline{S}_{Z_n}(f) \leq \overline{S}_{Z_1}(f).$$

Damit existieren die Grenzwerte

$$\underline{S}(f) = \lim_{n \rightarrow \infty} \underline{S}_{Z_n}(f) \leq \lim_{n \rightarrow \infty} \overline{S}_{Z_n}(f) = \overline{S}(f).$$

Wir verwenden die Tatsache, dass für  $f$  stetig die Oszillation lokal klein ist, genauer gilt

$$\text{osc}_f(\delta) \rightarrow 0 \quad \text{mit} \quad \delta \rightarrow 0 \quad \text{für} \quad f : [a, b] \rightarrow \mathbb{R} \text{ stetig.} \quad (1.5)$$

Dies kann mit dem Verfahren der Intervallhalbierung gezeigt werden, wir verzichten auf den Beweis. Wegen  $|I_k| \leq \Delta(Z_n) = \delta_n \rightarrow 0$  folgt aus (1.4) und (1.5)

$$\bar{S}_{Z_n}(f) - \underline{S}_{Z_n}(f) \leq \text{osc}_f(\delta_n)(b-a) \rightarrow 0.$$

Wir zeigen nun, dass jede Folge von Riemannsummen gegen  $S(f) := \bar{S}(f) = \underline{S}(f)$  konvergiert, sofern die Feinheit gegen Null geht. Betrachte dazu wieder eine beliebige Zerlegung  $Z$ . Nehmen wir zu  $Z$  einen Punkt  $\xi \in (x_{k-1}, x_k)$  hinzu, so gilt

$$\begin{aligned} \bar{S}_{Z \cup \{\xi\}}(f) - \bar{S}_Z(f) &= (\max_{I'_k} f) |I'_k| + (\max_{I''_k} f) |I''_k| - (\max_{I_k} f) |I_k| \\ &\geq -(\max_{I_k} |f|) |I_k| - (\max_{I_k} f) |I_k| \\ &\geq -2(\max_I |f|) \Delta(Z). \end{aligned}$$

Analog folgt für die Untersummen

$$\underline{S}_{Z \cup \{\xi\}}(f) - \underline{S}_Z(f) \leq 2(\max_I |f|) \Delta(Z).$$

Durch Induktion und mit (1.3) ergibt sich

$$\bar{S}_Z(f) \leq \bar{S}_{Z \cup Z_n}(f) + 2N_n(\max_I |f|) \Delta(Z) \leq \bar{S}_{Z_n}(f) + 2N_n(\max_I |f|) \Delta(Z) \quad (1.6)$$

$$\underline{S}_Z(f) \geq \underline{S}_{Z \cup Z_n}(f) - 2N_n(\max_I |f|) \Delta(Z) \geq \underline{S}_{Z_n}(f) - 2N_n(\max_I |f|) \Delta(Z) \quad (1.7)$$

Dabei ist  $N_n$  die Zahl der Unterteilungspunkte von  $Z_n$ . Nun ist  $\underline{S}_Z(f) \leq S_Z(f) \leq \bar{S}_Z(f)$  für jede Riemannsche Summe  $S_Z(f)$ , also mit (1.6), (1.7)

$$\underline{S}_{Z_n}(f) - 2N_n(\max_I |f|) \Delta(Z) \leq S_Z(f) \leq \bar{S}_{Z_n}(f) + 2N_n(\max_I |f|) \Delta(Z). \quad (1.8)$$

Zu  $\varepsilon > 0$  können wir ein  $n \in \mathbb{N}$  wählen, so dass  $\bar{S}_{Z_n}(f) - \frac{\varepsilon}{2} < S(f) < \underline{S}_{Z_n}(f) + \frac{\varepsilon}{2}$ . Sei weiter  $\Delta(Z) > 0$  so klein, dass  $2N_n(\max_I |f|) \Delta(Z) \leq \frac{\varepsilon}{2}$ . Aus (1.8) folgt

$$S(f) - \varepsilon < \underline{S}_{Z_n}(f) - \frac{\varepsilon}{2} \leq S_Z(f) \leq \bar{S}_{Z_n}(f) + \frac{\varepsilon}{2} < S(f) + \varepsilon.$$

Also konvergiert jede Folge Riemannscher Summen gegen  $S(f)$ , wenn die Feinheit der zugehörigen Zerlegungen gegen Null geht. Dies beweist Satz 1.1.

## 2 Ableitung und Integral

Wir kommen nun zu dem zentralen, von Newton und Leibniz studierten Zusammenhang zwischen Differentiation und Integration. Im folgenden bezeichnet  $I$  ein Intervall.

**Definition 2.1** Sei  $f : I \rightarrow \mathbb{R}$ . Eine differenzierbare Funktion  $F : I \rightarrow \mathbb{R}$  heißt *Stammfunktion* von  $f$ , wenn gilt:

$$F' = f \quad \Leftrightarrow \quad F'(x) = f(x) \quad \text{für alle } x \in I.$$

**Satz 2.1** Ist  $F$  eine Stammfunktion von  $f$  auf  $I$ , so ist jede Stammfunktion von  $f$  auf  $I$  von der Form  $F + c$ , für eine Konstante  $c \in \mathbb{R}$ .

BEWEIS: Sei  $G$  auch Stammfunktion von  $f$  auf  $I$ . Es folgt

$$(G - F)' = G' - F' = f - f = 0.$$

Nach Folgerung 2.1 in Kapitel 4 gibt es eine Konstante  $c \in \mathbb{R}$  mit  $G - F = c$ , also  $G = F + c$ .  $\square$

Folgerung 2.1 sagt aus, dass eine Lösung  $F$  der Gleichung  $F' = f$  bis auf eine additive Konstante  $c \in \mathbb{R}$  eindeutig bestimmt ist. Es schließt sich hier unmittelbar die Frage nach der Existenz an:

*Für welche  $f$  ist die Gleichung  $F' = f$  auf dem Intervall  $I$  lösbar?*

Um das zu beantworten, verallgemeinern wir die Definition des Integrals noch auf den Fall von Integrationsgrenzen  $b \leq a$ . Wir setzen

$$\int_a^b f(x) dx := - \int_b^a f(x) dx \quad \text{und} \quad \int_a^a f(x) dx = 0.$$

Es folgt dann für beliebige  $a, b, c \in \mathbb{R}$

$$\int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx. \quad (2.1)$$

Für  $a \leq b \leq c$  gilt dies nach Beispiel 1.3, und allgemein folgt (2.1) dann durch Vertauschung von  $a, b$  und  $c$ .

**Satz 2.2 (Hauptsatz der Differential- und Integralrechnung)** *Sei  $f : I \rightarrow \mathbb{R}$  stetig. Dann ist für jedes  $x_0 \in I$  die Funktion*

$$F : I \rightarrow \mathbb{R}, \quad F(x) = \int_{x_0}^x f(\xi) d\xi$$

*eine Stammfunktion von  $f$ , das heißt es gilt  $F'(x) = f(x)$  für alle  $x \in I$ .*

BEWEIS: Die Funktion  $F$  ist wohldefiniert nach Satz 1.1. Wir berechnen

$$\begin{aligned} \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| &= \frac{1}{|h|} \left| \int_{x_0}^{x+h} f(\xi) d\xi - \int_{x_0}^x f(\xi) d\xi - hf(x) \right| \\ &= \frac{1}{|h|} \left| \int_x^{x+h} (f(\xi) - f(x)) d\xi \right| \\ &\leq \frac{1}{|h|} |h| \sup_{|\xi-x| \leq |h|} |f(\xi) - f(x)| \quad (\text{Satz 1.4}). \end{aligned}$$

Da  $f$  im Punkt  $x$  stetig ist, geht die rechte Seite mit  $h \rightarrow 0$  gegen Null.  $\square$

**Folgerung 2.1** *Sei  $f : I \rightarrow \mathbb{R}$  stetig und  $F : I \rightarrow \mathbb{R}$  sei eine Stammfunktion von  $f$ . Dann gilt für jedes  $x_0 \in I$*

$$F(x) = F(x_0) + \int_{x_0}^x f(\xi) d\xi \quad \text{für alle } x \in I. \quad (2.2)$$

BEWEIS: Nach Satz 2.1 und Satz 2.2 gibt es ein  $c \in \mathbb{R}$  mit

$$F(x) = c + \int_{x_0}^x f(\xi) d\xi \quad \text{für alle } x \in I.$$

Setze  $x = x_0$  und erhalte  $c = F(x_0)$ . □

**Folgerung 2.2 (Berechnung von bestimmten Integralen mit einer Stammfunktion)**

Sei  $f : I = [a, b] \rightarrow \mathbb{R}$  stetig und  $F : I \rightarrow \mathbb{R}$  sei Stammfunktion von  $f$ . Dann gilt

$$\int_a^b f(x) dx = F(b) - F(a) =: [F(x)]_{x=a}^{x=b}$$

BEWEIS: Folgt aus (2.2) mit  $x_0 = a$ ,  $x = b$ . □

Die Folgerung eröffnet die Möglichkeit, viele Integrale durch Kenntnis von Stammfunktionen (bzw. Ableitungsregeln) auszurechnen. Hier eine kleine Aufleit-Tabelle:

$f(x)$	$F(x)$	Bemerkungen
$x^k$	$\frac{1}{k+1} x^{k+1}$	$k \in \mathbb{Z} \setminus \{-1\}$ , $x \neq 0$ für $k < 0$
$\frac{1}{x}$	$\ln x$	$x \in (0, \infty)$
$x^\alpha$	$\frac{1}{\alpha+1} x^{\alpha+1}$	$\alpha \in \mathbb{R} \setminus \{-1\}$ , $x \in (0, \infty)$
$e^{\lambda x}$	$\frac{1}{\lambda} e^{\lambda x}$	$\lambda \neq 0$
$\cos x$	$\sin x$	
$\sin x$	$-\cos x$	
$\frac{1}{\cos^2 x}$	$\tan x$	$x \notin (2\mathbb{Z} + 1)\frac{\pi}{2}$
$\frac{1}{\sin^2 x}$	$-\cot x$	$x \notin \mathbb{Z}\pi$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$x \in (-1, 1)$
$\frac{1}{\sqrt{1-x^2}}$	$-\arccos x$	$x \in (-1, 1)$
$\frac{1}{x^2+1}$	$\arctan x$	
$\cosh x$	$\sinh x$	
$\sinh x$	$\cosh x$	
$\frac{1}{\sqrt{x^2+1}}$	$\operatorname{Arsinh} x$	
$\frac{1}{\sqrt{x^2-1}}$	$\operatorname{Arcosh} x$	$x \in (1, \infty)$
$\frac{1}{1-x^2}$	$\operatorname{Artanh} x$	$x \in (-1, 1)$

Weiter können wir mit dem Hauptsatz bzw. Folgerung 2.2 aus den Differentiationsregeln von Kapitel 4.1 Integrationsregeln herleiten. Dies wird im Folgenden durchgeführt.

**Satz 2.3 (Partielle Integration)** Seien  $f, g \in C^1(I)$  mit  $I = [a, b]$ . Dann gilt

$$\int_a^b f(x)g'(x) dx = [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f'(x)g(x) dx.$$

BEWEIS: Es gilt nach der Produktregel

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x) \quad \text{auf } I = [a, b].$$

Folgerung 2.2 liefert die Behauptung. □

**Satz 2.4 (Substitutions- oder Transformationsregel)** Sei  $I = [a, b]$ ,  $I^* = [\alpha, \beta]$  und  $\varphi \in C^1(I)$  mit  $\varphi(I) \subset I^*$ . Dann gilt für  $f \in C^0(I^*)$

$$\int_{\varphi(a)}^{\varphi(b)} f(y) dy = \int_a^b f(\varphi(x)) \varphi'(x) dx.$$

BEWEIS: Wähle nach Satz 2.2 eine Stammfunktion  $F \in C^1(I^*)$  von  $f$ . Dann gilt

$$\begin{aligned} \int_{\varphi(a)}^{\varphi(b)} f(y) dy &= [F(y)]_{x=\varphi(a)}^{x=\varphi(b)} \quad (\text{Folgerung 2.2}) \\ &= [F(\varphi(x))]_{x=a}^{x=b} \\ &= \int_a^b (F \circ \varphi)'(x) dx \quad (\text{Folgerung 2.2}) \\ &= \int_a^b F'(\varphi(x)) \varphi'(x) dx \quad (\text{Kettenregel}) \\ &= \int_a^b f(\varphi(x)) \varphi'(x) dx. \end{aligned}$$

□

Für die Anwendung der Substitutionsregel ist folgendes *Kochrezept* nützlich: bei einem gegebenen Integral  $\int_{\alpha}^{\beta} f(y) dy$  möchten wir  $y = y(x)$  substituieren. Dazu berechnen wir

$$y = y(x) \quad \Rightarrow \quad dy = y'(x) dx.$$

Für die Intervallgrenzen bestimmen wir durch Auflösen nach  $x$  die Umkehrfunktion

$$x = x(y) \quad \Rightarrow \quad a = x(\alpha), b = x(\beta).$$

Damit gilt

$$\int_{\alpha}^{\beta} f(y) dy = \int_a^b f(y(x)) y'(x) dx.$$

Es folgen einige Anwendungen der Regeln. Ein Beispiel für die partielle Integration ist

**Beispiel 2.1**

$$\int_1^x \ln u du = \int_1^x 1 \cdot \ln u du = [u \ln u]_{u=1}^{u=x} - \int_1^x u \frac{1}{u} du = x \ln x - (x - 1).$$

Eine schöne Anwendung von Satz 2.3 ist wie folgt.

**Beispiel 2.2** Wir berechnen  $A_n = \int_0^{\pi/2} \sin^n x dx$ . Zunächst gilt

$$A_0 = \pi/2 \quad \text{und} \quad A_1 = [-\cos x]_{x=0}^{x=\pi/2} = 1.$$

Für  $n \geq 1$  erhalten wir durch partielle Integration

$$\begin{aligned} A_{n+1} &= \int_0^{\pi/2} \sin x \sin^n x dx \\ &= \underbrace{[-\cos x \sin^n x]_{x=0}^{x=\pi/2}}_{=0} + n \int_0^{\pi/2} \cos^2 x \sin^{n-1} x dx \\ &= n \int_0^{\pi/2} \sin^{n-1} x dx - n \int_0^{\pi/2} \sin^{n+1} x dx, \end{aligned}$$

wobei wir  $\cos^2 x = 1 - \sin^2 x$  benutzt haben. Es folgt

$$A_{n+1} = \frac{n}{n+1} A_{n-1} \quad (n \geq 1).$$

Durch Induktion erhalten wir

$$\begin{aligned} A_{2n} &= \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdot \dots \cdot \frac{1}{2} \cdot A_0 = \left( \prod_{k=1}^n \frac{2k-1}{2k} \right) \frac{\pi}{2}, \\ A_{2n+1} &= \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdot \dots \cdot \frac{2}{3} \cdot A_1 = \left( \prod_{k=1}^n \frac{2k}{2k+1} \right) 1. \end{aligned}$$

Als nächstes Beispiele zur Substitutionsregel.

**Beispiel 2.3 (Lineare Parameterwechsel)** Im Integral  $\int_{\alpha}^{\beta} f(y) dy$  substituieren wir

$$y = \frac{x - x_0}{\lambda} \quad \text{für } x_0 \in \mathbb{R}, \lambda > 0.$$

Es ist  $dy = 1/\lambda dx$  und  $x = x_0 + \lambda y$ . Also folgt

$$\int_{\alpha}^{\beta} f(y) dy = \frac{1}{\lambda} \int_{x_0+\lambda\alpha}^{x_0+\lambda\beta} f\left(\frac{x-x_0}{\lambda}\right) dx.$$

**Beispiel 2.4 (Integration von Ableitungen)**

$$\begin{aligned} \int_a^b \frac{f'(x)}{f(x)} dx &= \int_a^b (\ln f)'(x) dx = [\ln f(x)]_{x=a}^{x=b} \quad (f > 0), \\ \int_a^b u \sqrt{1+u^2} du &= \frac{1}{3} \int_a^b ((1+u^2)^{\frac{3}{2}})' du = \left[ \frac{1}{3} (1+u^2)^{\frac{3}{2}} \right]_{u=a}^{u=b}, \\ \int_a^b F'(f(x)) f'(x) dx &= [F(f(x))]_{x=a}^{x=b}. \end{aligned}$$

**Beispiel 2.5 (Flächeninhalt unter Hyperbel)** Zu berechnen ist das Integral

$$A(x) = \int_1^x \sqrt{u^2 - 1} du \quad \text{für } x \geq 1.$$

Wir substituieren  $u = \cosh t$  und erhalten  $du = \sinh t dt$ ,  $t = \operatorname{Arcosh} u$ , also mit  $\operatorname{Arcosh} 1 = 0$

$$\begin{aligned} A(x) &= \int_0^{\operatorname{Arcosh} x} \sinh^2 t dt \\ &= \frac{1}{4} \int_0^{\operatorname{Arcosh} x} (e^{2t} + e^{-2t} - 2) dt \\ &= \left[ \frac{1}{8} (e^{2t} - e^{-2t}) \right]_{t=0}^{t=\operatorname{Arcosh} x} - \frac{1}{2} \operatorname{Arcosh} x \\ &= \frac{1}{2} \left( \left[ \frac{e^t + e^{-t}}{2} \frac{e^t - e^{-t}}{2} \right]_{t=0}^{t=\operatorname{Arcosh} x} - \operatorname{Arcosh} x \right) \\ &= \frac{1}{2} (x \sqrt{x^2 - 1} - \operatorname{Arcosh} x). \end{aligned}$$

Für rationale Funktionen, also Quotienten von Polynomen, hat man ein spezielles Integrationsverfahren, die Partialbruchzerlegung, die wir hier nur an einem Beispiel vorführen:

**Beispiel 2.6 (Partialbruchzerlegung)** Um das Integral  $\int_{-1/2}^{1/2} \frac{dx}{1-x^2}$  zu berechnen, machen wir den Ansatz

$$\frac{1}{1-x^2} = \frac{1}{(1+x)(1-x)} \stackrel{!}{=} \frac{A}{1-x} + \frac{B}{1+x} = \frac{(A-B)x + (A+B)}{1-x^2}.$$

Der Koeffizientenvergleich ergibt  $A = B = 1/2$ , also folgt

$$\int_{-1/2}^{1/2} \frac{dx}{1-x^2} = \int_{-1/2}^{1/2} \left( \frac{1/2}{1-x} + \frac{1/2}{1+x} \right) dx = \frac{1}{2} \left[ \ln \frac{1+x}{1-x} \right]_{x=-1/2}^{x=1/2} = \frac{1}{2} (\ln 3 - \ln 1/3) = \ln 3.$$

Wir wollen auch Integrale über unendliche Intervalle berechnen, sowie Integrale mit Ausnahmestellen, wo die Funktion  $f(x)$  rechts- oder linksseitig nicht stetig ist. Wir fangen mit der Situation an, wo die Problemstelle die rechte (oder linke, analog) Intervallgrenze ist.

**Definition 2.2 (uneigentliches Integral)** Sei  $-\infty < a < b \leq \infty$ . Für eine stetige Funktion  $f : I = [a, b) \rightarrow \mathbb{R}$  setzen wir

$$\int_a^b f(x) dx = \lim_{\beta \nearrow b} \int_a^\beta f(x) dx,$$

falls der rechte Grenzwert existiert. Andernfalls heißt das Integral divergent.

Hier einige Beispiele von uneigentlichen Integralen.

**Beispiel 2.7**

$$\begin{aligned} \int_1^\infty x^\alpha dx &= \begin{cases} \lim_{R \nearrow \infty} \left[ \frac{x^{\alpha+1}}{\alpha+1} \right]_{x=1}^{x=R} = -\frac{1}{\alpha+1} & \text{für } \alpha < -1, \\ \text{divergent} & \text{für } \alpha \geq -1. \end{cases} \\ \int_0^1 x^\alpha dx &= \begin{cases} \lim_{\varepsilon \searrow 0} \left[ \frac{x^{\alpha+1}}{\alpha+1} \right]_{x=\varepsilon}^{x=1} = \frac{1}{\alpha+1} & \text{für } \alpha > -1, \\ \text{divergent} & \text{für } \alpha \leq -1. \end{cases} \\ \int_0^1 \frac{dx}{\sqrt{1-x^2}} &= \lim_{b \nearrow 1} \int_0^b \frac{dx}{\sqrt{1-x^2}} = \lim_{b \nearrow 1} \arcsin b = \pi/2. \end{aligned}$$

Als nächstes betrachten wir eine Funktion  $f : (a, b) \rightarrow \mathbb{R}$ , für die sowohl die linke als auch die rechte Intervallgrenze problematisch ist, zum Beispiel  $(a, b) = (-\infty, \infty)$ . Wir wählen dann einen Zwischenpunkt  $c \in (a, b)$  und setzen

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx = \lim_{\alpha \searrow a} \int_\alpha^c f(x) dx + \lim_{\beta \nearrow b} \int_c^\beta f(x) dx,$$

sofern die beiden Grenzwerte bzw. uneigentliche Integrale rechts existieren. Diese Definition hängt nicht von der Wahl von  $c$  ab, denn es folgt zum Beispiel für  $c' > c$

$$\int_\alpha^{c'} f(x) dx = \int_\alpha^c f(x) dx + \int_c^{c'} f(x) dx \xrightarrow{\alpha \searrow a} \int_a^c f(x) dx + \int_c^{c'} f(x) dx,$$

$$\int_{c'}^{\beta} f(x) dx = \int_c^{\beta} f(x) dx - \int_c^{c'} f(x) dx \xrightarrow{\beta \nearrow b} \int_c^b f(x) dx - \int_c^{c'} f(x) dx.$$

Also bleibt die Summe der Integrale gleich.

### Beispiel 2.8

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} &= \lim_{R \rightarrow \infty} \int_{-R}^0 \frac{dx}{1+x^2} + \lim_{R \rightarrow \infty} \int_0^R \frac{dx}{1+x^2} \\ &= \lim_{R \rightarrow \infty} [\arctan x]_{x=-R}^{x=0} + \lim_{R \rightarrow \infty} [\arctan x]_{x=0}^{x=R} \\ &= \frac{\pi}{2} + \frac{\pi}{2} = \pi. \end{aligned}$$

Ganz analog ist die Situation im Fall einer stetigen Funktion  $f : [a, b] \setminus \{c\} \rightarrow \mathbb{R}$  mit  $a < c < b$ , also mit einer Singularität im Innern des Intervalls. Wir definieren

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

wobei wieder beide uneigentlichen Integrale rechts existieren müssen.

**Beispiel 2.9** Für  $\alpha > -1$  berechnen wir

$$\int_{-1}^1 |x|^\alpha dx = \int_{-1}^0 (-x)^\alpha dx + \int_0^1 x^\alpha dx = 2 \int_0^1 x^\alpha dx = \frac{2}{\alpha+1}.$$

**Beispiel 2.10** Für die Funktion  $f(x) = \frac{1}{x}$  ist das uneigentliche Integral auf dem Intervall  $[-1, 1]$  nicht definiert, denn die einzelnen Integrale gehen gegen  $\pm\infty$ :

$$\begin{aligned} \int_{\varepsilon}^1 \frac{dx}{x} &= [\ln x]_{x=\varepsilon}^1 = -\ln \varepsilon \rightarrow \infty, \\ \int_{-1}^{-\varepsilon} \frac{dx}{x} &= [\ln(-x)]_{x=-1}^{x=-\varepsilon} = \ln \varepsilon \rightarrow -\infty. \end{aligned}$$

Allerdings existiert immer noch der Grenzwert

$$\lim_{\varepsilon \searrow 0} \left( \int_{-1}^{-\varepsilon} \frac{dx}{x} + \int_{\varepsilon}^1 \frac{dx}{x} \right) = \lim_{\varepsilon \searrow 0} (\ln \varepsilon - \ln \varepsilon) = 0.$$

Diesen Wert nennt man dann den Cauchy-Hauptwert des Integrals (Englisch: principal value). Solche Hauptwerte sind mit Vorsicht zu genießen, wir wollen sie nicht weiter betrachten.

Es stellt sich die Frage nach Kriterien zur Existenz uneigentlicher Integrale.

**Satz 2.5 (Vergleichstest)** Sei  $f : [0, \infty) \rightarrow \mathbb{R}$  stetig. Es gebe  $\alpha > 1$  und  $R > 0$  mit

$$|f(x)| \leq \frac{C}{x^\alpha} \quad \text{für alle } x \geq R,$$

wobei  $C < \infty$  Konstante. Dann konvergiert das unbestimmte Integral  $\int_0^\infty f(x) dx$ .

BEWEIS: Sei zunächst  $f \geq 0$ . Dann gilt für alle  $b \in [R, \infty)$

$$\begin{aligned} 0 &\leq \int_0^b f(x) dx \\ &\leq \int_0^R f(x) dx + \int_R^b \frac{C}{x^\alpha} dx \\ &\leq \int_a^R f(x) dx + C \int_R^\infty \frac{dx}{x^\alpha} < \infty. \end{aligned}$$

Die Funktion  $F(b) = \int_0^b f(x) dx$  ist monoton wachsend und nach oben beschränkt, also konvergiert das Integral mit  $b \rightarrow \infty$ . Hat  $f : [0, \infty) \rightarrow \mathbb{R}$  kein Vorzeichen, so betrachten wir

$$f^+(x) = \max(f(x), 0) \quad \text{und} \quad f^-(x) = -\min(f(x), 0).$$

Die Funktionen  $f^\pm$  sind stetig und erfüllen die Abschätzung

$$0 \leq f^\pm(x) \leq |f(x)| \leq \frac{C}{x^\alpha} \quad \text{für alle } x \geq R.$$

Also existieren die uneigentlichen Integrale von  $f^\pm$ . Aber  $f(x) = f^+(x) - f^-(x)$ , somit

$$\int_0^b f(x) dx = \int_0^b f^+(x) dx - \int_0^b f^-(x) dx,$$

und die gewünschte Konvergenz für  $b \rightarrow \infty$  folgt. □

**Beispiel 2.11** Als letztes Beispiel betrachten wir die Funktion

$$f : \left[ \frac{\pi}{2}, \infty \right) \rightarrow \mathbb{R}, \quad f(x) = \frac{\sin x}{x}.$$

Wir können Satz 2.5 nicht direkt anwenden, denn es gilt nur  $|f(x)| \leq 1/x$ , wir brauchen aber eine Potenz größer als Eins. Der Trick ist, die Konvergenz durch partielle Integration zu verbessern: es gilt

$$\int_{\frac{\pi}{2}}^b \sin x \frac{1}{x} dx = \left[ (-\cos x) \frac{1}{x} \right]_{x=\frac{\pi}{2}}^{x=b} - \int_{\frac{\pi}{2}}^b (-\cos x) \left( -\frac{1}{x^2} \right) dx.$$

Das rechte Integral erfüllt die Abschätzung aus Satz 2.5 mit  $\alpha = 2 > 1$ , also konvergiert es für  $b \rightarrow \infty$ . Wir erhalten

$$\lim_{b \rightarrow \infty} \int_{\frac{\pi}{2}}^b \frac{\sin x}{x} dx = - \int_{\frac{\pi}{2}}^\infty \frac{\cos x}{x^2} dx \in \mathbb{R}.$$

Meistens wird das Integral von  $\frac{\sin x}{x}$  auf  $[0, \infty)$  genommen. Im Nullpunkt gibt es aber kein Problem, denn dort konvergiert die Funktion gegen Eins. Für die Frage der Konvergenz des Integrals spielt der Anfangspunkt also keine Rolle.

### 3 Darstellung von Funktionen durch Reihen

Viele Funktionen in der Analysis können als unendliche Reihen definiert beziehungsweise dargestellt werden. Darum ist es wichtig, die Konvergenzfrage für Reihen zu untersuchen. Wichtige Beispiele sind die Exponentialfunktion und die trigonometrischen Funktionen.

**Definition 3.1** Seien  $a_n \in \mathbb{R}$  für  $n \in \mathbb{N}_0$ . Die Reihe mit den Gliedern  $a_n$  ist die Folge  $S_0 = a_0$ ,  $S_1 = a_0 + a_1$ ,  $S_2 = a_0 + a_1 + a_2$ ,  $\dots$ , bzw. allgemein

$$S_n = \sum_{k=0}^n a_k \quad \text{für alle } n \in \mathbb{N}_0.$$

Die Reihe heißt konvergent mit Wert  $S \in \mathbb{C}$ , wenn  $S_n$  gegen  $S$  konvergiert:

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

Die Zahl  $S_n$  heißt  $n$ -te Partialsumme der Reihe. Oft wird die Reihe in der Form  $a_0 + a_1 + a_2 + \dots$  angegeben. Leider ist es auch üblich, die Reihe selbst - unabhängig von der Frage der Konvergenz - mit dem Symbol  $\sum_{k=0}^{\infty} a_k$  zu bezeichnen.

**Beispiel 3.1** Die Reihe  $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots$  bzw.  $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ . Es gilt

$$S_1 = \frac{1}{1 \cdot 2} = \frac{1}{2}, \quad S_2 = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} = \frac{2}{3}, \quad S_3 = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} = \frac{3}{4}, \dots$$

Durch Induktion folgt  $S_n = \frac{n}{n+1}$ . Also ist die Reihe konvergent mit Wert  $\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1$ .

Für konvergente Reihen  $\sum_{k=0}^{\infty} a_k$  und  $\sum_{k=0}^{\infty} b_k$  und  $\lambda, \mu \in \mathbb{R}$  ist auch  $\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k)$  konvergent, und es gilt

$$\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k) = \lambda \sum_{k=0}^{\infty} a_k + \mu \sum_{k=0}^{\infty} b_k.$$

Dies ergibt sich aus dem entsprechenden Satz für Folgen. Durch Abändern, Hinzufügen oder Weglassen von endlich vielen Gliedern ändert sich nichts an der Konvergenz oder Divergenz, sondern nur der Wert der Reihe. Zum Beispiel haben wir

$$\sum_{k=n}^m a_k = \sum_{k=0}^m a_k - \sum_{k=0}^{n-1} a_k \xrightarrow{m \rightarrow \infty} \sum_{k=n}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n-1} a_k. \quad (3.1)$$

Wie bei Folgen kann der Anfangsindex einer Reihe statt  $k = 0$  auch eine andere Zahl sein.

**Beispiel 3.2 (Geometrische Reihe)** Die geometrische Reihe  $\sum_{k=0}^{\infty} x^k$  konvergiert genau für  $|x| < 1$ . Die Formel für die geometrische Summe, siehe Beispiel (3.2), ergibt

$$S_n = \sum_{k=0}^n x^k = \frac{1 - x^{n+1}}{1 - x} \rightarrow \frac{1}{1 - x} \quad \text{falls } |x| < 1.$$

Ist dagegen  $|x| \geq 1$ , so folgt  $|S_{n+1} - S_n| = |x|^{n+1} \geq 1$ , und  $S_n$  kann nicht konvergieren.

Im allgemeinen kann eine Reihe nicht einfach ausgerechnet werden. Stattdessen gibt es Tests, mit denen die Konvergenz gezeigt oder widerlegt werden kann, abhängig von den  $a_n$ . Die folgende Aussage ist vom zweiten Typ.

**Satz 3.1 (Nullfolgentest)** Ist  $\sum_{k=0}^{\infty} a_k$  konvergent, so folgt  $\lim_{k \rightarrow \infty} a_k = 0$ .

BEWEIS: Nach Voraussetzung ist die Folge der Partialsummen  $S_n = \sum_{k=0}^n a_k$  konvergent mit Grenzwert  $S$ , also folgt  $a_n = S_n - S_{n-1} \rightarrow S - S = 0$ .  $\square$

**Beispiel 3.3 (Harmonische Reihe)** Die Bedingung  $a_k \rightarrow 0$  ist *nicht hinreichend* für die Konvergenz der Reihe, zum Beispiel ist  $\sum_{k=1}^{\infty} 1/k = \infty$ , siehe Kapitel 2, Beispiel 4.10. Für jedes  $m \in \mathbb{N}_0$  ist die Summe der  $1/k$  mit  $2^m \leq k < 2^{m+1}$  mindestens  $1/2$ .

Wir wollen nun hinreichende Kriterien angeben, und beginnen mit einem einfachen Fall.

**Satz 3.2 (Reihen mit nichtnegativen Gliedern)** Eine Reihe  $\sum_{k=0}^{\infty} a_k$  mit  $a_k \geq 0$  konvergiert genau dann, wenn die Partialsummen  $S_n = \sum_{k=0}^n a_k$  nach oben beschränkt sind.

BEWEIS:  $S_n$  ist monoton wachsend, die Behauptung folgt aus Satz 4.6 (Kapitel 2).  $\square$

Ein effektives Kriterium ist der

**Satz 3.3 (Integralvergleichstest)** Sei  $f : [1, \infty) \rightarrow \mathbb{R}$  monoton fallend und nichtnegativ. Dann konvergiert die Reihe  $\sum_{k=1}^{\infty} f(k)$  genau dann, wenn  $\int_1^{\infty} f(x) dx$  konvergiert.

BEWEIS: Für  $x \in [k-1, k]$  gilt  $f(k) \leq f(x) \leq f(k-1)$ , wobei  $k \geq 2$ . Durch Integration folgt

$$f(k) \leq \int_{k-1}^k f(x) dx \leq f(k-1) \quad \Rightarrow \quad \sum_{k=2}^n f(k) \leq \int_1^n f(x) dx \leq \sum_{k=1}^{n-1} f(k).$$

Konvergiert das Integral, so sind die Partialsummen  $S_n$  nach oben beschränkt und die Reihe konvergiert nach Satz 3.2. Ist die Reihe konvergent, so ist das Integral eine monoton wachsende, nach oben beschränkte Funktion der oberen Grenze, also konvergent.  $\square$

**Beispiel 3.4 (die Reihen  $\sum_{k=1}^{\infty} k^\alpha$ )** Für  $\alpha < 0$  ist die Funktion  $f(x) = x^\alpha$  nichtnegativ und fallend auf  $[1, \infty)$ . Das Integral  $\int_1^{\infty} x^\alpha dx$  konvergiert genau für  $\alpha < -1$ , siehe Beispiel 2.7. Somit konvergieren auch die Reihen genau für  $\alpha < -1$ .

**Satz 3.4 (absolut konvergent  $\Rightarrow$  konvergent)** Falls  $\sum_{k=0}^{\infty} |a_k|$  konvergiert, so konvergiert auch die Reihe  $\sum_{k=0}^{\infty} a_k$  und es gilt die Abschätzung

$$\left| \sum_{k=0}^{\infty} a_k \right| \leq \sum_{k=0}^{\infty} |a_k|. \quad (3.2)$$

**Bezeichnung.** Eine Reihe mit  $\sum_{k=0}^{\infty} |a_k| < \infty$  heißt absolut konvergent.

BEWEIS: Für  $a_k^+ = \max(a_k, 0)$  und  $a_k^- = -\min(a_k, 0)$  folgt

$$\sum_{k=0}^n a_k^+ + \sum_{k=0}^n a_k^- = \sum_{k=0}^n |a_k| \leq \sum_{k=0}^{\infty} |a_k| < \infty.$$

Nach Satz 3.2 sind  $\sum_{k=0}^{\infty} a_k^+$  und  $\sum_{k=0}^{\infty} a_k^-$  konvergent. Wegen  $a_k = a_k^+ - a_k^-$  ist dann auch  $\sum_{k=0}^{\infty} a_k$  konvergent, und es folgt

$$\left| \sum_{k=0}^{\infty} a_k \right| = \lim_{n \rightarrow \infty} \left| \sum_{k=0}^n a_k \right| \leq \lim_{n \rightarrow \infty} \sum_{k=0}^n |a_k| = \sum_{k=0}^{\infty} |a_k|.$$

□

**Satz 3.5 (Majorantentest)** *Es gelte  $|a_k| \leq c_k$  wobei  $\sum_{k=0}^{\infty} c_k < \infty$ . Dann ist die Reihe  $\sum_{k=0}^{\infty} a_k$  absolut konvergent.*

BEWEIS: Die Summen  $\sum_{k=0}^n |a_k|$  sind durch  $\sum_{k=0}^{\infty} c_k < \infty$  beschränkt, also ist  $\sum_{k=0}^{\infty} |a_k|$  konvergent nach Satz 3.2. □

Man kann zeigen, dass die Reihe  $\sum_{k=1}^{\infty} (-1)^{k-1}/k = 1 - 1/2 + 1/3 - \dots$  konvergiert, obwohl die entsprechende Reihe mit positiven Vorzeichen – die harmonische Reihe – divergiert. Also ist absolute Konvergenz hinreichend, aber nicht notwendig für Konvergenz. Aber für unsere Zwecke ist absolute Konvergenz meistens passend. Es gibt eine Reihe weiterer Tests für absolute Konvergenz, zum Beispiel das Quotientenkriterium oder das Wurzelkriterium, für die wir aus Zeitgründen auf die Literatur verweisen.

Die einfachsten Funktionen, die wir kennen, sind natürlich die Polynome. Es ist naheliegend, neue Funktionen als Grenzwerte von Folgen bzw. Reihen von Polynomen zu suchen. Dies führt auf den Begriff der Potenzreihe.

**Definition 3.2 (Potenzreihen)** *Eine reelle Potenzreihe ist eine Reihe der Form*

$$P(x) = \sum_{k=0}^{\infty} a_k x^k \quad \text{mit } a_k \in \mathbb{R}.$$

Für jedes  $x \in \mathbb{R}$  haben wir eine Reihe von reellen Zahlen. Die Frage ist, für welche  $x$  die Reihe konvergiert. Es ist  $P(0) = 0$ ; schlimmstenfalls divergiert die Reihe für alle  $x \neq 0$ , dann ist sie natürlich nicht von Interesse. Ein Beispiel ist  $\sum_{n=0}^{\infty} n^n x^n$  (verwende den Nullfolgentest). Als anderes Extrem kann die Reihe für alle  $x \in \mathbb{R}$  konvergent sein, wie bei  $\sum_{k=0}^{\infty} x^k/k!$ . In diesem Fall liefert die Potenzreihe eine interessante Funktion, die Exponentialfunktion, an genau so was sind wir interessiert. Ein anderes Beispiel ist die geometrische Reihe  $\sum_{k=0}^{\infty} x^k$ , die genau für  $x \in (-1, 1)$  konvergiert. Allgemein kann die Menge der  $x$ , für die eine Potenzreihe konvergiert und damit eine Funktion definiert, wie folgt charakterisiert werden.

**Satz 3.6 (vom Konvergenzradius)** *Zu jeder Potenzreihe  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  gibt es genau ein  $R \in [0, \infty]$ , den Konvergenzradius, mit folgender Eigenschaft:*

$$P(x) \text{ ist } \begin{cases} \text{absolut konvergent} & \text{für } |x| < R, \\ \text{divergent} & \text{für } |x| > R. \end{cases}$$

Bemerkung. Für  $|x| = R$  ist keine allgemeine Aussage möglich.

BEWEIS: Die Eindeutigkeit von  $R$  ist klar. Zur Existenz definieren wir

$$R = \sup\{|x| : P(x) \text{ konvergiert}\} \in [0, \infty].$$

Für  $|x| > R$  ist dann offensichtlich  $P(x)$  divergent. Sei nun  $|x| < R$ , also  $P(x_0)$  konvergent für ein  $x_0$  mit  $|x| < |x_0|$ , also  $q := |x|/|x_0| < 1$ . Nach Satz 3.1 geht  $|a_k||x_0|^k$  gegen Null, also gibt es ein  $M \in [0, \infty)$  mit  $|a_k||x_0|^k \leq M$  für alle  $k$ . Es folgt

$$|a_k||x|^k = |a_k||x_0|^k \left(\frac{|x|}{|x_0|}\right)^k \leq Mq^k. \quad (3.3)$$

Die geometrische Reihe  $\sum_{k=0}^{\infty} q^k$  konvergiert, also folgt absolute Konvergenz aus dem Majorantenkriterium.  $\square$

Auf dem Konvergenzintervall  $(-R, R)$  definiert eine Potenzreihe  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  eine Funktion. Es stellt sich die Frage, ob  $P(x)$  differenzierbar ist und wenn ja, ob wir  $P'(x)$  durch gliedweise Differentiation berechnen können: gilt

$$\frac{d}{dx} \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \frac{d}{dx} a_k x^k = \sum_{k=1}^{\infty} k a_k x^{k-1}?$$

Allgemeiner betrachten wir eine Folge  $f_n : I \rightarrow \mathbb{R}$  von Funktionen, die punktweise (d.h. an jeder Stelle  $x \in I$ ) gegen eine Grenzfunktion  $f : I \rightarrow \mathbb{R}$  konvergieren:

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \text{für jedes } x \in I. \quad (3.4)$$

Wir interessieren uns dafür, ob wir Integral und Ableitung mit dem Grenzwert  $n \rightarrow \infty$  vertauschen können. Dies ist ein grundlegendes Problem, das in der Analysis immer wieder auftritt. Im allgemeinen ist die punktweise Konvergenz nicht einmal ausreichend für die Stetigkeit der Grenzfunktion – hier ein Beispiel.

**Beispiel 3.5** Die Folge  $f_n : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f_n(x) = \arctan(nx)$ , konvergiert punktweise gegen

$$f(x) = \begin{cases} \pi/2 & \text{für } x > 0 \\ -\pi/2 & \text{für } x < 0 \\ 0 & \text{für } x = 0. \end{cases}$$

In den nachfolgenden Vertauschungssätzen müssen wir also von  $f_n \rightarrow f$  mehr verlangen als nur die punktweise Konvergenz. Das Zauberwort ist hier gleichmäßige Konvergenz, das heißt

$$\|f_n - f\|_I = \sup_{x \in I} |f_n(x) - f(x)| \rightarrow 0 \quad \text{mit } n \rightarrow \infty. \quad (3.5)$$

Wir nennen  $\|f\|_I = \sup_{x \in I} |f(x)|$  die Supremumsnorm von  $f$  auf dem Intervall  $I$ .

**Satz 3.7 (Gleichmäßige Konvergenz und Stetigkeit)** *Seien  $f_n : I \rightarrow \mathbb{R}$  stetig. Falls die  $f_n$  gleichmäßig gegen  $f : I \rightarrow \mathbb{R}$  konvergieren, so ist  $f$  ebenfalls stetig auf  $I$ .*

BEWEIS: Sei  $x_k \rightarrow x_0 \in I$  gegeben, zu zeigen ist  $f(x_k) \rightarrow f(x_0)$ . Zu  $\varepsilon > 0$  wähle  $n \in \mathbb{N}$  mit  $\|f_n - f\|_I < \varepsilon/3$ . Dann folgt mit der Dreiecksungleichung

$$\begin{aligned} |f(x_k) - f(x_0)| &\leq |f(x_k) - f_n(x_k)| + |f_n(x_k) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &\leq \|f_n - f\|_I + |f_n(x_k) - f_n(x_0)| + 2\varepsilon/3. \end{aligned}$$

Da  $f_n$  stetig, gilt  $|f_n(x_k) - f_n(x_0)| < \varepsilon/3$  für  $k$  hinreichend groß, und damit  $|f(x_k) - f(x_0)| < \varepsilon$ .  $\square$

**Satz 3.8 (Vertauschung von Konvergenz und Integral)** Seien  $f_n : [a, b] \rightarrow \mathbb{R}$  stetig. Falls die  $f_n$  gleichmäßig gegen  $f : [a, b] \rightarrow \mathbb{R}$  konvergieren, so ist  $f$  stetig und es gilt

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

BEWEIS: Die Stetigkeit von  $f$  wurde im vorigen Satz gezeigt. Die Standardabschätzung des Integrals, siehe Satz 1.4, liefert

$$\left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| = \left| \int_a^b (f_n(x) - f(x)) dx \right| \leq \|f_n - f\|_I (b - a) \rightarrow 0.$$

□

**Satz 3.9 (Vertauschung von Konvergenz und Ableitung)** Seien  $f_n : I \rightarrow \mathbb{R}$  stetig differenzierbar, und  $f_n \rightarrow f$  punktweise auf  $I$ . Falls die Folge  $f'_n$  gleichmäßig gegen eine Funktion  $g$  konvergiert, so ist  $f$  stetig differenzierbar mit  $f' = g$ .

BEWEIS: Für  $x_0 \in I$  gilt nach Satz 2.2, dem Hauptsatz der Differential- und Integralrechnung,

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(\xi) d\xi \quad \text{für alle } x \in I.$$

Da  $f'_n$  gleichmäßig gegen  $g$  konvergiert, folgt nach Satz 3.8 mit  $n \rightarrow \infty$

$$f(x) = f(x_0) + \int_{x_0}^x g(\xi) d\xi \quad \text{für alle } x \in I.$$

Die Funktion  $g$  ist stetig nach Satz 3.7, also folgt wieder mit dem Hauptsatz  $f' = g$ . □

Wir kommen jetzt auf die Potenzreihen zurück. Bevor wir die Sätze anwenden können, müssen wir die Frage der Gleichmäßigkeit der Konvergenz klären. Sei  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R > 0$ . Es gilt

$$P(x) - P_n(x) = \sum_{k=n+1}^{\infty} a_k x^k \quad \text{mit } P_n(x) = \sum_{k=0}^n a_k x^k.$$

Wir schätzen wie in Satz 3.6 ab. Für  $\varrho < R$  wähle  $r \in (\varrho, R)$ , also  $q := \varrho/r < 1$ . Da  $P(r)$  konvergiert, gilt  $|a_k| r^k \leq M$  für alle  $k$ . Es folgt für  $n \rightarrow \infty$

$$\|P - P_n\|_{[-\varrho, \varrho]} \leq \sum_{k=n+1}^{\infty} |a_k| \varrho^k = \sum_{k=n+1}^{\infty} |a_k| r^k \left(\frac{\varrho}{r}\right)^k \leq M \sum_{k=n+1}^{\infty} q^k = \frac{M q^{n+1}}{1 - q} \rightarrow 0.$$

**Satz 3.10** Eine Potenzreihe  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  konvergiert auf jedem Teilintervall  $[-\varrho, \varrho]$  gleichmäßig, und  $P(x)$  ist stetig auf  $(-R, R)$ .

BEWEIS: Gleichmäßige Konvergenz wurde soeben gezeigt, Stetigkeit folgt mit Satz 3.7. □

Um die Ableitung einer Potenzreihe zu bilden, brauchen wir eine Hilfsaussage.

**Lemma 3.1** Sei  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  eine Potenzreihe mit Konvergenzradius  $R \in [0, \infty]$ . Dann hat die formal differenzierte Reihe

$$Q(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

denselben Konvergenzradius  $R$ .

BEWEIS: Da  $|a_k x^k| \leq |k a_k x^{k-1}| \cdot |x|$  für  $k \geq 1$ , hat  $Q(x)$  höchstens Konvergenzradius  $R$ . Zu  $r \in (0, R)$  gibt es ein  $M \in [0, \infty)$  mit  $|a_k| r^k \leq M$ , da  $P(r)$  konvergiert. Es folgt für  $0 < |x| < r$ , also  $q = |x|/r \in (0, 1)$ ,

$$k |a_k| |x|^{k-1} = \frac{|a_k| r^k}{|x|} k \left(\frac{|x|}{r}\right)^k \leq \frac{M}{|x|} k q^k.$$

Da  $\ln q < 0$ , konvergiert das Integral der Funktion  $f(s) = s q^s = s e^{s \ln q}$  auf  $[0, \infty)$  (verwende partielle Integration). Also ist die rechte Reihe konvergent nach Satz 3.3, und  $Q(x)$  ist absolut konvergent für  $|x| < R$ .  $\square$

**Satz 3.11 (Differenzierbarkeit von Potenzreihen)** Eine Potenzreihe  $P(x) = \sum_{k=0}^{\infty} a_k x^k$  mit Konvergenzradius  $R > 0$  ist auf dem Intervall  $(-R, R)$  differenzierbar, und die Ableitung kann gliedweise berechnet werden:

$$P'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} \quad \text{für alle } x \in (-R, R). \quad (3.6)$$

BEWEIS: Die Funktionen  $P_n(x) = \sum_{k=0}^n a_k x^k$  konvergieren auf  $(-R, R)$  punktweise gegen  $P(x)$ , und nach Lemma 3.1 und Satz 3.10 konvergieren die  $P'_n$  gleichmäßig auf jedem Teilintervall  $[-\varrho, \varrho] \subset (-R, R)$  gegen die stetige Funktion  $Q : (-R, R) \rightarrow \mathbb{R}$ ,  $Q(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$ . Die Behauptung folgt aus Satz 3.9.  $\square$

Jetzt stehen alle Hilfsmittel zur Verfügung, um einige interessante Potenzreihenentwicklungen herzuleiten.

**Beispiel 3.6** Für  $x \in (-1, 1) \subset \mathbb{R}$  gilt die Reihendarstellung

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - + \dots \quad (3.7)$$

Es  $\ln(1+x)' = (1+x)^{-1} = \sum_{k=0}^{\infty} (-1)^k x^k$  für  $x \in (-1, 1)$ . Die Reihe konvergiert gleichmäßig auf  $[0, x]$ , also können wir Satz 3.8 anwenden:

$$\ln(1+x) = \int_0^x \frac{d\xi}{1+\xi} = \int_0^x \sum_{k=0}^{\infty} (-1)^k \xi^k d\xi = \sum_{k=0}^{\infty} (-1)^k \int_0^x \xi^k d\xi = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

**Beispiel 3.7** Ähnlich zeigen wir für  $x \in (-1, 1) \subset \mathbb{R}$  die Reihenentwicklung

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - + \dots \quad (3.8)$$

Denn  $\arctan'(x) = (1+x^2)^{-1} = \sum_{k=0}^{\infty} (-1)^k x^{2k}$ , und es folgt mit Satz 3.8

$$\arctan x = \int_0^x \frac{d\xi}{1+\xi^2} = \int_0^x \sum_{k=0}^{\infty} (-1)^k \xi^{2k} d\xi = \sum_{k=0}^{\infty} (-1)^k \int_0^x \xi^{2k} d\xi = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}.$$

## 4 Taylorentwicklung

Sei  $f \in C^k(I)$ , das heißt  $f$  hat stetige Ableitungen bis zur Ordnung  $k$ . Die Idee der Taylorentwicklung ist es, die Funktion  $f(x)$  mit einem Polynom zu vergleichen, das mit  $f(x)$  an einer festen Stelle  $x_0$  „von höherer Ordnung“ übereinstimmt. Genauer betrachten wir das  $k$ -te Taylorpolynom

$$P_k(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j. \quad (4.1)$$

Dies hat offenbar die Eigenschaft  $P^{(j)}(x_0) = f^{(j)}(x_0)$  für  $j = 0, 1, \dots, k$ . Es sollte dann auch nahe bei  $x_0$  die Funktion gut approximieren, und das soll quantifiziert werden. Die Fehlerfunktion

$$R_k : I \rightarrow \mathbb{R}, \quad R_k(x) = f(x) - P_k(x) \quad (4.2)$$

heißt Restglied  $k$ -ter Ordnung der Taylorentwicklung in  $x_0$ . Knackpunkt bei der Taylorentwicklung ist die Abschätzung dieses Restglieds und damit eine Aussage darüber, wie gut die Funktion durch das Taylorpolynom approximiert wird. Hierfür gibt es verschiedene mögliche Darstellungen von  $R_k$ .

**Satz 4.1 (Taylorentwicklung)** Sei  $f \in C^{k+1}(I)$  für ein  $k \in \mathbb{N}_0$  und  $x_0 \in I$ . Dann gilt

$$f(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + R_k(x) \quad \text{für } x \in I, \quad (4.3)$$

wobei das Restglied  $R_k(x)$  folgende Darstellungen besitzt:

$$R_k(x) = \frac{1}{k!} \int_{x_0}^x (x - y)^k f^{(k+1)}(y) dy \quad (\text{Cauchy}) \quad (4.4)$$

$$= \frac{f^{(k+1)}(\xi)}{(k+1)!} (x - x_0)^{k+1} \quad \text{für ein } \xi \in [x_0, x] \quad (\text{Lagrange}). \quad (4.5)$$

BEWEIS: Wir zeigen erst (4.4) durch Induktion über  $k \in \mathbb{N}_0$ . Der Fall  $k = 0$  folgt aus dem Hauptsatz der Differential- und Integralrechnung:

$$R_0(x) = f(x) - f(x_0) = \int_{x_0}^x f'(y) dy.$$

Der Induktionsschritt ist beruht auf partieller Integration, und zwar gilt für  $k \geq 1$

$$\begin{aligned} R_k(x) &= R_{k-1}(x) - \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= \frac{1}{(k-1)!} \int_{x_0}^x (x - y)^{k-1} f^{(k)}(y) dy - \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= \frac{1}{(k-1)!} \left[ -\frac{(x - y)^k}{k} f^{(k)}(y) \right]_{y=x_0}^{y=x} + \frac{1}{k!} \int_{x_0}^x (x - y)^k f^{(k+1)}(y) dy - \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= \frac{1}{k!} \int_{x_0}^x (x - y)^k f^{(k+1)}(y) dy. \end{aligned}$$

Die Lagrangedarstellung folgt hieraus mit dem Mittelwertsatz der Integralrechnung, Satz 1.5. Sei zum Beispiel  $x \geq x_0$ , dann gilt für ein  $\xi \in [x_0, x]$

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{k!} \int_{x_0}^x (x-y)^k dy = \frac{f^{(k+1)}(\xi)}{(k+1)!} (x-x_0)^{k+1}.$$

□

**Beispiel 4.1** Betrachte für  $x \in (-1, 1)$  die Funktion  $f(x) = (1-x)^{-1/2}$ , mit Ableitungen

$$f'(x) = \frac{1}{2}(1-x)^{-3/2} \quad \text{und} \quad f''(x) = \frac{3}{4}(1-x)^{-5/2}.$$

Es gilt  $f(0) = 1$  und  $f'(0) = 1/2$ , also lautet das Taylorpolynom der Ordnung Eins in  $x_0 = 0$

$$P_1(x) = f(0) + f'(0)x = 1 + \frac{1}{2}x,$$

mit der Lagrange-Restglieddarstellung

$$R_1(x) = \frac{f''(\xi)}{2}x^2 = \frac{3}{8}(1-\xi)^{-5/2}x^2 \quad \text{für ein } \xi \in [0, x].$$

Als Anwendung erhalten wir für die relativistische Energie eines Teilchens mit Ruhemasse  $m_0$  und Geschwindigkeit  $v$ , wenn wir  $\beta = v/c$  setzen,

$$E = \frac{m_0 c^2}{\sqrt{1-\beta^2}} = m_0 c^2 \left( 1 + \frac{1}{2}\beta^2 + \frac{f''(\xi)}{2}\beta^4 \right) = m_0 c^2 + \frac{1}{2}m_0 v^2 + \Delta E.$$

Dabei ist der erste Term die Ruheenergie und der zweite die klassische kinetische Energie. Für den relativistischen Korrekturterm ergibt sich aus der Restglieddarstellung die Abschätzung

$$\frac{\Delta E}{E_{kin}} = f''(\xi)\beta^2 \leq f''(\beta^2)\beta^2 < 0,008 \quad \text{für } \beta \leq 0,1.$$

Bei Geschwindigkeiten  $v \leq \frac{1}{10}c$  beträgt die relativistische Korrektur weniger als ein Prozent der klassischen kinetischen Energie.

**Definition 4.1** Für  $f \in C^\infty(I)$  und  $x_0 \in I$  heißt die Reihe

$$P(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!} (x-x_0)^j$$

Taylorreihe von  $f$  mit Zentrum  $x_0$ .

Die Taylorreihe ist eine Potenzreihe, wengleich mit Zentrum  $x_0$  statt dem Nullpunkt. Wir können aber alle Aussagen zur Konvergenz anwenden. Nach Satz 3.6 gibt es ein  $R \in [0, \infty]$ , den Konvergenzradius, so dass die Reihe für  $|x-x_0| < R$  absolut konvergiert, für  $|x-x_0| > R$  dagegen divergiert. Es ist nun naheliegend zu vermuten, dass die Reihe auf  $(-R, R)$  gegen  $f(x)$  konvergiert. Das ist im allgemeinen aber falsch, zum Beispiel kann man zeigen, dass folgende Funktion in  $C^\infty(\mathbb{R})$  ist:

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

Offenbar ist  $f^{(j)}(0) = 0$  für alle  $j$ , und somit sind alle Taylorpolynome Null, die Taylorreihe liefert die Nullfunktion und nicht die Funktion  $f(x)$ .

**Satz 4.2 (Darstellbarkeit durch die Taylorreihe)** Sei  $f \in C^\infty(I)$ , und  $P(x)$  sei die Taylorreihe von  $f(x)$  mit Zentrum  $x_0 \in I$ . Für jedes  $x \in I$  gilt dann

$$f(x) = P(x) \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} R_k(x) = 0.$$

BEWEIS: Wegen  $f(x) - P_k(x) = R_k(x)$  ist die Aussage trivial.  $\square$

Wird die Funktion  $f(x)$  auf  $(x_0 - \delta, x_0 + \delta)$  durch eine Reihe  $\sum_{j=0}^{\infty} a_j(x - x_0)^j$  dargestellt, so hat diese Reihe Konvergenzradius mindestens  $\delta > 0$  und kann auf  $(x_0 - \delta, x_0 + \delta)$  gliedweise differenziert werden nach Satz 3.11. Somit ist  $f(x)$  auf  $(x_0 - \delta, x_0 + \delta)$  unendlich oft differenzierbar, und es folgt

$$f^{(k)}(x_0) = \frac{d^k}{dx^k} \sum_{j=0}^{\infty} a_j(x - x_0)^j \Big|_{x=x_0} = \sum_{j=0}^{\infty} \frac{d^k}{dx^k} a_j(x - x_0)^j \Big|_{x=x_0} = k! a_k.$$

Die darstellende Reihe muss also die Taylorreihe sein. Funktionen  $f : I \rightarrow \mathbb{R}$ , die in der Nähe jedes Punkts  $x_0 \in I$  durch ihre Taylorreihe dargestellt werden, nennt man (reell-) analytisch.

Mit der Taylorreihe leiten wir nun einige wichtige Reihenentwicklungen ab.

**Beispiel 4.2** Für die Funktionen  $\cos x$  und  $\sin x$  gilt

$$\cos^{(j)}(0) = \begin{cases} (-1)^k & \text{für } j = 2k \\ 0 & \text{für } j = 2k + 1 \end{cases} \quad \sin^{(j)}(0) = \begin{cases} (-1)^k & \text{für } j = 2k + 1 \\ 0 & \text{für } j = 2k. \end{cases}$$

Daraus folgen die Taylorentwicklungen, für geeignete  $\xi \in [0, x]$ ,

$$\begin{aligned} \cos x &= \sum_{k=0}^n \frac{(-1)^k}{(2k)!} x^{2k} + \cos^{(2n+2)}(\xi) \frac{x^{2n+2}}{(2n+2)!}, \\ \sin x &= \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} x^{2k+1} + \sin^{(2n+3)}(\xi) \frac{x^{2n+3}}{(2n+3)!}. \end{aligned}$$

Da  $x^k/k! \rightarrow 0$  mit  $k \rightarrow \infty$ , gehen die Restglieder mit  $n \rightarrow \infty$  gegen Null. Also folgen die Reihendarstellungen

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} \pm \dots \quad \text{für alle } x \in \mathbb{R}, \quad (4.6)$$

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} \pm \dots \quad \text{für alle } x \in \mathbb{R}. \quad (4.7)$$

Weiter bekommen wir

$$e^{ix} = \cos x + i \sin x = \sum_{k=0}^{\infty} \frac{(ix)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(ix)^{2k+1}}{(2k+1)!} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = \exp(ix).$$

Die Glieder der rechten Reihe sind komplexe Zahlen. Unsere Resultate zur Konvergenz von Reihen gelten aber ganz entsprechend (wie auch bei Folgen). Mit der anschaulichen Definition von  $\cos x$  und  $\sin x$  ist schwierig zu argumentieren, die Begründung der Additionstheoreme und der Ableitungen von  $\cos x$  und  $\sin x$  erforderte nicht offensichtliche, geometrische Überlegungen. Mit den Reihenentwicklungen haben wir jetzt (endlich) eine analytische Formel.

**Beispiel 4.3** Wir betrachten für  $\alpha \in \mathbb{R}$  die Funktion  $f(x) = (1+x)^\alpha$ . Im Fall  $\alpha = n \in \mathbb{N}_0$  gilt nach dem Binomischen Lehrsatz

$$f(x) = \sum_{k=0}^n \binom{n}{k} x^k \quad \text{für alle } x \in \mathbb{R}.$$

Ab jetzt sei  $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$ . Dann gilt  $f^{(k)}(0) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k} \neq 0$  für  $k \in \mathbb{N}_0$ , also ist die Taylorreihe von  $f(x)$  mit Zentrum  $x_0 = 0$  die Binomialreihe

$$B_\alpha(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots$$

Um den Konvergenzradius der Reihe zu bestimmen, berechnen wir für die Glieder  $a_k = \binom{\alpha}{k} x^k$

$$\frac{|a_{k+1}|}{|a_k|} = \frac{|\alpha - k|}{k+1} |x| \rightarrow |x| \quad \text{mit } k \rightarrow \infty.$$

Im Fall  $|x| > 1$  sind die Glieder dem Betrag nach wachsend für große  $k$ , die Reihe kann also nicht konvergieren nach Satz 3.1. Für  $|x| < 1$  fallen die Glieder der Reihe dagegen mit geometrischer Rate, und zwar gilt für jedes  $q \in (|x|, 1)$

$$\frac{|a_{k+1}|}{|a_k|} \leq q < 1 \quad \text{für hinreichend große } k.$$

Die geometrische Reihe  $q^k$  liefert also eine Majorante, der Konvergenzradius ist  $R = 1$ . Die Frage ist nun, ob die Funktion  $f(x)$  auf dem Intervall  $(-1, 1)$  durch die Reihe dargestellt wird. Da die Abschätzung des Restglieds etwas kompliziert ist, vor allem im Fall  $x < 0$ , gehen wir anders vor und berechnen auf  $(-1, 1)$  mit Satz 3.11 die Ableitung

$$B'_\alpha(x) = \sum_{k=0}^{\infty} (k+1) \binom{\alpha}{k+1} x^k = \sum_{k=0}^{\infty} (k+1) \frac{\alpha-k}{k+1} \binom{\alpha}{k} x^k = \alpha B_\alpha(x) - x B'_\alpha(x),$$

das heißt  $B'_\alpha(x) = \frac{\alpha}{1+x} B_\alpha(x)$ . Es folgt mit  $g(x) = (1+x)^{-\alpha}$

$$(gB_\alpha)'(x) = g'(x)B_\alpha(x) + g(x)B'_\alpha(x) = g(x)B_\alpha(x) \left( -\frac{\alpha}{1+x} + \frac{\alpha}{1+x} \right) = 0.$$

Wegen  $B_\alpha(0) = 1 = f(0)$  ergibt sich die folgende Darstellung (Newton 1665)

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{für alle } x \in (-1, 1). \quad (4.8)$$

Oft wird die Näherung  $(1+x)^\alpha \approx 1 + \alpha x$  für  $|x| \ll 1$  benutzt, siehe Beispiel 4.1.

# Kapitel 5

## Lineare Algebra

### 1 Vektorräume

**Definition 1.1 (Vektorraum)** *Ein reeller Vektorraum (oder  $\mathbb{R}$ -Vektorraum) ist eine nicht-leere Menge  $V$ , auf der eine Vektoraddition und eine Skalarmultiplikation gegeben sind:*

$$\begin{aligned} V \times V &\rightarrow V & (\mathbf{v}, \mathbf{w}) &\mapsto \mathbf{v} + \mathbf{w} && \text{(Vektoraddition)} \\ \mathbb{R} \times V &\rightarrow V & (\lambda, \mathbf{v}) &\mapsto \lambda \mathbf{v} && \text{(Skalarmultiplikation)} \end{aligned}$$

Dabei wird verlangt:

(a)  $(V, +)$  ist eine additive Gruppe:

(a1)  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$  (Assoziativgesetz)

(a2)  $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$  (Kommutativgesetz)

(a3) Es gibt ein Element  $\mathbf{0} \in V$ , den Nullvektor, mit  $\mathbf{v} + \mathbf{0} = \mathbf{v}$  für alle  $\mathbf{v}$

(a4) Zu jedem  $\mathbf{v} \in V$  gibt es genau ein inverses Element  $-\mathbf{v}$ , also  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$

(b) Es sollen für alle  $\mathbf{v}, \mathbf{w} \in V$  und  $\lambda, \mu \in \mathbb{R}$  folgende Rechenregeln gelten:

(b1)  $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$

(b2)  $1 \cdot \mathbf{v} = \mathbf{v}$

(b3)  $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$

(b4)  $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ .

Es handelt sich hier um eine minimale Liste von Regeln, aus denen sich bei Bedarf alles weitere herleiten lässt. Zum Beispiel folgt, dass der Nullvektor  $\mathbf{0} \in V$  sowie das zu einem  $\mathbf{v} \in V$  inverse Element eindeutig bestimmt sind. Das zentrale Beispiel für uns ist der  $\mathbb{R}^n$ , mit der üblichen komponentenweisen Addition und Skalarmultiplikation:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \mathbf{x} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}, \quad \text{wobei } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in \mathbb{R}.$$

Im  $\mathbb{R}^n$  hatten wir die Regeln bereits in Kapitel 1, Abschnitt 4, formuliert. Wir schreiben die Vektoren nun ohne Pfeil, weil sie nicht unbedingt im Anschauungsraum liegen.

**Beispiel 1.1 (Funktionsraum)** Sei  $M$  eine beliebige Menge. Dann ist die Menge  $\mathcal{F}(M)$  aller Funktionen  $f : M \rightarrow \mathbb{R}$  ein Vektorraum, mit Addition und Skalarmultiplikation

$$\begin{aligned} f + g \in \mathcal{F}(M), \quad (f + g)(x) &= f(x) + g(x) \\ \lambda f \in \mathcal{F}(M), \quad (\lambda f)(x) &= \lambda f(x). \end{aligned}$$

Nullvektor ist die Nullfunktion,  $f(x) = 0$  für alle  $x \in M$ , und das zu  $f \in \mathcal{F}(M)$  inverse Element ist  $-f \in \mathcal{F}(M)$ ,  $(-f)(x) = -f(x)$ . Die Verifikation der Regeln ist sehr einfach.

In dieser Vorlesung geht es aber hauptsächlich um den  $\mathbb{R}^n$ . Die abstrakte Definition ist auch da erforderlich und nützlich: betrachte zum Beispiel eine Ebene  $E$  im  $\mathbb{R}^3$ , die den Nullpunkt enthält. Es ist nicht schwer zu sehen, dass  $E$  mit der Vektoraddition und Skalarmultiplikation des  $\mathbb{R}^3$  selbst ein Vektorraum ist (siehe unten). Im  $\mathbb{R}^3$  haben wir die Koordinaten  $\mathbf{x} = (x_1, x_2, x_3)$  bezüglich der Standardbasis  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ . Auf  $E$  können wir Koordinaten  $(y_1, y_2)$  definieren, indem wir zwei nichtparallele Vektoren  $\mathbf{w}_1, \mathbf{w}_2 \in E$  wählen und jedes  $\mathbf{w} \in E$  in der Form  $\mathbf{w} = y_1 \mathbf{w}_1 + y_2 \mathbf{w}_2$  schreiben. Jedes  $\mathbf{w} \in E$  wird so mit einem Punkt  $(y_1, y_2)$  im  $\mathbb{R}^2$  identifiziert. Jedoch ist diese Identifikation nicht kanonisch, eine andere Wahl  $\mathbf{w}'_1, \mathbf{w}'_2$  führt zu anderen Koordinaten  $(y'_1, y'_2) \in \mathbb{R}^2$ . Ein Vorteil der allgemeinen Definition des Vektorraums ist, dass sie nicht auf die Wahl von Koordinaten Bezug nimmt.

**Definition 1.2 (Untervektorraum)** Sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Eine Menge  $X \subset V$  heißt Untervektorraum, wenn  $0 \in X$  und wenn folgende Regeln gelten:

$$\begin{aligned} \mathbf{x}, \mathbf{y} \in X &\Rightarrow \mathbf{x} + \mathbf{y} \in X \\ \mathbf{x} \in X, \lambda \in \mathbb{R} &\Rightarrow \lambda \mathbf{x} \in X. \end{aligned}$$

Bemerkung.  $X$  ist dann selbst ein  $\mathbb{R}$ -Vektorraum, mit den Operationen  $+$  und  $\cdot$  von  $V$ .

**Beispiel 1.2** Sei  $X$  die Menge aller Lösungen  $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$  der Gleichung <sup>1</sup>

$$x_1 + 2x_2 - x_3 = 0.$$

Der Nullvektor gehört dazu, und mit  $\mathbf{x}, \mathbf{y}$  sind auch  $\mathbf{x} + \mathbf{y}$  und  $\lambda \mathbf{x}$  Lösungen:

$$\begin{aligned} x_1 + y_1 + 2(x_2 + y_2) - (x_3 + y_3) &= x_1 + 2x_2 - x_3 + y_1 + 2y_2 - y_3 = 0, \\ \lambda x_1 + 2\lambda x_2 - \lambda x_3 &= \lambda(x_1 + 2x_2 - x_3) = 0. \end{aligned}$$

Also ist  $X$  ein Untervektorraum des  $\mathbb{R}^3$ . Dagegen ist die Lösungsmenge  $Y$  der Gleichung

$$x_1 + 2x_2 - x_3 = 3$$

kein Untervektorraum. Das ergibt sich schon daraus, dass der Nullvektor nicht dazugehört. Auch lösen  $\mathbf{x} + \mathbf{y}$  sowie  $\lambda \mathbf{x}$  (außer für  $\lambda = 1$ ) nicht wieder die Gleichung.

**Beispiel 1.3** Für jedes  $\mathbf{v} \in V$  ist die Menge

$$\mathbb{R} \cdot \mathbf{v} = \{\lambda \mathbf{v} : \lambda \in \mathbb{R}\}$$

ein Untervektorraum.

---

<sup>1</sup>In eingerückten Formeln schreiben wir Vektoren des  $\mathbb{R}^n$  in der Regel als Spaltenvektoren. Im Text verwenden wir aus Platzgründen Zeilenvektoren.

**Beispiel 1.4 ( $C^k(I)$ -Vektorräume)** Sei  $I \subset \mathbb{R}$  ein Intervall. Dann ist der Raum der stetigen Funktionen  $C^0(I)$  ein Untervektorraum des Raums  $\mathcal{F}(I)$  aller Funktionen auf  $I$ . Denn die Nullfunktion ist stetig, und nach Satz 5.2 in Kapitel 2 gilt

$$\begin{aligned} f, g \text{ stetig auf } I &\Rightarrow f + g \text{ stetig auf } I \\ \lambda \in \mathbb{R}, f \text{ stetig auf } I &\Rightarrow \lambda f \text{ stetig auf } I. \end{aligned}$$

Nach Satz 2.1 in Kapitel 3 ist die Summe von differenzierbaren Funktionen wieder differenzierbar mit  $(f + g)'(x) = f'(x) + g'(x)$ , ebenso  $(\lambda f)'(x) = \lambda f'(x)$ . Es folgt, dass auch der Raum  $C^1(I)$  der stetig differenzierbaren Funktionen ein Unterraum ist, sowie allgemeiner  $C^k(I)$  für  $k \in \mathbb{N}$  und  $C^\infty(I)$ . Wir haben also eine Sequenz von Unterräumen

$$C^\infty(I) \subset \dots \subset C^k(I) \subset \dots \subset C^1(I) \subset C^0(I) \subset \mathcal{F}(I).$$

**Beispiel 1.5 (Lösungsraum einer linearen Differentialgleichung)** Sei  $I \subset \mathbb{R}$  ein Intervall, und es seien  $a_0, a_1 \in C^0(I)$  gegeben. Betrachte

$$X = \{f \in C^2(I) : f''(t) + a_1(t)f'(t) + a_0(t)f(t) = 0 \text{ für alle } t \in I\}.$$

Lösen  $f$  und  $g$  die Differentialgleichung, so auch  $f + g$  und  $\lambda f$ . Daher ist  $X$  ein Untervektorraum von  $C^2(I)$ .

**Beispiel 1.6 (Durchschnitt von Unterräumen)** Seien  $X_i, i \in I$ , Unterräume des Vektorraums  $V$ . Dann ist

$$X = \bigcap_{i \in I} X_i \subset V$$

wieder ein Unterraum von  $V$ . Denn es ist  $\mathbf{0} \in X_i$  für alle  $i \in I$ , und mit  $\mathbf{x}, \mathbf{y} \in X_i$  für alle  $i \in I$  folgt auch  $\mathbf{x} + \mathbf{y} \in X_i$  sowie  $\lambda \mathbf{x} \in X_i$  für alle  $i \in I$ .

Im Gegensatz zum Durchschnitt ist die Vereinigung  $X \cup Y$  von zwei Unterräumen  $X, Y$  im allgemeinen kein Unterraum. Daher folgender Begriff.

**Satz 1.1 (Summe von Unterräumen)** Sei  $V$  ein  $\mathbb{R}$ -Vektorraum, und  $X, Y \subset V$  seien Unterräume. Dann ist

$$X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y\}$$

wieder ein Unterraum von  $V$ .

BEWEIS: Nach Voraussetzung enthalten  $X, Y$  beide den Nullvektor, also auch  $X + Y$ . Sind  $\mathbf{z}_1 = \mathbf{x}_1 + \mathbf{y}_1, \mathbf{z}_2 = \mathbf{x}_2 + \mathbf{y}_2 \in X + Y$ , so folgt

$$\begin{aligned} \mathbf{z}_1 + \mathbf{z}_2 &= (\mathbf{x}_1 + \mathbf{x}_2) + (\mathbf{y}_1 + \mathbf{y}_2) \in X + Y, \\ \lambda \mathbf{z}_1 &= \lambda \mathbf{x}_1 + \lambda \mathbf{y}_1 \in X + Y. \end{aligned}$$

□

**Definition 1.3 (direkte Summe)** Zwei Unterräume  $X, Y \subset V$  heißen komplementär, bzw.  $V$  heißt direkte Summe von  $X$  und  $Y$  (Notation:  $V = X \oplus Y$ ), wenn gilt:

$$X + Y = V \quad \text{und} \quad X \cap Y = \{0\}.$$

**Beispiel 1.7** Betrachte im  $\mathbb{R}^2$  die Geraden

$$X = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad Y = \mathbb{R} \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

Wir behaupten  $\mathbb{R}^2 = X \oplus Y$ , falls  $b \neq 0$ . Ist  $\mathbf{x} \in X \cap Y$ , so gibt es  $\lambda, \mu \in \mathbb{R}$  mit

$$\mathbf{x} = \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mu \begin{pmatrix} a \\ b \end{pmatrix} \quad \Rightarrow \quad \mu = 0 \quad \Rightarrow \quad \mathbf{x} = 0.$$

Weiter gilt für  $\mathbf{x} \in \mathbb{R}^2$  die Darstellung

$$\mathbf{x} = \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{mit} \quad \lambda = x_1 - \frac{x_2 a}{b}, \quad \mu = \frac{x_2}{b}.$$

Dies zeigt  $\mathbb{R}^2 = X \oplus Y$ .

**Definition 1.4 (erzeugter Untervektorraum)** Sei  $\emptyset \neq M \subset V$ . Wir definieren den von  $M$  erzeugten Untervektorraum oder Span von  $M$  (Notation:  $\text{Span } M$ ) als die Menge aller  $\mathbf{x} \in V$  der Form

$$\mathbf{x} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k \quad \text{wobei} \quad \mathbf{x}_i \in M, \lambda_i \in \mathbb{R}, k \in \mathbb{N}. \quad (1.1)$$

Eine (endliche) Summe wie in (1.1) heißt Linearkombination von  $\mathbf{x}_1, \dots, \mathbf{x}_k$ .

Wir wollen verifizieren, dass  $\text{Span } M$  ein Untervektorraum ist. Den Nullvektor kriegen wir mit  $\mathbf{0} = 0 \cdot \mathbf{x} \in \text{Span } M$  für irgendein  $\mathbf{x} \in M \neq \emptyset$ . Seien  $\mathbf{x}, \mathbf{y} \in \text{Span } M$ . Wir können annehmen, dass  $\mathbf{x}, \mathbf{y}$  Linearkombinationen derselben Menge  $\mathbf{x}_1, \dots, \mathbf{x}_k \in M$  sind, sonst vereinigen wir die beiden Mengen und setzen die zusätzlichen Koeffizienten gleich Null. Sei also  $\mathbf{x} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k$  und  $\mathbf{y} = \mu_1 \mathbf{x}_1 + \dots + \mu_k \mathbf{x}_k$ . Dann folgt

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (\lambda_1 + \mu_1) \mathbf{x}_1 + \dots + (\lambda_k + \mu_k) \mathbf{x}_k \in \text{Span } M, \\ \alpha \mathbf{x} &= (\alpha \lambda_1) \mathbf{x}_1 + \dots + (\alpha \lambda_k) \mathbf{x}_k \in \text{Span } M. \end{aligned}$$

Damit sind die Unterraumkriterien gezeigt. Enthält ein Unterraum  $X \subset V$  die Menge  $M$ , so enthält er schon ganz  $\text{Span } M \subset X$ . Somit ist  $\text{Span } M$  der Durchschnitt aller Unterräume, die  $M$  enthalten. Verwenden wir diese Charakterisierung im Fall  $M = \emptyset$ , so folgt  $\text{Span } \emptyset = \{0\}$  als sinnvolle Festsetzung.

**Definition 1.5 (endlichdimensional)** Eine Vektorraum  $V$  heißt endlichdimensional, wenn es eine endliche Menge  $M \subset V$  gibt mit  $V = \text{Span } M$ . Andernfalls heißt  $V$  unendlichdimensional.

**Beispiel 1.8 (Polynomraum)** Eine Funktion  $P : \mathbb{R} \rightarrow \mathbb{R}$  heißt Polynom vom Grad höchstens  $n$ , wenn es Koeffizienten  $a_0, a_1, \dots, a_n \in \mathbb{R}$  gibt mit

$$P(x) = a_0 + a_1 x + \dots + a_n x^n \quad \text{für alle } x \in \mathbb{R}.$$

Ist dabei  $a_n \neq 0$ , so hat  $P(x)$  genau den Grad  $n$ . Sind  $P, Q$  Polynome vom Grad höchstens  $n$  mit Koeffizienten  $a_i, b_i$ , so sind  $P + Q$  und  $\lambda P$  ebenfalls Polynome vom Grad höchstens  $n$ , mit Koeffizienten  $a_i + b_i$  bzw.  $\lambda a_i$ . Somit ist die Menge  $\mathbf{P}_n(\mathbb{R})$  aller Polynome vom Grad

höchstens  $n$  ein Untervektorraum von  $\mathcal{F}(\mathbb{R})$ . Jedes Element von  $\mathbf{P}_n(\mathbb{R})$  ist Linearkombination der Monome  $x^i$  mit  $i = 0, \dots, n$ , somit ist der Raum endlichdimensional. Der Raum aller reellen Polynome

$$\mathbf{P}(\mathbb{R}) = \bigcup_{n=0}^{\infty} \mathbf{P}_n(\mathbb{R})$$

ist ebenfalls ein Unterraum von  $\mathcal{F}(\mathbb{R})$ . Denn die Summe von zwei Polynomen vom Grad höchstens  $m$  bzw.  $n$  ist Polynom vom Grad höchstens  $\max(m, n)$ . Der Raum  $\mathbf{P}(\mathbb{R})$  ist jedoch nicht endlichdimensional: angenommen  $\mathbf{P}(\mathbb{R})$  ist Span der Polynome  $P_i(x) \in \mathbf{P}_{n_i}(\mathbb{R})$  mit  $i = 1, \dots, r$ . Ist  $n > \max_{i=1, \dots, r} n_i$ , so ist  $P(x) = x^n$  nicht als Linearkombination der  $P_i$  darstellbar. Denn sonst gibt es  $\lambda_i \in \mathbb{R}$  mit

$$x^n - \sum_{i=1}^r \lambda_i P_i(x) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Das Polynom hat den Grad  $n$ , kann also nach Lemma 2.3 in Kapitel 2 höchstens  $n$  Nullstellen haben, ein Widerspruch.

**Definition 1.6 (lineare Abhängigkeit)** Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$  heißen linear abhängig, wenn es  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  gibt, die nicht alle Null sind und so dass gilt:

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k = \mathbf{0}.$$

Andernfalls heißen  $\mathbf{v}_1, \dots, \mathbf{v}_k$  linear unabhängig.

**Bemerkung.** Um nachzuweisen, dass  $\mathbf{v}_1, \dots, \mathbf{v}_k$  linear unabhängig sind, muss folgende Implikation gezeigt werden:

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k = \mathbf{0} \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_k = 0.$$

**Beispiel 1.9** Sind die nachfolgenden Vektoren  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$  linear abhängig?

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}.$$

Dazu ist zu klären, ob die Gleichung  $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \lambda_3 \mathbf{v}_3 = \mathbf{0}$  von Null verschiedene Lösungen  $\lambda_1, \lambda_2, \lambda_3$  hat. Ausgeschrieben lautet die Gleichung

$$\begin{array}{rcl} \lambda_1 & & + \lambda_3 = 0 \\ 2\lambda_1 & + 2\lambda_2 & + 4\lambda_3 = 0 \\ 3\lambda_1 & + \lambda_2 & + 4\lambda_3 = 0 \end{array}$$

Das System hat die Lösung  $(\lambda_1, \lambda_2, \lambda_3) = (1, 1, -1)$ , mit anderen Worten gilt  $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$ . Also sind die Vektoren linear abhängig.

**Satz 1.2 (Austauschsatz)** Sei  $V$  ein Vektorraum mit  $V = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ . Sind  $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$  linear unabhängig, so gilt nach Umnummerierung der  $\mathbf{v}_j$

$$V = \text{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_m\}, \quad \text{insbesondere } k \leq m.$$

BEWEIS: Durch Induktion: es sei schon nach Umnummerierung

$$V = \text{Span} \{ \mathbf{x}_1, \dots, \mathbf{x}_\ell, \mathbf{v}_{\ell+1}, \dots, \mathbf{v}_m \} \quad \text{für ein } \ell \in \{0, \dots, k-1\}.$$

Der Induktionsanfang ist der Fall  $\ell = 0$ . Schreibe  $\mathbf{x}_{\ell+1}$  als Linearkombination

$$\mathbf{x}_{\ell+1} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_\ell \mathbf{x}_\ell + \lambda_{\ell+1} \mathbf{v}_{\ell+1} + \dots + \lambda_m \mathbf{v}_m.$$

Es gibt ein  $j \in \{\ell+1, \dots, m\}$  mit  $\lambda_j \neq 0$  (insbesondere ist  $\ell+1 \leq m$ ), sonst wären  $\mathbf{x}_1, \dots, \mathbf{x}_{\ell+1}$  linear abhängig.  $\mathbf{v}_j$  ist dann Linearkombination der verbleibenden Vektoren, und es folgt

$$V = \text{Span} \{ \mathbf{x}_1, \dots, \mathbf{x}_{\ell+1}, \mathbf{v}_{\ell+1}, \dots, \cancel{\mathbf{v}_j}, \dots, \mathbf{v}_m \}.$$

Nach Umnummerierung ist der Induktionsschritt vollzogen. Der Prozess stoppt erst dann, wenn alle  $\mathbf{x}_\ell$  eingefügt sind. Im letzten Schritt ist  $k = \ell + 1 \leq m$ .  $\square$

**Definition 1.7 (Basis)** Sei  $V$  ein endlichdimensionaler Vektorraum. Eine nummerierte Menge  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  heißt Basis von  $V$ , wenn gilt:

- (1)  $\text{Span} \{ \mathbf{v}_1, \dots, \mathbf{v}_n \} = V$
- (2)  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sind linear unabhängig.

Änderung der Reihenfolge der Vektoren ergibt eine andere Basis.

**Satz 1.3 (Existenz einer Basis und Dimension)** Jeder endlichdimensionale Vektorraum  $V$  hat eine Basis, und je zwei Basen von  $V$  haben gleichviele Elemente. Diese Anzahl nennen wir die Dimension  $\dim V$  von  $V$ .

BEWEIS: Da  $V$  endlichdimensional ist, gilt  $V = \text{Span} \{ \mathbf{v}_1, \dots, \mathbf{v}_m \}$  für ein  $m \in \mathbb{N}$ . Nach dem Austauschsatz hat eine linear unabhängige Menge höchstens  $m$  Elemente. Seien  $\mathbf{x}_1, \dots, \mathbf{x}_k$  maximal viele linear unabhängige Vektoren. Wir zeigen, dass dies eine Basis ist. Für jedes  $\mathbf{v} \in V$  ist  $\mathbf{v}, \mathbf{x}_1, \dots, \mathbf{x}_k$  linear abhängig, das heißt es gibt  $\lambda, \lambda_1, \dots, \lambda_k$  nicht alle Null mit

$$\lambda \mathbf{v} + \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k = \mathbf{0}.$$

Es muss  $\lambda \neq 0$  sein, sonst wären  $\mathbf{x}_1, \dots, \mathbf{x}_k$  linear abhängig. Also folgt

$$\mathbf{v} = -\frac{1}{\lambda} (\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k) \in \text{Span} \{ \mathbf{x}_1, \dots, \mathbf{x}_k \}.$$

Also gilt  $V = \text{Span} \{ \mathbf{x}_1, \dots, \mathbf{x}_k \}$ , und  $\mathbf{x}_1, \dots, \mathbf{x}_k$  ist eine Basis.  $\square$

**Beispiel 1.10** Im  $\mathbb{R}^n$  haben wir die Standardbasis

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{für } i = 1, \dots, n.$$

Dabei steht die 1 an der  $i$ -ten Stelle. Es ist klar dass  $\mathbb{R}^n = \text{Span}\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , denn

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \quad \text{für jedes } \mathbf{x} \in \mathbb{R}^n.$$

Andererseits sind die  $\mathbf{e}_1, \dots, \mathbf{e}_n$  linear unabhängig, denn

$$\mathbf{0} = \sum_{i=1}^n \lambda_i \mathbf{e}_i = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \Rightarrow \lambda_i = 0 \text{ für } i = 1, \dots, n.$$

Die Dimension von  $\mathbb{R}^n$  ist somit gleich  $n$ , wie es sein sollte.

**Beispiel 1.11** Sei  $X = \{\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 : 3x_1 + x_2 + x_3 = 0\}$ . Wir behaupten, dass folgende Vektoren eine Basis von  $X$  bilden:

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}.$$

Zunächst gilt  $\mathbf{v}_1, \mathbf{v}_2 \in X$ , und die Vektoren sind linear unabhängig:

$$\mathbf{0} = \lambda \mathbf{v}_1 + \mu \mathbf{v}_2 = \begin{pmatrix} -\lambda - \mu \\ 3\lambda \\ 3\mu \end{pmatrix} \Rightarrow \lambda = \mu = 0.$$

Weiter gilt für  $\mathbf{x} \in X$  die Darstellung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} \quad \text{für } \lambda = \frac{x_2}{3} \text{ und } \mu = \frac{x_3}{3}.$$

Beachte dabei  $x_1 = -(x_2 + x_3)/3$  wegen  $\mathbf{x} \in X$ . Also ist  $\mathbf{v}_1, \mathbf{v}_2$  eine Basis.

**Beispiel 1.12** Die Monome  $P_i(x) = x^i$ ,  $i = 0, \dots, n$ , sind Basis des Raums  $\mathbf{P}_n(\mathbb{R})$  der Polynome auf  $\mathbb{R}$  vom Grad höchstens  $n$ . Denn zu  $P \in \mathbf{P}_n(\mathbb{R})$  gibt es  $a_0, \dots, a_n \in \mathbb{R}$  mit

$$P = \sum_{i=0}^n a_i P_i \Leftrightarrow P(x) = a_0 + a_1 x + \dots + a_n x^n \quad \text{für alle } x \in \mathbb{R}.$$

Außerdem sind die Monome linear unabhängig, denn sei

$$a_0 + a_1 x + \dots + a_n x^n = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Dann verschwinden alle Koeffizienten, sonst hätte das Polynom höchstens  $n$  Nullstellen.

**Satz 1.4 (Existenz eines Komplements)** Sei  $V$  ein endlichdimensionaler Vektorraum. Dann hat jeder Unterraum  $X \subset V$  eine Komplement  $Y$ , also  $V = X \oplus Y$ . Insbesondere gilt

$$\dim V = \dim X + \dim Y.$$

BEWEIS: Variante des Beweises von Satz 1.3. Wähle erst eine Basis  $\mathbf{x}_1, \dots, \mathbf{x}_k$  von  $X$ . Diese bekommen wir als maximale linear unabhängige Menge in  $X$ . Nach dem Austauschatz ist die Zahl der Elemente durch  $\dim V < \infty$  beschränkt. Wähle nun eine maximale, linear unabhängige Menge  $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$  in  $V$ , und setze  $Y = \text{Span}\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ . Man sieht dann leicht  $V = X \oplus Y$ . Da  $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$  Basis von  $Y$ , folgt die Formel für die Dimensionen.  $\square$

Der folgende Satz verallgemeinert diese Situation dahingehend, dass die Räume nicht notwendig komplementär sind.

**Satz 1.5 (Dimension einer Summe)** *Für endlichdimensionale Unterräume  $X, Y$  eines Vektorraums  $V$  gilt die Formel*

$$\dim(X + Y) = \dim X + \dim Y - \dim(X \cap Y).$$

BEWEIS: Seien  $\mathbf{x}_1, \dots, \mathbf{x}_m$  und  $\mathbf{y}_1, \dots, \mathbf{y}_n$  Basen von  $X$  bzw.  $Y$ . Wähle weiter eine Basis  $\mathbf{v}_1, \dots, \mathbf{v}_k$  von  $X \cap Y$ . Nach dem Austauschatz können wir annehmen, dass

$$\mathbf{x}_i = \mathbf{y}_i = \mathbf{v}_i \quad \text{für } i = 1, \dots, k.$$

Wir zeigen, dass  $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_m, \mathbf{y}_{k+1}, \dots, \mathbf{y}_n$  eine Basis von  $X + Y$  ist. Daraus folgt

$$\dim(X + Y) = k + (m - k) + (n - k) = m + n - k = \dim X + \dim Y - \dim(X \cap Y).$$

Es ist klar, dass die Vektoren  $X + Y$  aufspannen, zu zeigen ist die Unabhängigkeit. Sei

$$\mathbf{v} + \mathbf{x} + \mathbf{y} = 0 \quad \text{wobei } \mathbf{v} = \sum_{i=1}^k \alpha_i \mathbf{v}_i, \mathbf{x} = \sum_{j=k+1}^m \lambda_j \mathbf{x}_j, \mathbf{y} = \sum_{\ell=k+1}^n \mu_\ell \mathbf{y}_\ell.$$

Es folgt  $\mathbf{x} = -\mathbf{v} - \mathbf{y} \in Y$ , also ist  $\mathbf{x} \in X \cap Y$ . Es gibt dann  $\beta_1, \dots, \beta_k$  mit

$$\sum_{i=1}^k \beta_i \mathbf{v}_i = \mathbf{x} = \sum_{j=k+1}^m \lambda_j \mathbf{x}_j.$$

Da  $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_m$  Basis von  $X$  ist, folgt  $\lambda_j = 0$  (und  $\beta_i = 0$ ). Analog sehen wir  $\mu_\ell = 0$  und schließlich  $\alpha_i = 0$ . Damit ist die lineare Unabhängigkeit bewiesen.  $\square$

## 2 Lineare Abbildungen und Matrizen

**Definition 2.1 (lineare Abbildung)** *Seien  $V, W$   $\mathbb{R}$ -Vektorräume.  $A : V \rightarrow W$  heißt lineare Abbildung (Notation:  $A \in L(V, W)$ ), wenn gilt:*

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= A(\mathbf{x}) + A(\mathbf{y}) && \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \\ A(\lambda \mathbf{x}) &= \lambda A(\mathbf{x}) && \text{für alle } \mathbf{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}. \end{aligned}$$

Bei linearen Abbildungen ist die Notation  $A\mathbf{x}$  statt  $A(\mathbf{x})$  üblich. Die Linearität kann auch so ausgedrückt werden, dass  $A$  mit beliebigen Linearkombinationen vertauscht:

$$A(\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k) = \lambda_1 A\mathbf{x}_1 + \dots + \lambda_k A\mathbf{x}_k. \quad (2.2)$$

**Beispiel 2.1 (Drehung im  $\mathbb{R}^2$ )** Die Drehung im  $\mathbb{R}^2$  um den Winkel  $\alpha$  ist

$$D_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2, D_\alpha \mathbf{x} = \begin{pmatrix} (\cos \alpha) x_1 - (\sin \alpha) x_2 \\ (\sin \alpha) x_1 + (\cos \alpha) x_2 \end{pmatrix}.$$

Es ist klar, dass  $D_\alpha(\mathbf{x} + \mathbf{y}) = D_\alpha \mathbf{x} + D_\alpha \mathbf{y}$  sowie  $D_\alpha(\lambda \mathbf{x}) = \lambda D_\alpha(\mathbf{x})$ . Wir schreiben auch

$$D_\alpha \mathbf{x} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

**Definition 2.2** Für eine lineare Abbildung  $A : V \rightarrow W$  definieren wir:

- (1) Der Kern von  $A$  ist  $\ker A = \{v \in V : Av = 0\}$ ,
- (2) Das Bild von  $A$  ist  $\text{Bild } A = \{Av : v \in V\}$ .

Überlegen Sie, dass  $\ker A$  ein Unterraum von  $V$  ist, und  $\text{Bild } A$  ein Unterraum von  $W$ . Die Zahl  $\text{rang } A = \dim \text{Bild } A$  heißt Rang von  $A$ .

**Beispiel 2.2 (Projektion)** Sei  $V = X \oplus Y$ . Dann hat jedes  $\mathbf{v} \in V$  eine Zerlegung

$$\mathbf{v} = \mathbf{x} + \mathbf{y} \quad \text{mit } \mathbf{x} \in X, \mathbf{y} \in Y.$$

Die Zerlegung ist eindeutig, denn ist  $\mathbf{x}_1 + \mathbf{y}_1 = \mathbf{x}_2 + \mathbf{y}_2$  mit  $\mathbf{x}_{1,2} \in X$  und  $\mathbf{y}_{1,2} \in Y$ , so folgt

$$\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{y}_2 - \mathbf{y}_1 \in X \cap Y = \{0\},$$

also  $\mathbf{x}_1 = \mathbf{x}_2$  und  $\mathbf{y}_1 = \mathbf{y}_2$ . Wir erhalten damit eine wohldefinierte Abbildung

$$P : V \rightarrow X, P\mathbf{v} = \mathbf{x} \quad \text{für } \mathbf{v} = \mathbf{x} + \mathbf{y}.$$

Für  $\mathbf{v}_{1,2} = \mathbf{x}_{1,2} + \mathbf{y}_{1,2}$  ergibt sich

$$P(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{x}_1 + \mathbf{x}_2 = P\mathbf{v}_1 + P\mathbf{v}_2 \quad \text{und} \quad P(\lambda \mathbf{v}_1) = \lambda \mathbf{x}_1 = \lambda P\mathbf{v}_1.$$

Folglich ist  $P$  eine lineare Abbildung. Es gilt  $\text{Bild } P = X$ , genauer  $P|_X = \text{Id}_X$ , und  $\ker P = Y$ , denn für  $\mathbf{v} = \mathbf{x} + \mathbf{y}$  ist  $P\mathbf{v} = \mathbf{0}$  genau wenn  $\mathbf{x} = \mathbf{0}$ . Die Abbildung  $P$  heißt Projektion auf  $X$  bezüglich der Zerlegung  $V = X \oplus Y$ . Beachten Sie, dass  $P$  auch von der Wahl des Komplements  $Y$  abhängt.

**Lemma 2.1** Für eine lineare Abbildung  $A : V \rightarrow W$  gilt:

$$\begin{aligned} A \text{ injektiv} &\Leftrightarrow \ker A = \{\mathbf{0}\} \\ A \text{ surjektiv} &\Leftrightarrow \text{Bild } A = W. \end{aligned}$$

BEWEIS: Es gilt

$$A\mathbf{v}_1 = A\mathbf{v}_2 \Leftrightarrow A(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0} \Leftrightarrow \mathbf{v}_1 - \mathbf{v}_2 \in \ker A.$$

Ist  $\ker A = \{\mathbf{0}\}$ , so folgt aus  $A\mathbf{v}_1 = A\mathbf{v}_2$  auch  $\mathbf{v}_1 = \mathbf{v}_2$ , das heißt  $A$  ist injektiv. Ist umgekehrt  $A$  injektiv, so folgt  $A\mathbf{v} \neq A \cdot \mathbf{0} = \mathbf{0}$  für  $\mathbf{v} \neq \mathbf{0}$ , also  $\ker A = \{\mathbf{0}\}$ . Die zweite Aussage ist trivial.  $\square$

**Satz 2.1 (Dimensionsformel)** Sei  $A \in L(V, W)$  mit  $\dim V < \infty$ . Dann gilt

$$\dim V = \dim \ker A + \dim \text{Bild } A.$$

BEWEIS: Sei  $X$  ein Komplement von  $\ker A$ , siehe Satz 1.4, und  $\mathbf{x}_1, \dots, \mathbf{x}_k$  sei eine Basis von  $X$ . Wir zeigen, dass  $A\mathbf{x}_1, \dots, A\mathbf{x}_k$  eine Basis von  $\text{Bild } A$  ist. Jedes  $\mathbf{v} \in V$  hat eine Darstellung

$$\mathbf{v} = \mathbf{x}_0 + \sum_{i=1}^k \lambda_i \mathbf{x}_i \quad \text{mit } \mathbf{x}_0 \in \ker A, \lambda_i \in \mathbb{R}.$$

Es folgt

$$A\mathbf{v} = \sum_{i=1}^k \lambda_i A\mathbf{x}_i.$$

Dies beweist  $\text{Bild } A = \text{Span} \{A\mathbf{x}_1, \dots, A\mathbf{x}_k\}$ . Weiter schließen wir

$$\begin{aligned} \lambda_1 A\mathbf{x}_1 + \dots + \lambda_k A\mathbf{x}_k = \mathbf{0} &\Rightarrow A(\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k) = \mathbf{0} \\ &\Rightarrow \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k \in \ker A \cap X \\ &\Rightarrow \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k = \mathbf{0} \\ &\Rightarrow \lambda_i = 0 \quad \text{für } i = 1, \dots, k. \end{aligned}$$

Die  $A\mathbf{x}_1, \dots, A\mathbf{x}_k$  sind damit linear unabhängig, also eine Basis von  $\text{Bild } A$ . Es folgt

$$\dim V = \dim \ker A + \dim X = \dim \ker A + \dim \text{Bild } A.$$

□

**Satz 2.2** Jede Matrix  $a \in \mathbb{R}^{m \times n}$  definiert eine lineare Abbildung  $A \in L(\mathbb{R}^n, \mathbb{R}^m)$  durch

$$(A\mathbf{x})_i = \sum_{j=1}^n a_{ij} x_j \quad \text{für } i = 1, \dots, m, \quad (2.3)$$

oder ausführlicher

$$A\mathbf{x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}. \quad (2.4)$$

Die hierdurch gegebene Zuordnung zwischen  $\mathbb{R}^{m \times n}$  und  $L(\mathbb{R}^n, \mathbb{R}^m)$  ist bijektiv.

BEWEIS: Durch (2.3) ist eine lineare Abbildung definiert, denn es gilt

$$\begin{aligned} A(\mathbf{x} + \mathbf{y})_i &= \sum_{j=1}^n a_{ij}(x_j + y_j) = \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n a_{ij}y_j = (A\mathbf{x})_i + (A\mathbf{y})_i \\ A(\lambda\mathbf{x})_i &= \sum_{j=1}^n a_{ij}(\lambda x_j) = \lambda \sum_{j=1}^n a_{ij}x_j = \lambda(A\mathbf{x})_i. \end{aligned}$$

Wir können die lineare Abbildung alternativ wie folgt hinschreiben:

$$A\mathbf{x} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}. \quad (2.5)$$

Somit ist  $A\mathbf{x}$  Linearkombination der Spaltenvektoren von  $A$  mit den Koeffizienten  $x_j$  für  $j = 1, \dots, n$ . Es bezeichne nun  $\mathbf{e}_j$ ,  $j = 1, \dots, n$ , die Standardbasis des  $\mathbb{R}^n$ . Sei  $A \in L(\mathbb{R}^n, \mathbb{R}^m)$  gegeben. Dann definieren wir eine Matrix  $a = (a_{ij}) \in \mathbb{R}^{m \times n}$  durch

$$A\mathbf{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{für } j = 1, \dots, n.$$

Wir schreiben also den Vektor  $A\mathbf{e}_j$  in die  $j$ -te Spalte von  $a$ . Ist  $a = (a_{ij}) \in \mathbb{R}^{m \times n}$  gegeben und  $A$  die zugehörige lineare Abbildung, so folgt aus (2.5)

$$A\mathbf{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Das ist genau die  $j$ -te Spalte von  $a$ , wir erhalten also die Matrix  $a$  zurück. Ist umgekehrt  $A \in L(\mathbb{R}^n, \mathbb{R}^m)$  gegeben und  $a \in \mathbb{R}^{m \times n}$  die zugehörige Matrix, so folgt für  $\mathbf{x} \in \mathbb{R}^n$

$$\begin{aligned} x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} &= x_1 A\mathbf{e}_1 + \dots + x_n A\mathbf{e}_n \\ &= A(x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n) = A\mathbf{x}. \end{aligned}$$

Wir erhalten also nun die Abbildung  $A$  zurück, so dass die Zuordnungen zueinander invers und damit bijektiv sind.  $\square$

Im folgenden werden wir zwischen  $A \in \mathbb{R}^{m \times n}$  und  $A \in L(\mathbb{R}^m, \mathbb{R}^n)$  nicht mehr unterscheiden, allerdings schreiben wir die Matrix als  $A = (a_{ij})$ , das scheint üblich zu sein. Es ist auf Folgendes hinzuweisen: sind  $\mathbf{v}_1, \dots, \mathbf{v}_n$  und  $\mathbf{w}_1, \dots, \mathbf{w}_m$  irgendwelche Basen des  $\mathbb{R}^n$  bzw.  $\mathbb{R}^m$ , so kann man jeder linearen Abbildung  $A \in L(\mathbb{R}^n, \mathbb{R}^m)$  auch eine Matrix zuordnen durch

$$A\mathbf{v}_j = \sum_{i=1}^m \tilde{a}_{ij} \mathbf{w}_i.$$

Diese Matrix ist im allgemeinen nicht dieselbe wie für die Standardbasen. Unsere Identifikation zwischen  $\mathbb{R}^{m \times n}$  und  $L(\mathbb{R}^n, \mathbb{R}^m)$  verwendet stillschweigend, dass die Standardbasen zugrunde liegen.

**Satz 2.3 (Orthogonales Komplement)** *Es bezeichne  $\langle \mathbf{x}, \mathbf{y} \rangle$  das Standardskalarprodukt im  $\mathbb{R}^n$ . Für einen Unterraum  $X \subset \mathbb{R}^n$  definiere*

$$X^\perp = \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ für alle } \mathbf{x} \in X\}.$$

*Dann ist  $X^\perp$  auch ein Unterraum von  $\mathbb{R}^n$  und es gilt  $\mathbb{R}^n = X \oplus X^\perp$ .*

BEWEIS: Es ist leicht zu sehen, dass  $X^\perp$  ein Untervektorraum ist. Wähle eine Basis  $\mathbf{x}_1, \dots, \mathbf{x}_k$  von  $X$  und betrachte die lineare Abbildung

$$F: \mathbb{R}^n \rightarrow \mathbb{R}^k, F\mathbf{v} = \begin{pmatrix} \langle \mathbf{v}, \mathbf{x}_1 \rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{x}_k \rangle \end{pmatrix}.$$

Sei  $\mathbf{x} \in X$  mit  $F\mathbf{x} = \mathbf{0}$ , also  $\langle \mathbf{x}, \mathbf{x}_j \rangle = 0$  für  $j = 1, \dots, k$ . Da  $\mathbf{x}$  Linearkombination der  $\mathbf{x}_j$ , folgt  $\langle \mathbf{x}, \mathbf{x} \rangle = 0$  und hieraus  $\mathbf{x} = \mathbf{0}$ . Somit ist  $F|_X: X \rightarrow \mathbb{R}^k$  injektiv, und wegen  $\dim X = \dim \mathbb{R}^k$  auch surjektiv. Nun gilt  $X^\perp = \ker F$ , und die Dimensionsformel für  $F$  selbst ergibt

$$\dim X^\perp = n - \dim \text{Bild } F = n - k.$$

Weiter wegen  $X \cap X^\perp = \{0\}$  nach Satz 1.5

$$\dim(X + X^\perp) = \dim X + \dim X^\perp = k + n - k = n,$$

und somit  $X \oplus X^\perp = \mathbb{R}^n$ . □

**Definition 2.3 (transponierte Matrix)** Für  $A \in \mathbb{R}^{m \times n}$  bezeichnet  $A^T \in \mathbb{R}^{n \times m}$  die Matrix mit den Einträgen  $a_{ji}^T = a_{ij}$  für  $1 \leq j \leq n, 1 \leq i \leq m$ .

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

**Lemma 2.2** Für  $A \in \mathbb{R}^{m \times n}$  gilt bezüglich der Standardskalarprodukte auf  $\mathbb{R}^m$  bzw.  $\mathbb{R}^n$

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A^T \mathbf{w} \rangle \quad \text{für alle } \mathbf{v} \in \mathbb{R}^n, \mathbf{w} \in \mathbb{R}^m.$$

BEWEIS: Wir berechnen

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^m (A\mathbf{v})_i w_i = \sum_{i=1}^m \left( \sum_{j=1}^n a_{ij} v_j \right) w_i = \sum_{j=1}^n v_j \left( \sum_{i=1}^m a_{ji}^T w_i \right) = \sum_{j=1}^n v_j (A^T \mathbf{w})_j = \langle \mathbf{v}, A^T \mathbf{w} \rangle.$$

□

Der Rang von  $A \in \mathbb{R}^{m \times n} = L(\mathbb{R}^n, \mathbb{R}^m)$  ist die Dimension des Raums Bild  $A$ . Wie wir oben gesehen haben, ist das Bild genau der Span der Spaltenvektoren von  $A$ , und darum nennt man  $\text{rang } A$  auch den Spaltenrang. Die Spaltenvektoren von  $A^T \in \mathbb{R}^{n \times m} = L(\mathbb{R}^m, \mathbb{R}^n)$  sind gerade die Zeilenvektoren von  $A$ , daher ist  $\text{rang } A^T$  der Zeilenrang von  $A$ . Die folgende Relation ist nicht direkt offensichtlich.

**Satz 2.4 (Zeilenrang gleich Spaltenrang)** Für  $A \in \mathbb{R}^{m \times n}$  gilt

$$\dim \text{Bild } A^T = \dim \text{Bild } A.$$

BEWEIS: Für die lineare Abbildung  $A^T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  ergibt die Dimensionsformel

$$\dim \text{Bild } A^T = m - \dim \ker A^T.$$

Berechne nun für  $\mathbf{w} \in \mathbb{R}^m$

$$\begin{aligned} \mathbf{w} \in \ker A^T &\Leftrightarrow \langle \mathbf{v}, A^T \mathbf{w} \rangle = 0 \quad \text{für alle } \mathbf{v} \in \mathbb{R}^n \\ &\Leftrightarrow \langle A\mathbf{v}, \mathbf{w} \rangle = 0 \quad \text{für alle } \mathbf{v} \in \mathbb{R}^n \\ &\Leftrightarrow \mathbf{w} \in \text{Bild } A^\perp. \end{aligned}$$

Also gilt  $\dim \ker A^T = \dim \text{Bild } A^\perp = m - \dim \text{Bild } A$ , und die Behauptung folgt.  $\square$

Zum Schluss des Abschnitts betrachten wir die Hintereinanderschaltung linearer Abbildungen. Sind  $B : X \rightarrow Y$  und  $A : Y \rightarrow Z$  lineare Abbildungen, so ist  $A \circ B = AB$  ebenfalls linear:

$$AB(\mathbf{x}_1 + \mathbf{x}_2) = A(B\mathbf{x}_1 + B\mathbf{x}_2) = AB\mathbf{x}_1 + AB\mathbf{x}_2 \quad \text{und} \quad AB(\lambda\mathbf{x}) = A(\lambda B\mathbf{x}) = \lambda AB\mathbf{x}.$$

Sei nun  $B \in \mathbb{R}^{m \times n} = L(\mathbb{R}^n, \mathbb{R}^m)$  und  $A \in \mathbb{R}^{\ell \times m} = L(\mathbb{R}^m, \mathbb{R}^\ell)$ . Wir berechnen die Matrix  $AB \in L(\mathbb{R}^n, \mathbb{R}^\ell) = \mathbb{R}^{\ell \times n}$ :

$$(AB)_{ik} = (AB\mathbf{e}_k)_i = \sum_{j=1}^m a_{ij}(B\mathbf{e}_k)_j = \sum_{j=1}^m a_{ij}b_{jk} \quad \text{für } i = 1, \dots, \ell, k = 1, \dots, n.$$

Der Eintrag  $(AB)_{ik}$  ergibt sich also durch Multiplikation und Summation der Einträge der  $i$ -ten Zeile von  $A$  mit der  $k$ -ten Spalte von  $B$ .

$$\begin{pmatrix} \dots & \dots & \dots \\ a_{i1} & \dots & a_{im} \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \vdots & b_{1k} & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & b_{mk} & \vdots \end{pmatrix} = \begin{pmatrix} \dots & \vdots & \dots \\ \dots & \sum_{j=1}^m a_{ij}b_{jk} & \dots \\ \dots & \vdots & \dots \end{pmatrix}.$$

Wir nennen  $AB$  das Produkt der Matrizen  $A$  und  $B$ . Beachten Sie, dass  $AB$  nur dann definiert ist, wenn die Zahl der Spalten von  $A$  gleich der Zahl der Zeilen von  $B$  ist. Das Produkt hat folgende Eigenschaften:

- (1) Das Produkt ist assoziativ. Dies folgt daraus, dass die Hintereinanderschaltung von Abbildungen assoziativ ist.
- (2) Das Produkt ist im allgemeinen nicht kommutativ:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

- (3) Das Produkt ist im allgemeinen nicht nullteilerfrei.
- (4) Es gilt  $A(B + C) = AB + AC$ ,  $(A + B)C = AB + AC$  (wenn definiert).
- (5) Für  $A \in \mathbb{R}^{m \times n}$  und die Einheitsmatrix  $E_m \in \mathbb{R}^{m \times m}$  bzw.  $E_n \in \mathbb{R}^{n \times n}$  gilt

$$AE_n = E_m A = A.$$



Also gilt  $A\mathbf{x} = \mathbf{b}$  genau wenn  $\mathbf{x} - \mathbf{x}_0 \in \ker A$ . Aussage (4) folgt aus der Dimensionsformel:

$$\dim \mathbb{R}^n - \dim \text{Bild}A = \dim \ker A.$$

□

Die Hauptfrage ist natürlich, wie werden die Lösungen eines Gleichungssystems in der Praxis berechnet? Bei wichtigen Anwendungen, etwa der Berechnung einer dreidimensionalen Strömung oder einer Wärmeverteilung, treten sehr große Gleichungssysteme auf (zum Beispiel  $10^6$  Unbekannte). Eine allgemeine, systematische Methode ist der Algorithmus von Gauß. Er beruht auf folgenden Tatsachen:

*Die Lösungsmenge eines linearen Gleichungssystems  $A\mathbf{x} = \mathbf{b}$  ändert sich bei folgenden Umformungen der erweiterten Matrix  $(A|\mathbf{b})$  nicht:*

- (1) Vertauschen der Zeilen  $i_1$  und  $i_2$
- (2) Multiplikation der Zeile  $i$  mit  $\lambda \neq 0$
- (3) Addition des  $\lambda$ -fachen der Zeile  $i_1$  zur Zeile  $i_2$ , wobei  $\lambda \in \mathbb{R}$ .

Der Grund ist, dass die Zeilenumformungen wie folgt rückgängig gemacht werden können:

- (1) Vertauschen der Zeilen  $i_1$  und  $i_2$
- (2) Multiplikation der Zeile  $i$  mit  $1/\lambda$
- (3) Addition des  $(-\lambda)$ -fachen der Zeile  $i_1$  zur Zeile  $i_2$

**Beispiel 3.1** Wir zeigen exemplarisch, wie ein Gleichungssystem durch elementare Zeilenumformungen der erweiterten Koeffizientenmatrix gelöst wird. Die Matrix bezeichnen wir immer mit  $A = (a_{ij})$ , sie ändert sich in jedem Schritt. Betrachte

$$A = \left( \begin{array}{ccccc|c} 0 & 0 & 0 & -4 & -3 & 1 \\ 0 & -4 & 2 & 2 & 3 & 3 \\ 0 & 6 & -3 & 1 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & -2 \end{array} \right).$$

*1. Schritt:* Suche die erste Spalte, in der nicht alle Einträge Null sind, das ist hier die Spalte  $j_1 = 2$ . Wenn nötig, vertausche dann Zeilen so dass  $a_{1j_1} \neq 0$ . Dieses Matrix-Element heißt erstes Pivot-Element (*französisch:* Dreh- und Angelpunkt). Im Beispiel vertauschen wir die Zeilen  $i = 1$  und  $i = 4$ . Stand:

$$A = \left( \begin{array}{ccccc|c} 0 & 2 & -1 & 1 & 0 & -2 \\ 0 & -4 & 2 & 2 & 3 & 3 \\ 0 & 6 & -3 & 1 & 1 & 0 \\ 0 & 0 & 0 & -4 & -3 & 1 \end{array} \right)$$

*2. Schritt:* Addiere zur  $i$ -ten Zeile,  $i = 2, \dots, m$ , das  $-a_{ij_1}/a_{1j_1}$ -fache der ersten Zeile. Hierdurch werden die Elemente  $a_{ij_1}$  mit  $i = 2, \dots, m$  zu Null. Stand:

$$A = \left( \begin{array}{ccccc|c} 0 & 2 & -1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 4 & 3 & -1 \\ 0 & 0 & 0 & -2 & 1 & 6 \\ 0 & 0 & 0 & -4 & -3 & 1 \end{array} \right)$$

3. *Schritt.* Suche die nächste Spalte  $j_2 > j_1$ , in der nicht alle Einträge  $a_{ij_2}$ ,  $i = 2, \dots, m$ , Null sind. Durch evtl. Vertauschen von Zeilen erhalte  $a_{2j_2} \neq 0$ , dies ist das nächste Pivotelement. Addiere nun zur  $i$ -ten Zeile,  $i = 3, \dots, m$ , das  $-a_{ij_2}/a_{2j_2}$ -fache der zweiten Zeile. Dies macht alle Einträge  $a_{ij_2}$ ,  $i = 3, \dots, m$ , zu Null. Im Beispiel ist  $j_2 = 4$  und  $a_{24} \neq 0$ , so dass kein Tausch notwendig ist. Stand:

$$A = \left( \begin{array}{ccccc|c} 0 & 2 & -1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 4 & 3 & -1 \\ 0 & 0 & 0 & 0 & \frac{5}{2} & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

*Letzter Schritt.* Führe das Verfahren bis zum Pivotelement  $a_{rj_r}$  durch, so dass alle Einträge  $a_{ij}$  mit  $i > r$  Null sind. Im Beispiel ist  $r = 3$  und  $j_r = 5$ .

Durch diesen Prozess ergibt sich eine Matrix  $A$  in Zeilenstufenform: es gibt ein  $r \leq \min(m, n)$  und  $1 \leq j_1 < \dots < j_r \leq n$ , so dass gilt:

- Die Elemente  $a_{1j_1}, \dots, a_{rj_r}$  sind ungleich Null (Pivotelemente)
- Die Elemente  $a_{ij}$  sind Null für  $i = 1, \dots, r$ ,  $1 \leq j < j_r$ , und für  $i > r$ .

Ist  $A$  in Zeilenstufenform, so können wir die Lösungen der Gleichung  $A\mathbf{x} = \mathbf{b}$  induktiv ausrechnen. Ist  $b_i \neq 0$  für ein  $i > r$ , so gibt es keine Lösung. Andernfalls erhalten wir der Reihe nach durch Auflösen der Gleichungen  $i = r, \dots, 1$  in dieser Reihenfolge

$$x_{j_i} = \frac{1}{a_{ij_i}} \left( b_i - \sum_{j=j_i+1}^n a_{ij} x_j \right) \quad \text{für } i = r, \dots, 1.$$

Die  $n - r$  Koordinaten  $x_j$  mit  $j \neq j_1, \dots, j_r$  sind dabei frei wählbar, im Fall  $\mathbf{b} = \mathbf{0}$  erhalten wir so alle Lösungen der homogenen Gleichung. Es ist leicht zu sehen, dass die ersten  $r$  Zeilen der Zeilenstufenform  $A^{ZS}$  linear unabhängig sind, also hat diese den Zeilenrang und damit auch Spaltenrang  $r$ . Aber die Zeilenumformungen ändern nicht die Lösungen der homogenen Gleichung, das heißt den Kern. Nach der Dimensionsformel ist somit

$$\dim \text{Bild } A = n - \dim \ker A = n - \dim \ker A^{ZS} = \dim \text{Bild } A^{ZS} = r,$$

das heißt  $r$  ist auch der Rang der ursprünglich gegebenen Matrix  $A$ , während sich das Bild selbst unter den Zeilenumformungen natürlich ändert.

## 4 Die inverse Matrix

In diesem Abschnitt geht es um  $n \times n$ -Matrizen  $A$ , also mit gleichvielen Zeilen wie Spalten. Die zugehörige lineare Abbildung lautet dann  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ .

**Definition 4.1** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt invertierbar, wenn es  $B \in \mathbb{R}^{n \times n}$  gibt mit

$$AB = BA = E_n. \tag{4.8}$$

Die Matrix  $B$  ist durch (4.8) eindeutig bestimmt, wird mit  $A^{-1}$  bezeichnet und heißt die zu  $A$  inverse Matrix oder Inverse von  $A$ .

Die Eindeutigkeit der inversen Matrix ergibt sich wie folgt.

**Lemma 4.1** Zu  $A \in \mathbb{R}^{n \times n}$  gebe es  $B, C \in \mathbb{R}^{n \times n}$  mit  $BA = E_n = AC$ . Dann gilt  $B = C$ , und  $A$  ist invertierbar.

BEWEIS:  $B = BE_n = B(AC) = (BA)C = E_n C = C$ . □

**Beispiel 4.1** Die Einheitsmatrix  $E_n$  ist invertierbar, es ist  $E_n^{-1} = E_n$ .

**Beispiel 4.2** Betrachte eine Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2 \times 2} \quad \text{mit } ad - bc \neq 0.$$

Dann ist  $A$  invertierbar, und zwar gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In der Tat sehen wir  $AA^{-1} = A^{-1}A = E_2$  wie verlangt.

**Lemma 4.2** Seien  $A, B \in \mathbb{R}^{n \times n}$  invertierbar. Dann gilt:

- (1)  $A^{-1}$  ist invertierbar mit  $(A^{-1})^{-1} = A$ .
- (2)  $AB$  ist invertierbar mit  $(AB)^{-1} = B^{-1}A^{-1}$ .
- (3)  $A^T$  ist invertierbar mit  $(A^T)^{-1} = (A^{-1})^T$ .

BEWEIS: Aussage (1) gilt wegen  $AA^{-1} = A^{-1}A = E_n$ . Für Aussage (2) berechnen wir

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AE_nA^{-1} = E_n \\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}E_nB = E_n. \end{aligned}$$

Für Aussage (3) berechnen wir zuerst (vgl. Aufgabe 1, Serie 14)

$$(AB)_{ik}^T = (AB)_{ki} = \sum_{j=1}^n A_{kj}B_{ji} = \sum_{j=1}^n (B^T)_{ij}(A^T)_{jk} = (B^T A^T)_{ik}.$$

Somit ist  $(AB)^T = B^T A^T$ . Daraus folgt weiter

$$\begin{aligned} A^T(A^{-1})^T &= (A^{-1}A)^T = (E_n)^T = E_n \\ (A^{-1})^T A^T &= (AA^{-1})^T = (E_n)^T = E_n. \end{aligned}$$

Damit ist das Lemma bewiesen. □

Wir bringen nun die linearen Abbildungen ins Spiel.

**Satz 4.1**  $A \in \mathbb{R}^{n \times n}$  ist genau dann invertierbar, wenn die Abbildung  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  bijektiv ist. In diesem Fall ist die Umkehrabbildung durch die Matrix  $A^{-1}$  gegeben.

BEWEIS: Wie auf Seite 89 gezeigt, gilt

$$A(B\mathbf{x}) = (AB)\mathbf{x} \quad \text{für } A, B \in \mathbb{R}^{n \times n}.$$

Dabei steht rechts das Matrixprodukt. Für  $A \in \mathbb{R}^{n \times n}$  invertierbar folgt

$$A(A^{-1}\mathbf{x}) = E_n\mathbf{x} = \mathbf{x} \quad \text{und} \quad A^{-1}(A\mathbf{x}) = E_n\mathbf{x} = \mathbf{x} \quad \text{für alle } \mathbf{x} \in \mathbb{R}^n.$$

Also ist  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  bijektiv, und die Umkehrabbildung ist durch die Matrix  $A^{-1}$  gegeben. Sei umgekehrt  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  bijektiv, und  $B$  die Umkehrabbildung von  $A$ . Dann ist  $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear, denn es gilt für  $\mathbf{v} = A\mathbf{x}$  und  $\mathbf{w} = A\mathbf{y}$ ,

$$\begin{aligned} B(\mathbf{v} + \mathbf{w}) &= B(A\mathbf{x} + A\mathbf{y}) = B(A(\mathbf{x} + \mathbf{y})) = \mathbf{x} + \mathbf{y} = B\mathbf{v} + B\mathbf{w}, \\ B(\lambda\mathbf{v}) &= B(\lambda A\mathbf{x}) = B(A(\lambda\mathbf{x})) = \lambda\mathbf{x} = \lambda B\mathbf{v}. \end{aligned}$$

Also können wir  $B$  als  $n \times n$ -Matrix auffassen. Es folgt

$$(AB)\mathbf{e}_j = A(B\mathbf{e}_j) = \mathbf{e}_j \quad \text{für } j = 1, \dots, n,$$

das heißt  $AB = E_n$ . Analog sehen wir  $BA = E_n$ , also ist  $A$  invertierbar. □

**Satz 4.2** Für eine Matrix  $A \in \mathbb{R}^{n \times n}$  sind äquivalent:

- (a)  $A$  ist invertierbar
- (b)  $\text{rang } A = n$ .
- (c)  $\ker A = \{0\}$ .
- (d) Es gibt ein  $C \in \mathbb{R}^{n \times n}$  mit  $AC = E_n$ .
- (e) Es gibt ein  $B \in \mathbb{R}^{n \times n}$  mit  $BA = E_n$ .

Falls (d) oder (e) gilt, so ist  $C = A^{-1}$  bzw.  $B = A^{-1}$ .

BEWEIS: Die Dimensionsformel, siehe Satz 2.1, liefert

$$\text{rang } A = n - \dim \ker A.$$

Also sind (b) und (c) äquivalent. Aber (b) und (c) zusammen sind äquivalent zu (a), vgl. Lemma 2.1. Damit sind (a), (b) und (c) untereinander äquivalent. Nun folgen (d) und (e) trivial aus (a). Umgekehrt folgt (b) aus (d):

$$\text{Bild } A \supset \text{Bild } AC = \text{Bild } E_n = \mathbb{R}^n.$$

Schließlich folgt (c) aus (e):

$$A\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad B A \mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}.$$

Damit sind alle fünf Aussagen äquivalent. Außerdem  $B = C = A^{-1}$  nach Lemma 4.1. □

Natürlich ergibt sich die Frage, wie die inverse Matrix bestimmt werden kann. Um eine Formel herzuleiten, brauchen wir Determinanten, das müssen wir in den Sommer verschieben. Wir können aber die Matrix  $A^{-1}$  mit dem Gauß-Algorithmus berechnen. Dies beruht auf folgender Tatsache:

**Lemma 4.3** Die elementaren Zeilenumformungen einer Matrix  $A \in \mathbb{R}^{n \times n}$  entsprechen einer Multiplikation von links mit einer invertierbaren Matrix.

BEWEIS: Die Vertauschung der Zeilen  $i_1$  und  $i_2$  ergibt sich durch Linksmultiplikation mit

$$P^{i_1, i_2} = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 0 & 1 & & \vdots \\ \vdots & & 1 & 0 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Es ist leicht zu sehen, dass  $(P^{i_1, i_2})^{-1} = P^{i_1, i_2}$ . Die Multiplikation der  $i$ -ten Zeile mit  $\lambda \neq 0$  ergibt sich durch Linksmultiplikation mit der Matrix

$$Q^i(\lambda) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \lambda & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Es ist  $Q(\lambda)^{-1} = Q(\frac{1}{\lambda})$ . Schließlich ist die Addition des  $\lambda$ -fachen der Zeile  $i_1$  zur Zeile  $i_2$  gegeben durch Linksmultiplikation mit der Matrix

$$R^{i_1, i_2}(\lambda) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & 0 & & \vdots \\ \vdots & & \lambda & 1 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Man prüft nach, dass auch diese Matrix invertierbar ist mit  $R^{i_1, i_2}(\lambda) = R^{i_1, i_2}(-\lambda)$ . □

Wir berechnen die inverse Matrix nun für

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

Als erstes schreiben wir  $A$  und rechts daneben die Einheitsmatrix hin:

$$\left( \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array} \right).$$

Jetzt bringen wir  $A$  auf Zeilenstufenform, und wenden gleichzeitig rechts dieselben Operationen an (ähnlich wie bei der rechten Seite). Wir beginnen mit der ersten Spalte, also  $j_1 = 1$ .

$$\left( \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -3 & -1 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array} \right).$$

Jetzt kommt die zweite Spalte, also  $j_2 = 2$ .

$$\left( \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -3 & -1 & 1 & 0 \\ 0 & 0 & 5 & 1 & -2 & 1 \end{array} \right).$$

Die Zeilenstufenform ist eine obere Dreiecksmatrix mit Diagonalelementen ungleich Null. Für eine invertierbare Matrix muss das so sein:  $A \in \mathbb{R}^{n \times n}$  hat den Rang  $n$ , das heißt die Zahl der Stufen ist gleich  $n$ . Wegen  $1 \leq j_1 < \dots < j_n = n$  muss  $(j_1, \dots, j_n) = (1, \dots, n)$  gelten, die Pivotelemente stehen also auf der Diagonalen. Als nächstes teilen wir die Zeilen durch die Diagonalelemente:

$$\left( \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{array} \right).$$

Jetzt machen wir die Nichtdiagonalelemente in der  $j$ -ten Spalte zu Null durch Addition von Vielfachen der  $j$ -ten Zeile. Wir beginnen mit  $j = n = 3$ :

$$\left( \begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{2}{5} & \frac{6}{5} & -\frac{3}{5} \\ 0 & 1 & 0 & \frac{2}{5} & -\frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{array} \right).$$

Dann kommt die vorletzte Spalte dran, also hier  $j = 2$ .

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{5} & \frac{4}{5} & \frac{3}{5} \\ 0 & 1 & 0 & \frac{2}{5} & -\frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{array} \right).$$

Die Matrix  $B$  rechts ist nun die Inverse von  $A$ , zum Beispiel rechnen wir nach

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & 4 & 3 \\ 2 & 1 & -3 \\ 1 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Warum funktioniert das? Jede Umformung ist eine Linksmultiplikation mit einer Elementarmatrix, also haben wir links am Ende die Gleichung  $QA = E_3$ , wobei  $Q$  ein Produkt von Elementarmatrizen ist. Da wir auf  $E_3$  rechts dieselben Operationen angewandt haben, folgt

$$B = QE_3 = Q, \quad \text{also } BA = QA = E_3.$$

Somit ist  $B$  Linksinverse zu  $A$ , und  $B = A^{-1}$  nach Satz 4.2.

## 5 Determinanten

Wir beginnen die Vorlesung mit folgendem Satz.

**Satz 5.1 (Determinante)** *Es gibt genau eine Funktion*

$$\det : \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{n \text{ Faktoren}} \rightarrow \mathbb{R}, (a_1, \dots, a_n) \mapsto \det(a_1, \dots, a_n),$$

mit folgenden Eigenschaften:

(1) *det ist multilinear, d.h. linear in jedem Eintrag:*

$$\det(\dots, \lambda a_i + \mu b_i, \dots) = \lambda \det(\dots, a_i, \dots) + \mu \det(\dots, b_i, \dots).$$

(2) *det ist alternierend: bei Vertauschung von zwei Einträgen ändert sich das Vorzeichen,*

$$\det(\dots, a_i, \dots, a_j, \dots) = -\det(\dots, a_j, \dots, a_i, \dots).$$

(3) *det ist normiert: für die Standardbasis gilt*

$$\det(e_1, \dots, e_n) = 1.$$

Wir betrachten wir erst die Spezialfälle  $n = 2$  und  $n = 3$ . Zuvor noch eine Tatsache.

**Lemma 5.1** *Sind die Vektoren  $a_1, \dots, a_n \in \mathbb{R}^n$  linear abhängig, so gilt*

$$\det(a_1, \dots, a_n) = 0.$$

*Das ist insbesondere dann der Fall, wenn zwei der  $a_k$  gleich sind.*

BEWEIS: Steht derselbe Vektor  $a$  an den Stellen  $i$  und  $j$ , so folgt durch Vertauschung

$$\det(\dots, a, \dots, a, \dots) = -\det(\dots, a, \dots, a, \dots) = 0.$$

Sind  $a_1, \dots, a_n$  linear abhängig, so gibt es ein  $k$  mit  $a_k = \sum_{i=1, i \neq k}^n \lambda_i a_i$  für geeignete  $\lambda_i \in \mathbb{R}$ , und es folgt

$$\begin{aligned} \det(a_1, \dots, a_k, \dots, a_n) &= \det(a_1, \dots, \sum_{i=1, i \neq k}^n \lambda_i a_i, \dots, a_n) \\ &= \sum_{i=1, i \neq k}^n \lambda_i \det(a_1, \dots, a_i, \dots, a_n) = 0. \end{aligned}$$

□

**Beispiel 5.1** Wir berechnen die Determinante von  $(a, b) \in \mathbb{R}^2 \times \mathbb{R}^2$ .

$$\begin{aligned} \det(a, b) &= \det(a_1 e_1 + a_2 e_2, b_1 e_1 + b_2 e_2) \\ &= a_1 b_1 \det(e_1, e_1) + a_1 b_2 \det(e_1, e_2) + a_2 b_1 \det(e_2, e_1) + a_2 b_2 \det(e_2, e_2) \\ &= a_1 b_2 - a_2 b_1. \end{aligned}$$

Die Zahl  $|\det(a, b)|$  ist der Flächeninhalt  $A(a, b)$  des von  $a$  und  $b$  aufgespannten Parallelogramms. Um das zu sehen, seien zunächst  $a, b$  zueinander senkrecht, also

$$\langle a, b \rangle = a_1 b_1 + a_2 b_2 = 0.$$

Dann folgt  $a_1^2 b_1^2 + 2a_1 a_2 b_1 b_2 + a_2^2 b_2^2 = 0$  und weiter

$$\begin{aligned} A(a, b)^2 &= |a|^2 |b|^2 \\ &= (a_1^2 + a_2^2)(b_1^2 + b_2^2) \\ &= \underbrace{a_1^2 b_1^2 + 2a_1 a_2 b_1 b_2 + a_2^2 b_2^2}_{=0} + a_1^2 b_2^2 - 2a_1 a_2 b_1 b_2 + a_2^2 b_1^2 \\ &= (a_1 b_2 - a_2 b_1)^2 = \det(a, b)^2. \end{aligned}$$

Allgemein gibt es ein  $\lambda \in \mathbb{R}$ , so dass  $a$  und  $b - \lambda a$  senkrecht sind. Das von  $a, b - \lambda a$  aufgespannte Rechteck hat auch Flächeninhalt  $A(a, b)$  (Scherung eines Parallelogramms). Also

$$A(a, b) = A(a, b - \lambda a) = |\det(a, b - \lambda a)| = |\det(a, b) - \lambda \det(a, a)| = |\det(a, b)|.$$

**Beispiel 5.2** Als nächstes kommt der Fall  $n = 3$ . Beim Ausmultiplizieren können wir im ersten Eintrag den  $i$ -ten, im zweiten den  $j$ -ten und im dritten den  $k$ -ten Basisvektor nehmen. Es ergibt sich

$$\begin{aligned} \det(a, b, c) &= \det\left(\sum_{i=1}^3 a_i e_i, \sum_{j=1}^3 b_j e_j, \sum_{k=1}^3 c_k e_k\right) \\ &= \sum_{i,j,k=1}^3 a_i b_j c_k \det(e_i, e_j, e_k). \end{aligned}$$

Allerdings sind alle Terme Null, für die zwei der drei Zahlen  $i, j, k$  (oder alle drei) übereinstimmen. Wir brauchen mit  $(ijk)$  also nur die Umordnungen von  $(123)$  zu durchlaufen. Das Vorzeichen  $\varepsilon_{ijk} \in \{+1, -1\}$  ist wie folgt gegeben:

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{für } (ijk) = (123), (312), (231) \\ -1 & \text{für } (ijk) = (213), (321), (132). \end{cases}$$

Im Fall  $\varepsilon_{ijk} = -1$  lässt sich  $ijk$  durch eine Vertauschung in  $(123)$  überführen, es folgt  $\det(e_i, e_j, e_k) = -1$ . Im Fall  $\varepsilon_{ijk} = +1$  braucht man zwei (oder keine) Vertauschung, es folgt  $\det(e_i, e_j, e_k) = +1$ . Damit gilt, wenn wir die Menge der Umordnungen von  $\{1, 2, 3\}$  mit  $\mathbb{S}_3$  bezeichnen:

$$\begin{aligned} \det(a, b, c) &= \sum_{(ijk) \in \mathbb{S}_3} \varepsilon_{ijk} a_i b_j c_k \\ &= a_1 b_2 c_3 + a_3 b_1 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 - a_3 b_2 c_1 - a_1 b_3 c_2. \end{aligned}$$

Sei  $a \times b$  das Kreuzprodukt der Vektoren  $a, b$ . Dann gilt, siehe Abschnitt 4 in Kapitel 1,

$$\det(a, b, c) = \langle a \times b, c \rangle.$$

Wie in Kapitel 1, Abschnitt 4 gezeigt, steht der Vektor  $a \times b$  senkrecht auf die von  $a, b$  aufgespannte Ebene, seine Länge ist

$$|a \times b| = |a| |b| \sin \angle(a, b).$$

Die rechte Seite ist wiederum der Flächeninhalt des von  $a, b$  aufgespannten Parallelogramms. Wir behaupten, dass  $|\det(a, b, c)|$  das Volumen  $V(a, b, c)$  des von  $a, b, c$  aufgespannten Parallelotops ist. Dieses ist die Menge

$$P = \{\lambda a + \mu b + \gamma c : 0 \leq \lambda, \mu, \gamma \leq 1\}.$$

Ist  $c$  senkrecht auf  $a, b$ , so ist  $c$  parallel zu  $a \times b$  und es folgt

$$|\det(a, b, c)| = |\langle a \times b, c \rangle| = |a \times b| |c| = A(a, b) |c| = V(a, b, c).$$

Allgemein gibt es einen Vektor  $v = \lambda a + \mu b$ , so dass  $c - v$  auf  $a, b$  senkrecht steht. Das von  $a, b, c - v$  aufgespannte Parallelotop hat dann ebenfalls das Volumen  $V(a, b, c)$  (Scherungsinvarianz des Volumens). Somit folgt

$$V(a, b, c) = V(a, b, c - v) = |\det(a, b, c - v)| = |\det(a, b, c) - \det(a, b, v)| = |\det(a, b, c)|.$$

Damit haben wir in zwei und drei Dimensionen eine geometrische Interpretation der Determinante, jedenfalls bis aufs Vorzeichen.

*Beweis von Satz 5.1 (Skizze)* Wir leiten aus den Eigenschaften eine Formel her, was die Eindeutigkeit zeigt. Sei  $f : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$  eine Funktion mit den Eigenschaften (1) und (2). Indem wir die Multilinearität benutzen und ausmultiplizieren, erhalten wir

$$f(a_1, \dots, a_n) = \sum_{\sigma_1, \dots, \sigma_n=1}^n (a_1)_{\sigma_1} \cdot \dots \cdot (a_n)_{\sigma_n} \det(e_{\sigma_1}, \dots, e_{\sigma_n}).$$

Sind (mindestens) zwei der  $\sigma_1, \dots, \sigma_n$  gleich, so ist der Term Null. Somit ist nur über die Umordnungen  $\sigma \in \mathbb{S}_n$  von  $1, \dots, n$  zu summieren. Wir brauchen folgende Fakten: jede Umordnung lässt sich durch endlich viele Vertauschungen in die Standardreihenfolge  $(1 \dots n)$  bringen. Die Zahl dieser Vertauschungen ist nicht eindeutig bestimmt, aber sie ist entweder stets gerade oder stets ungerade. Man spricht dann von einer geraden bzw. einer ungeraden Umordnung, und definiert

$$\varepsilon(\sigma) = \begin{cases} +1 & \text{falls } \sigma \text{ gerade,} \\ -1 & \text{falls } \sigma \text{ ungerade.} \end{cases}$$

Da jede Vertauschung ein Minuszeichen liefert, folgt nun

$$f(e_{\sigma_1}, \dots, e_{\sigma_n}) = \varepsilon(\sigma) f(e_1, \dots, e_n),$$

und damit für die gesamte Summe

$$f(a_1, \dots, a_n) = f(e_1, \dots, e_n) \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) (a_1)_{\sigma_1} \cdot \dots \cdot (a_n)_{\sigma_n}. \quad (5.9)$$

Für die Determinante soll ja noch die Normierung (3) gelten, wir erhalten

$$\det(a_1, \dots, a_n) = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) (a_1)_{\sigma_1} \cdot \dots \cdot (a_n)_{\sigma_n}. \quad (5.10)$$

Dies zeigt die Eindeutigkeit. Für die Existenz definiere die Determinante durch (5.10). Es sind dann (1), (2) und (3) zu zeigen. Die Multilinearität ist klar, und aus  $(e_j)_\sigma = \delta_{\sigma j}$  folgt

$$\det(e_1, \dots, e_n) = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) \delta_{\sigma_1 1} \cdot \dots \cdot \delta_{\sigma_n n} = \varepsilon(\text{id}) = 1.$$

Ist  $\tau$  eine Vertauschung, so durchläuft  $\sigma\tau$  für  $\sigma \in \mathbb{S}_n$  alle Umordnungen, und es gilt  $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$ . Somit gilt auch die Eigenschaft (2):

$$\begin{aligned} \det(a_{\tau_1}, \dots, a_{\tau_n}) &= \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma\tau) (a_{\tau_1})_{\sigma_{\tau_1}} \cdot \dots \cdot (a_{\tau_n})_{\sigma_{\tau_n}} \\ &= \varepsilon(\tau) \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) (a_1)_{\sigma_1} \cdot \dots \cdot (a_n)_{\sigma_n} \\ &= \varepsilon(\tau) \det(a_1, \dots, a_n). \end{aligned}$$

□

*Zur Definition des Vorzeichens:* jede Umordnung ergibt sich aus  $\text{id} = (1 \dots n)$  durch höchstens  $n - 1$  Vertauschungen. Für  $n = 1$  gibt es nur  $\text{id}$ , und wir brauchen null Vertauschungen. Für  $\sigma \in \mathbb{S}_n$  mit  $n \geq 2$  unterscheiden wir zwei Fälle:

**Fall 1**  $\sigma_n = n$ , also  $\sigma = (\sigma', n)$  mit  $\sigma' \in \mathbb{S}_{n-1}$ .

Per Induktion brauchen wir höchstens  $n - 2$  Vertauschungen für  $\sigma'$ , also auch für  $\sigma$ .

**Fall 2**  $\sigma_k = n$  für ein  $k < n$ , also  $\sigma = \tau_{kn}(\sigma', n)$  mit  $\sigma' \in \mathbb{S}_{n-1}$ .

Zu den höchstens  $n - 2$  Vertauschungen für  $\sigma'$  kommt noch  $\tau_{kn}$  hinzu, also maximal  $n - 1$ .

Wir definieren  $\varepsilon : \mathbb{S}_n \rightarrow \{\pm 1\}$  induktiv: für  $n = 1$  sei  $\varepsilon(\text{id}) = 1$ , und für  $n \geq 2$  sei

$$\varepsilon(\sigma) = \begin{cases} \varepsilon(\sigma') & \text{falls } \sigma = (\sigma', n) \text{ mit } \sigma' \in \mathbb{S}_{n-1} \\ -\varepsilon(\sigma') & \text{falls } \sigma = \tau_{kn}(\sigma', n) \text{ mit } \sigma' \in \mathbb{S}_{n-1}, k < n. \end{cases}$$

Wir behaupten, dass  $\varepsilon(\sigma)$  unter Vertauschungen sein Vorzeichen wechselt. Sei  $\tau_{ij}$  die Vertauschung des  $i$ -ten und  $j$ -ten Eintrags, mit  $1 \leq i, j \leq n$  und  $i \neq j$ . Dann gelten folgende Tabellen:

**Fall 1**  $\sigma = (\sigma', n)$  mit  $\sigma' \in \mathbb{S}_{n-1}$ , also  $\varepsilon(\sigma) = \varepsilon(\sigma')$

$\tau_{ij}$	$\tau_{ij}\sigma$	$\varepsilon(\tau_{ij}\sigma)$
$i, j < n$	$(\tau_{ij}\sigma', n)$	$\varepsilon(\tau_{ij}\sigma') = -\varepsilon(\sigma')$
$i < j = n$	$\tau_{in}(\sigma', n)$	$-\varepsilon(\sigma')$

**Fall 2**  $\sigma = \tau_{kn}(\sigma', n)$  mit  $\sigma' \in \mathbb{S}_{n-1}$ ,  $k < n$ , also  $\varepsilon(\sigma) = -\varepsilon(\sigma')$

$\tau_{ij}$	$\tau_{ij}\sigma = \tau_{ij}\tau_{kn}(\sigma', n)$	$\varepsilon(\tau_{ij}\sigma)$
$i, j < n, k \neq i, j$	$\tau_{kn}(\tau_{ij}\sigma', n)$	$-\varepsilon(\tau_{ij}\sigma') = \varepsilon(\sigma')$
$i, j < n, k = i$	$\tau_{jn}(\tau_{ij}\sigma', n)$	$-\varepsilon(\tau_{ij}\sigma') = \varepsilon(\sigma')$
$i < j = n, k \neq i$	$\tau_{kn}(\tau_{ik}\sigma', n)$	$-\varepsilon(\tau_{ik}\sigma') = \varepsilon(\sigma')$
$i < j = n, k = i$	$(\sigma', n)$	$\varepsilon(\sigma')$

Lässt sich  $\sigma$  durch  $k$  Vertauschungen aus id erzeugen, so ist  $(-1)^k = \varepsilon(\sigma)$ . Durch  $\varepsilon(\sigma)$  ist also bestimmt, ob  $k$  gerade oder ungerade ist. Lassen sich  $\sigma_{1,2}$  durch  $k_{1,2}$  Umordnungen erzeugen, so folgt

$$\varepsilon(\sigma_1\sigma_2) = (-1)^{k_1+k_2} = (-1)^{k_1}(-1)^{k_2} = \varepsilon(\sigma_1)\varepsilon(\sigma_2).$$

Aus dem Beweis von Satz 5.1 halten wir folgendes Ergebnis fest.

**Folgerung 5.1** *Ist  $f : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$  multilinear und alternierend, so gilt*

$$f(a_1, \dots, a_n) = f(e_1, \dots, e_n) \det(a_1, \dots, a_n).$$

Ein Tupel von  $n$  Vektoren  $a_1, \dots, a_n \in \mathbb{R}^n$  kann als Matrix  $A \in \mathbb{R}^{n \times n}$  aufgefasst werden. An dieser Stelle wollen wir die  $a_j$  als Spaltenvektoren nehmen.

**Definition 5.1 (Determinante einer Matrix)** *Für eine Matrix  $A \in \mathbb{R}^{n \times n}$  mit Spaltenvektoren  $a_1, \dots, a_n$  setzen wir*

$$\det(A) = \det(a_1, \dots, a_n) = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) a_{\sigma_1 1} \cdot \dots \cdot a_{\sigma_n n}.$$

Seien  $A, B \in \mathbb{R}^{n \times n}$ . Ist  $B = (b_1, \dots, b_n)$  mit Spaltenvektoren  $b_j$ , so gilt für das Produkt

$$AB = (Ab_1, \dots, Ab_n). \quad (5.11)$$

Das können wir in Koordinaten nachrechnen, es gilt

$$(AB)_{ik} = \sum_{j=1}^n a_{ij} b_{jk} = \sum_{j=1}^n a_{ij} (b_k)_j = (Ab_k)_i.$$

Damit kommen wir zu einer zentralen Eigenschaft der Determinante.

**Satz 5.2 (Determinantenmultiplikationssatz)** *Es gilt*

$$\det(AB) = \det(A) \det(B) \quad \text{für alle } A, B \in \mathbb{R}^{n \times n}.$$

BEWEIS: Betrachte  $f(b_1, \dots, b_n) := \det(Ab_1, \dots, Ab_n) = \det(AB)$ . Die Funktion  $f$  ist multilinear und alternierend, also gilt nach Folgerung 5.1

$$\begin{aligned} \det(AB) &= f(b_1, \dots, b_n) \\ &= f(e_1, \dots, e_n) \det(b_1, \dots, b_n) \\ &= \det(Ae_1, \dots, Ae_n) \det(b_1, \dots, b_n) \\ &= \det(A) \det(B). \end{aligned}$$

□

**Satz 5.3 (Invertierbarkeitstest)** *Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau invertierbar, wenn  $\det(A) \neq 0$ . In diesem Fall gilt*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

BEWEIS: Ist  $A$  nicht invertierbar, so ist  $A$  nicht surjektiv nach der Dimensionsformel. Dann sind die Spaltenvektoren linear abhängig, und es folgt  $\det(A) = 0$  aus Lemma 5.1. Ist  $A$  invertierbar, so folgt

$$1 = \det(E_n) = \det(A^{-1}A) = \det(A^{-1}) \det(A).$$

Also ist  $\det(A) \neq 0$  und die Formel gilt. □

Damit haben wir einen einfachen Test, ob eine gegebene quadratische Matrix invertierbar ist. Rechentechisch betrachtet ist allerdings der Gauß-Algorithmus effektiver, die Summe in der Determinante enthält nämlich  $n!$  Summanden mit jeweils  $n$  Faktoren.

**Satz 5.4** Für  $A \in \mathbb{R}^{n \times n}$  gilt

$$\det(A^T) = \det(A).$$

BEWEIS: Um diese Eigenschaft zu zeigen, müssen wir auf die Formel für die Determinante zurückgreifen. Wegen  $a_{ij}^T = a_{ji}$  gilt

$$\det(A^T) = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) a_{\sigma_1 1}^T \cdots a_{\sigma_n n}^T = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) a_{1\sigma_1} \cdots a_{n\sigma_n}.$$

Für  $j = 1, \dots, n$  sei  $\tau_j$  die Nummer mit  $\sigma_{\tau_j} = j$ ; man nennt  $\tau$  die inverse Umordnung zu  $\sigma$ . Nun ist  $a_{\tau_j \sigma_{\tau_j}} = a_{\tau_j j}$ , und die  $\tau_j$  durchlaufen alle  $i = 1, \dots, n$ . Es folgt

$$a_{1\sigma_1} \cdots a_{n\sigma_n} = a_{\tau_1 1} \cdots a_{\tau_n n}.$$

Ergibt sich  $\sigma$  durch  $k$  und  $\tau$  durch  $\ell$  Vertauschungen aus  $\text{id}$ , so kommen wir mit  $k + \ell$  Vertauschungen zurück zu  $\text{id}$ :

$$(1 \dots n) \xrightarrow{k \text{ Vertauschungen}} (\sigma_1 \dots \sigma_n) \xrightarrow{\ell \text{ Vertauschungen}} (\sigma_{\tau_1} \dots \sigma_{\tau_n}) = (1 \dots n).$$

Da  $\text{id}$  gerade, sind  $k, \ell$  entweder beide gerade oder beide ungerade. Dies zeigt  $\varepsilon(\tau) = \varepsilon(\sigma)$ . Mit  $\sigma$  durchläuft  $\tau$  die Menge aller Umordnungen genau einmal, also

$$\det(A^T) = \sum_{\tau \in \mathbb{S}_n} \varepsilon(\tau) a_{\tau_1 1} \cdots a_{\tau_n n} = \det(A).$$

□

Die Zeilenvektoren von  $A$  sind die Spaltenvektoren von  $A^T$ . Es folgt

**Folgerung 5.2**  $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  ist multilinear und alternierend in den Zeilenvektoren.

Wie gesagt ist die Summe über alle Umordnungen groß und daher meistens nicht das Mittel der Wahl, wenn es um die Berechnung von Determinanten geht. Es dürfte nicht überraschend sein, dass der Gaußalgorithmus eine effektive Alternative ist.

**Folgerung 5.3** Bei einer elementaren Zeilentransformation  $A \rightarrow B$  ändert sich die Determinante wie folgt:

- (1) Multiplikation einer Zeile mit  $\lambda \in \mathbb{R}$ :  $\det(B) = \lambda \det(A)$
- (2) Vertauschung von zwei Zeilen:  $\det(B) = -\det(A)$

(3) *Addition eines Vielfachen einer Zeile zu einer anderen Zeile:*  $\det(B) = \det(A)$ .

**Satz 5.5 (Determinante von Dreiecksmatrizen)** *Die Determinante einer oberen Dreiecksmatrix  $D$  mit Diagonalelementen  $d_{11}, \dots, d_{nn}$  ist*

$$\det(D) = d_{11} \cdot \dots \cdot d_{nn}.$$

BEWEIS: Wir betrachten unsere Formel

$$\det(D) = \sum_{\sigma \in \mathbb{S}_n} \varepsilon(\sigma) d_{\sigma_1 1} \cdot \dots \cdot d_{\sigma_n n}.$$

Es ist  $\sigma_{ij} = 0$  für  $i > j$ . Also müssen wir nur die Umordnungen berücksichtigen, für die  $\sigma(k) \leq k$  für alle  $k$  ist. Das leistet aber nur  $\sigma = \text{id}$ , denn aus  $\sigma_i = i$  für  $i < k$  folgt  $\sigma_k \geq k$ , also gilt  $\sigma_k = k$  für alle  $k$ . Es folgt  $\det D = d_{11} \cdot \dots \cdot d_{nn}$ .  $\square$

Zur Berechnung der Determinante muss der Gaußalgorithmus also nur soweit durchgeführt werden, bis man eine obere Dreiecksmatrix hat. Zusätzlich muss das Produkt der anfallenden Faktoren  $\lambda$  bei Multiplikation bzw.  $-1$  bei Vertauschung von Zeilen abgespeichert werden, am Ende ist durch diese Zahl zu teilen.

**Beispiel 5.3** Eine beliebige Verfahren zur Berechnung der Determinante von  $3 \times 3$ -Matrizen ist die sogenannte Entwicklung nach einer Zeile oder Spalte. Zum Beispiel geht das Entwickeln nach der zweiten Spalte so:

$$\begin{aligned} \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} &= -b_1 \cdot \det \begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix} + b_2 \cdot \det \begin{pmatrix} a_1 & c_1 \\ a_3 & c_3 \end{pmatrix} - b_3 \cdot \det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \\ &= -b_1(a_2c_3 - a_3c_2) + b_2(a_1c_3 - a_3c_1) - b_3(a_1c_2 - a_2c_1) \\ &= a_1b_2c_3 + a_3b_1c_2 + a_2b_3c_1 - a_2b_1c_3 - a_3b_2c_1 - a_1b_3c_2. \end{aligned}$$

Streicht man die Zeile und Spalte, in der ein Matrixelement steht, so ergibt sich eine  $2 \times 2$ -Matrix und die zugehörige Subdeterminante. Die Determinante der  $3 \times 3$ -Matrix kann berechnet werden, indem man die Einträge einer Spalte (oder Zeile) mit den jeweiligen Subdeterminanten multipliziert und dann aufaddiert. Allerdings müssen die Summanden dabei ein Vorzeichen kriegen nach folgendem Schema:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

**Satz 5.6 (Laplaceentwicklung)** *Für  $A \in \mathbb{R}^{n \times n}$  bezeichne  $S_{ij}(A) \in \mathbb{R}^{(n-1) \times (n-1)}$  die Matrix, die durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte von  $A$  entsteht. Dann ergibt sich  $\det(A)$  durch Entwickeln nach der  $j$ -ten Spalte (oder der  $i$ -ten Zeile):*

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(S_{ij}(A)) \quad \left( = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(S_{ij}(A)) \right).$$

BEWEIS: Wir wollen nach der  $j$ -ten Spalte entwickeln, und führen folgende Notation ein:

$$\text{für } x \in \mathbb{R}^n : \quad x' = (x_1, \dots, \cancel{x_i}, \dots, x_n) \in \mathbb{R}^{n-1}$$

$$\text{für } y \in \mathbb{R}^{n-1} : \quad \bar{y} = (y_1, \dots, \overset{i}{0}, \dots, y_{n-1}) \in \mathbb{R}^n.$$

Es folgt  $\bar{x}' = (x_1, \dots, \overset{i}{0}, \dots, x_n) = x - x_i e_i$ . Seien  $a_1, \dots, a_n$  die Spalten von  $A$ . Dann gilt

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_{j-1}, \sum_{i=1}^n a_{ij} e_i, a_{j+1}, \dots, a_n) \\ &= \sum_{i=1}^n a_{ij} \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n) \\ &= \sum_{i=1}^n a_{ij} \det(\bar{a}'_1, \dots, \bar{a}'_{j-1}, e_i, \bar{a}'_{j+1}, \dots, \bar{a}'_n). \end{aligned}$$

Die letzte Gleichheit gilt, weil die Determinante verschwindet, wenn  $e_i$  zwei- oder mehrfach auftritt. Wir benötigen also die Determinante

$$\det(\bar{a}'_1, \dots, \bar{a}'_{j-1}, e_i, \bar{a}'_{j+1}, \dots, \bar{a}'_n) = \det \begin{pmatrix} a_{11} & \dots & \overset{j}{0} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ i & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Dazu betrachten wir die Funktion  $f : \mathbb{R}^{n-1} \times \dots \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$  mit

$$f(y_1, \dots, y_{n-1}) = \det(\bar{y}_1, \dots, \bar{y}_{j-1}, e_i, \bar{y}_j, \dots, \bar{y}_{n-1}).$$

$f$  ist multilinear und alternierend. Bezeichnet  $\varepsilon_1, \dots, \varepsilon_{n-1}$  die Standardbasis des  $\mathbb{R}^{n-1}$ , so gilt nach (5.9) für alle  $y \in \mathbb{R}^{n-1}$

$$f(y_1, \dots, y_{n-1}) = f(\varepsilon_1, \dots, \varepsilon_{n-1}) \det(y_1, \dots, y_{n-1}).$$

Nun ist  $\bar{\varepsilon}_k = e_k$  für  $k < i$ , und  $\bar{\varepsilon}_k = e_{k+1}$  für  $k \geq i$ . Daraus folgt

$$\begin{aligned} f(\varepsilon_1, \dots, \varepsilon_{n-1}) &= \begin{cases} \det(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_j, \overset{j}{e_i}, e_{j+1}, \dots, e_n) & \text{für } i < j \\ \det(e_1, \dots, e_{j-1}, \overset{j}{e_i}, e_j, \dots, e_{i-1}, e_{i+1}, \dots, e_n) & \text{für } i > j \\ \det(e_1, \dots, e_n) & \text{für } i = j \end{cases} \\ &= (-1)^{|i-j|} = (-1)^{i+j}. \end{aligned}$$

Also folgt

$$\det(\bar{a}'_1, \dots, \bar{a}'_{j-1}, e_i, \bar{a}'_{j+1}, \dots, \bar{a}'_n) = (-1)^{i+j} \det(a'_1, \dots, a'_{j-1}, a'_{j+1}, \dots, a'_n),$$

und insgesamt erhalten wir

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(a'_1, \dots, a'_{j-1}, a'_{j+1}, \dots, a'_n) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det S_{ij}(A).$$

□

**Satz 5.7 (Kofaktor-Matrix)** Für  $A \in \mathbb{R}^{n \times n}$  ist die Kofaktormatrix definiert durch

$$\operatorname{cof}(A) \in \mathbb{R}^{n \times n}, \operatorname{cof}(A)_{ij} = (-1)^{i+j} \det S_{ji}(A).$$

Es gilt  $A \operatorname{cof}(A) = \operatorname{cof}(A) A = \det(A) E_n$ . Insbesondere, wenn  $A$  invertierbar ist,

$$A^{-1} = \frac{1}{\det(A)} \operatorname{cof}(A).$$

BEWEIS: Wir berechnen, vgl. Beweis von Satz 5.6,

$$\begin{aligned} \det(A) \delta_{ik} &= \det(a_1, \dots, \overset{i}{a}_k, \dots, a_n) \\ &= \sum_{j=1}^n a_{jk} \det(a_1, \dots, \overset{i}{e}_j, \dots, a_n) \\ &= \sum_{j=1}^n a_{jk} (-1)^{i+j} \det S_{ji}(A) \\ &= \sum_{j=1}^n \operatorname{cof}(A)_{ij} a_{jk}. \end{aligned}$$

Also gilt  $\operatorname{cof}(A)A = \det(A)E_n$ . Nun ist leicht zu sehen, dass  $\operatorname{cof}(A^T) = \operatorname{cof}(A)^T$ . Mit  $\det(A^T) = \det(A)$  nach Satz 5.4 folgt

$$(A \operatorname{cof}(A))^T = \operatorname{cof}(A)^T A^T = \operatorname{cof}(A^T) A^T = \det(A^T) E_n = \det(A) E_n.$$

Dies zeigt die zweite Gleichung, der Satz ist bewiesen.  $\square$

**Satz 5.8 (Cramersche Regel)** Sei  $A \in \mathbb{R}^{n \times n}$  invertierbar mit Spalten  $a_j$ . Die Lösung des Gleichungssystems  $Ax = b$  lautet

$$x_j = \frac{\det(a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n)}{\det A} \quad \text{für } j = 1, \dots, n,$$

BEWEIS: Wir berechnen,

$$\begin{aligned} \det(a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n) &= \det(a_1, \dots, a_{j-1}, Ax, a_{j+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{j-1}, \sum_{k=1}^n x_k a_k, a_{j+1}, \dots, a_n) \\ &= x_j \det(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n). \end{aligned}$$

$\square$

Wir wollen zum Schluss des Abschnitts nochmal auf die Bedeutung der Determinante in der Volumenberechnung eingehen. Dabei setzen wir voraus, dass das Volumen vernünftig definiert ist. Sei  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  gegeben, und sei

$$Q = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \text{ für } i = 1, \dots, n\}.$$

Dann ist  $AQ$  ein Parallelotop mit den Kanten  $Ae_1, \dots, Ae_n$ . Wie bereits diskutiert, gilt

$$V(AQ) = |\det(Ae_1, \dots, Ae_n)| = |\det(A)|.$$

Betrachte für  $v \in \mathbb{R}^n$  und  $\lambda > 0$  den skalierten und verschobenen Würfel

$$v + \lambda Q = \{v + \lambda x : x \in Q\}.$$

Es folgt  $A(v + \lambda Q) = \{Av + \lambda Ax : x \in Q\} = Av + \lambda AQ$ , also

$$V(A(v + \lambda Q)) = V(Av + \lambda AQ) = \lambda^n V(AQ) = |\det(A)| V(v + \lambda Q).$$

Hier haben wir benutzt, dass das Volumen translationsinvariant ist und bei Streckung um den Faktor  $\lambda > 0$  mit der Potenz  $\lambda^n$  skaliert. Betrachte nun in  $\mathbb{R}^n$  alle Gitterwürfel mit ganzzahligen Ecken und Kantenlänge Eins. Bezeichne mit  $\mathcal{Q}_\lambda$  die Menge aller skalierten Gitterwürfel mit Kantenlänge  $\lambda > 0$ . Da die Seitenflächen der Würfel Volumen Null haben, erhalten wir für eine Vereinigung  $F$  von endlich vielen Würfeln aus  $\mathcal{Q}_\lambda$  durch Summation

$$V(AF) = |\det(A)| V(F).$$

Definiere nun für  $B \subset \mathbb{R}^n$  beschränkt eine Innen- und eine Außenfigur wie folgt:

$$\begin{aligned} F_\lambda &= \text{Vereinigung der Würfel in } \mathcal{Q}_\lambda, \text{ die in } B \text{ enthalten sind,} \\ F^\lambda &= \text{Vereinigung der Würfel in } \mathcal{Q}_\lambda, \text{ die } B \text{ treffen.} \end{aligned}$$

Wegen  $F_\lambda \subset B \subset F^\lambda$  erhalten wir

$$|\det(A)| V(F_\lambda) = V(AF_\lambda) \leq V(AB) \leq V(AF^\lambda) = |\det(A)| V(F^\lambda).$$

Wird die Gitterlänge verkleinert, so wird  $B$  genauer approximiert. Es ist zu erwarten, dass

$$V(F_\lambda), V(F^\lambda) \rightarrow V(B) \quad \text{mit } \lambda \searrow 0.$$

Das ist tatsächlich für jede Menge richtig, deren Rand nicht allzu exotisch ist.

**Satz 5.9 (Transformationsatz)** Für  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  und jede Menge  $B \subset \mathbb{R}^n$  gilt

$$V(AB) = |\det(A)| V(B).$$

## 6 Lineare Abbildungen und Koordinatendarstellungen

Wir beginnen diesen Abschnitt mit einem Einschub zur Geometrie, um mehr Beispiele zur Verfügung zu haben. Ein Euklidischer Vektorraum ist ein  $\mathbb{R}$ -Vektorraum  $V$  mit einem Skalarprodukt  $\langle v, w \rangle$ , das heißt es gilt (vgl. Kapitel I, Abschnitt 4)

$$\begin{aligned} \text{Symmetrie:} & \quad \langle v, w \rangle = \langle w, v \rangle \\ \text{Bilinearität:} & \quad \langle \lambda v_1 + \mu v_2, w \rangle = \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle \\ \text{Positivität:} & \quad \langle v, v \rangle = \|v\|^2 \geq 0, \quad \text{Gleichheit} \Leftrightarrow v = 0. \end{aligned}$$

In jedem Euklidischen Vektorraum hat man die Euklidische Norm (oder Länge)

$$\|v\| = \sqrt{\langle v, v \rangle} \quad \text{für } v \in V.$$

In Kapitel 1, Abschnitt 4, wurden folgende Eigenschaften der Norm gezeigt:

$$\begin{aligned} \text{Homogenität:} & \quad \|\lambda v\| = |\lambda| \|v\| \\ \text{Dreiecksungleichung:} & \quad \|v + w\| \leq \|v\| + \|w\| \\ \text{Positivität:} & \quad \|v\| \geq 0, \quad \text{Gleichheit} \Leftrightarrow v = 0. \end{aligned}$$

In einem Euklidischen Vektorraum sind gewisse Basen besonders praktisch.

**Definition 6.1** Eine Basis  $\mathcal{B} = \{e_1, \dots, e_n\}$  eines Euklidischen Vektorraums  $V$  heißt Orthonormalbasis, wenn

$$\langle e_i, e_j \rangle = \delta_{ij} \quad \text{für } i, j = 1, \dots, n.$$

Die Vektoren stehen also paarweise aufeinander senkrecht und sind normiert. Eine Menge solcher Vektoren (die nicht notwendig eine Basis bilden) nennt man Orthonormalsystem. Zum Beispiel ist  $\{e_1, e_2\} \subset \mathbb{R}^3$  ein Orthonormalsystem bezüglich des Standardskalarprodukts, und  $\{e_1, e_2, e_3\}$  ist eine Orthonormalbasis des  $\mathbb{R}^3$ .

**Beispiel 6.1** Auf dem Raum  $C^0([-\pi, \pi])$  betrachten wir das Riemann-Integral

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) dx.$$

Die Symmetrie und Bilinearität von  $\langle f, g \rangle$  sind klar. Weiter gilt

$$\langle f, f \rangle = \int_{-\pi}^{\pi} |f(x)|^2 dx \geq 0.$$

Ist  $f$  nicht die Nullfunktion, so ist  $\langle f, f \rangle > 0$  aus Stetigkeitsgründen, somit haben wir ein Skalarprodukt. Für  $u_k(x) = \frac{1}{\sqrt{\pi}} \cos kx$  sowie  $v_k(x) = \frac{1}{\sqrt{\pi}} \sin kx$  mit  $k \in \mathbb{N}$  ergibt sich

$$\begin{aligned} \langle u_k, u_\ell \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(kx) \cos(\ell x) dx = \delta_{k\ell} \\ \langle v_k, v_\ell \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(kx) \sin(\ell x) dx = \delta_{k\ell} \\ \langle u_k, v_\ell \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(kx) \sin(\ell x) dx = 0. \end{aligned}$$

Wir nehmen noch die konstante Funktion  $u_0(x) = \frac{1}{\sqrt{2\pi}}$  hinzu. Dann gilt

Die Funktionen  $u_k, v_\ell$  mit  $k \in \mathbb{N}_0$  und  $\ell \in \mathbb{N}$  bilden ein Orthonormalsystem.

Die endlichen Linearkombinationen der  $u_k, v_\ell$  nennt man trigonometrische Polynome. Die  $u_k, v_\ell$  bilden keine Basis von  $C^0([-\pi, \pi])$ , denn nicht jede stetige Funktion ist ein trigonometrisches Polynom. Zum Beispiel ist die Betragsfunktion  $f(x) = |x|$  kein trigonometrisches Polynom, denn diese Funktion ist nicht differenzierbar. Man kann aber zeigen: für jede Funktion  $f \in C^0([-\pi, \pi])$  gibt es trigonometrische Polynome  $f_n(x)$  der Form

$$f_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)) \quad \text{mit } a_k, b_k \in \mathbb{R},$$

so dass die Funktion  $f$  durch die  $f_n$  wie folgt approximiert wird:

$$\|f - f_n\| = \left( \int_{-\pi}^{\pi} |f(x) - f_n(x)|^2 dx \right)^{1/2} \rightarrow 0 \quad \text{mit } n \rightarrow \infty.$$

Tatsächlich ist das der Hauptsatz in der Theorie der Fourierreihen.

Der folgende Satz garantiert nicht nur die Existenz einer Orthonormalbasis, er liefert auch eine praktische Konstruktionsanleitung.

**Satz 6.1 (Gram-Schmidt Orthonormalisierungsverfahren)** *Jeder endlichdimensionale Euklidische Vektorraum  $V$  hat eine Orthonormalbasis.*

BEWEIS: Wähle eine beliebige Basis  $\{v_1, \dots, v_n\}$ . Wir konstruieren eine Orthonormalbasis  $\{e_1, \dots, e_n\}$  mit  $\text{Span}\{e_1, \dots, e_k\} = \text{Span}\{v_1, \dots, v_k\}$  für  $k = 1, \dots, n$ . Im ersten Schritt sei  $e_1 = \frac{v_1}{|v_1|}$ . Sind  $e_1, \dots, e_{k-1}$  schon gefunden, so bilden wir den Vektor

$$\tilde{e}_k := v_k - \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j.$$

Es ist  $\tilde{e}_k \neq 0$ , denn sonst wäre  $v_k$  eine Linearkombination von  $e_1, \dots, e_{k-1}$  und damit auch von  $v_1, \dots, v_{k-1}$  im Widerspruch zur Annahme. Wir können daher setzen

$$e_k = \frac{\tilde{e}_k}{|\tilde{e}_k|}.$$

Nun gilt  $\langle e_k, e_k \rangle = 1$  nach Definition. Für  $i < k$  berechnen wir induktiv

$$\langle e_i, \tilde{e}_k \rangle = \langle e_i, v_k \rangle - \sum_{j=1}^{k-1} \langle v_k, e_j \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}} = \langle e_i, v_k \rangle - \langle v_k, e_i \rangle = 0.$$

Es folgt  $\langle e_i, e_k \rangle = 0$  für  $i < k$ , und der Satz ist bewiesen. □

Wir kommen jetzt zum eigentlichen Thema des Abschnitts. Sei  $f : V \rightarrow W$  eine lineare Abbildung zwischen  $\mathbb{R}$ -Vektorräumen  $V, W$ , das heißt es gilt

$$f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2) \quad \text{für alle } v_{1,2} \in V, \lambda, \mu \in \mathbb{R}.$$

Bisher haben wir den Fall  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  betrachtet, und die Abbildung  $f \in L(\mathbb{R}^n, \mathbb{R}^m)$  mit einer Matrix  $F \in \mathbb{R}^{m \times n}$  identifiziert: der  $j$ -te Spaltenvektor von  $F$  ist der Vektor  $f(e_j) \in \mathbb{R}^m$ , also das Bild des  $j$ -ten Vektors der Standardbasis im  $\mathbb{R}^n$ . Es gilt dann

$$Fv = \sum_{j=1}^n v_j F e_j = \sum_{j=1}^n v_j f(e_j) = f(v).$$

Hier steht links die Multiplikation der Matrix  $F \in \mathbb{R}^{m \times n}$  mit dem Vektor  $v \in \mathbb{R}^n$ , rechts der Bildvektor  $f(v)$ . In diesem Sinn sind  $f$  und  $F$  ein- und dasselbe.

Bei Anwendungsproblemen befinden wir uns aber im Anschauungsraum, und wir müssen erst eine Basis wählen, um Koordinaten zu definieren. Oft ist die Wahl der Basis nicht zwingend, und es stellt sich die Frage wie unsere Beschreibung von der Wahl einer Basis abhängt. Selbst für lineare Abbildungen vom  $\mathbb{R}^n$  in den  $\mathbb{R}^m$  ist die Standardbasis nicht immer die beste Wahl.

**Beispiel 6.2** Sei  $v \in \mathbb{R}^3$  ein Einheitsvektor, also  $|v| = 1$ , und

$$s : \mathbb{R}^3 \rightarrow \mathbb{R}^3, s(x) = x - 2\langle v, x \rangle v.$$

Die Abbildung  $s(x)$  ist linear, denn es gilt

$$\begin{aligned} s(\lambda x_1 + \mu x_2) &= \lambda x_1 + \mu x_2 - 2\langle v, \lambda x_1 + \mu x_2 \rangle v \\ &= \lambda(x_1 - 2\langle v, x_1 \rangle v) + \mu(x_2 - 2\langle v, x_2 \rangle v) \\ &= \lambda s(x_1) + \mu s(x_2). \end{aligned}$$

Für die Standardbasis berechnen wir

$$s(e_j) = e_j - 2\langle v, e_j \rangle v = e_j - 2v_j \sum_{i=1}^3 v_i e_i = (1 - 2v_j^2)e_j - \sum_{i=1, i \neq j}^3 2v_i v_j,$$

das heißt  $s$  hat bezüglich der Standardbasis die Matrix

$$S = \begin{pmatrix} 1 - 2v_1^2 & -2v_1 v_2 & -2v_1 v_3 \\ -2v_2 v_1 & 1 - 2v_2^2 & -2v_2 v_3 \\ -2v_3 v_1 & -2v_3 v_2 & 1 - 2v_3^2 \end{pmatrix}$$

Wir können die Abbildung jedoch besser in einer anderen Basis verstehen, und zwar sei  $\mathcal{B} = \{v_1, v_2, v_3\}$  mit  $v_1 = v$  und  $v_{2,3} \perp v$ . Dann gilt

$$s(v_j) = v_j - 2\langle v, v_j \rangle v = \begin{cases} -v_j & \text{falls } j = 1, \\ v_j & \text{falls } j = 2, 3. \end{cases}$$

Tragen wir die Koordinaten von  $S(v_j)$  bezüglich der Basis  $\mathcal{B}$  in die  $j$ -te Spalte einer Matrix ein, so ergibt sich

$$S_{\mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Es ist nicht zu bestreiten, dass diese Matrix einfacher ist und besser wiedergibt wie  $s$  abbildet:  $s$  ist eine Spiegelung an der Ebene senkrecht zu  $v_1 = v$ .

**Beispiel 6.3** Betrachte den Lösungsraum

$$V = \{u \in C^2(\mathbb{R}) : u'' - u = 0\}.$$

Man kann zeigen, dass  $V$  ein zweidimensionaler Vektorraum ist; eine Basis ist

$$\mathcal{A} = \{\cosh(x), \sinh(x)\}.$$

Der Ableitungsoperator  $\frac{d}{dx}$  bildet  $V$  in sich ab und ist linear, genauer gilt

$$\frac{d}{dx} \cosh(x) = \sinh(x) \quad \text{und} \quad \frac{d}{dx} \sinh(x) = \cosh(x).$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix so erhalten wir die Darstellung

$$\left(\frac{d}{dx}\right)_{\mathcal{A}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Es geht aber noch einfacher. Bezüglich der Basis  $\mathcal{B} = \{e^x, e^{-x}\}$  hat  $\frac{d}{dx}$  die Matrixdarstellung

$$\left(\frac{d}{dx}\right)_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die Beispiele zeigen, dass die Wahl einer passenden Basis hilfreich sein kann, um eine lineare Abbildung zu verstehen. Wir wollen das jetzt systematisieren.

**Definition 6.2 (Matrix bzgl. gegebener Basen)** Seien  $V, W$  Vektorräume mit Basen  $\mathcal{A} = \{v_1, \dots, v_n\}$  und  $\mathcal{B} = \{w_1, \dots, w_m\}$ . Wir ordnen einer linearen Abbildung  $f : V \rightarrow W$  eine Matrix  $f_{\mathcal{B}\mathcal{A}} \in \mathbb{R}^{m \times n}$  zu, indem wir den Vektor  $f(v_j)$  in die Basis  $w_1, \dots, w_m$  entwickeln und die Koeffizienten in die  $j$ -te Spalte eintragen:

$$f(v_j) = \sum_{i=1}^m F_{ij} w_i \quad \text{für } F = f_{\mathcal{B}\mathcal{A}}.$$

Ein Vektor  $x \in V$  hat bzgl. der Basis  $\mathcal{A} = \{v_1, \dots, v_n\}$  den zugehörigen Koordinatenvektor

$$x_{\mathcal{A}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{wobei } x = \sum_{j=1}^n x_j v_j. \quad (6.12)$$

**Beispiel 6.4** Sei  $\mathcal{E} = \{e_1, e_2\}$  die Standardbasis und  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$f_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wir betrachten die (ebenfalls orthonormale Basis)  $\mathcal{A} = \{v_1, v_2\}$  mit

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Dann ist  $f(v_1) = 3v_1$  sowie  $f(v_2) = v_2$ , die Matrixdarstellung lautet also

$$f_{\mathcal{A}\mathcal{A}} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Eine lineare Abbildung kann nach wie vor als Multiplikation von Matrix mit Koordinatenvektor interpretiert werden. Genauer gilt:

**Satz 6.2 (Lineare Abbildungen operieren als Matrizen)** Seien  $V, W$  Vektorräume mit Basen  $\mathcal{A} = \{v_1, \dots, v_n\}$  und  $\mathcal{B} = \{w_1, \dots, w_m\}$ . Ist  $f : V \rightarrow W$  linear, so gilt

$$f(x)_{\mathcal{B}} = f_{\mathcal{B}\mathcal{A}} \cdot x_{\mathcal{A}}.$$

BEWEIS: Mit  $x = \sum_{j=1}^n x_j v_j$  und  $F = f_{\mathcal{B}\mathcal{A}} \in \mathbb{R}^{m \times n}$  gilt

$$f(x) = \sum_{j=1}^n x_j f(v_j) = \sum_{j=1}^n x_j \sum_{i=1}^m F_{ij} w_i = \sum_{i=1}^m \left( \sum_{j=1}^n F_{ij} x_j \right) w_i.$$

Das bedeutet  $(f(x)_{\mathcal{B}})_i = (F \cdot x_{\mathcal{A}})_i$  bzw.  $f(x)_{\mathcal{B}} = f_{\mathcal{B}\mathcal{A}} \cdot x_{\mathcal{A}}$ . □

Das ist ganz analog zur Situation mit Standardbasen, nur muss klar sein, welche Basis zugrundeliegt. Ist  $\mathcal{A} = \{v_1, \dots, v_n\}$  eine (nichtstandard-) Basis des  $\mathbb{R}^n$ , so hat ein Vektor  $x \in \mathbb{R}^n$  zum einen die Koordinaten bzgl. der Standardbasis  $e_1, \dots, e_n$ , also  $x = \sum_{j=1}^n x_j e_j$ , wird aber bzgl. der Basis  $\mathcal{A}$  durch einen anderen Koordinatenvektor  $x' = x_{\mathcal{A}} \in \mathbb{R}^n$  dargestellt, also  $x = \sum_{j=1}^n x'_j v_j$ .

**Beispiel 6.5** Wir betrachten weiter das Beispiel 6.4. Die Vektoren der Standardbasis lassen sich in die Basis  $\mathcal{A} = \{v_1, v_2\}$  entwickeln, es gilt

$$e_1 = \frac{1}{\sqrt{2}}(v_1 - v_2) \quad \text{und} \quad e_2 = \frac{1}{\sqrt{2}}(v_1 + v_2).$$

Für einen beliebigen Vektor  $x \in \mathbb{R}^2$  folgt daraus

$$x = x_1 e_1 + x_2 e_2 = \frac{1}{\sqrt{2}}((x_1 + x_2)v_1 + (x_2 - x_1)v_2),$$

das heißt der Koordinatenvektor von  $x$  bzgl.  $\mathcal{A}$  ist

$$x_{\mathcal{A}} = \frac{1}{\sqrt{2}} \begin{pmatrix} x_1 + x_2 \\ x_2 - x_1 \end{pmatrix}.$$

Nach Satz 6.2 hat der Bildvektor  $f(x)$  nun bzgl.  $\mathcal{A}$  den Koordinatenvektor

$$f(x)_{\mathcal{A}} = f_{\mathcal{A}\mathcal{A}} \cdot x_{\mathcal{A}} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} x_1 + x_2 \\ x_2 - x_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3(x_1 + x_2) \\ x_2 - x_1 \end{pmatrix}.$$

Stimmt das? Wir entwickeln dazu wieder zurück in die Standardbasis:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2}}(3(x_1 + x_2)v_1 + (x_2 - x_1)v_2) \\ &= \frac{1}{\sqrt{2}}\left(3(x_1 + x_2)\frac{1}{\sqrt{2}}(e_1 + e_2) + (x_2 - x_1)\frac{1}{\sqrt{2}}(e_2 - e_1)\right) \\ &= (2x_1 + x_2)e_1 + (x_1 + 2x_2)e_2. \end{aligned}$$

Das ist in der Tat die Abbildung  $f(x)$  in der Standardbasis.

**Satz 6.3 (Kürzungsregel für Basen)** Seien  $f : U \rightarrow V$  und  $g : V \rightarrow W$  linear, und seien  $\mathcal{A}$ ,  $\mathcal{B}$  und  $\mathcal{C}$  Basen der Räume  $U$ ,  $V$  und  $W$ . Dann gilt für die zugehörigen Matrizen

$$(g \circ f)_{\mathcal{C}\mathcal{A}} = g_{\mathcal{C}\mathcal{B}} \cdot f_{\mathcal{B}\mathcal{A}}.$$

BEWEIS: Seien  $\dim U = n$ ,  $\dim V = m$  und  $\dim W = \ell$ . Wir schreiben

$$\begin{aligned} \mathcal{A} &= \{a_1, \dots, a_n\} \\ \mathcal{B} &= \{b_1, \dots, b_m\} \\ \mathcal{C} &= \{c_1, \dots, c_\ell\}. \end{aligned}$$

Mit  $F = f_{\mathcal{B}\mathcal{A}}$  und  $G = g_{\mathcal{C}\mathcal{B}}$  gilt dann

$$f(a_k) = \sum_{j=1}^m F_{jk} b_j \quad \text{und} \quad g(b_j) = \sum_{i=1}^{\ell} G_{ij} c_i.$$

Somit ergibt sich

$$(g \circ f)(a_k) = \sum_{j=1}^m F_{jk} g(b_j) = \sum_{j=1}^m F_{jk} \sum_{i=1}^{\ell} G_{ij} c_i = \sum_{i=1}^{\ell} \left( \sum_{j=1}^m G_{ij} F_{jk} \right) c_i = \sum_{i=1}^{\ell} (GF)_{ik} c_i.$$

Also ist  $(g \circ f)_{\mathcal{C}\mathcal{A}} = GF = g_{\mathcal{C}\mathcal{B}} \cdot f_{\mathcal{B}\mathcal{A}}$  wie behauptet.  $\square$

**Satz 6.4 (Matrix der inversen Abbildung)** Sei  $f : V \rightarrow W$  linear und invertierbar. Ist  $\mathcal{A}$  Basis von  $V$  und  $\mathcal{B}$  Basis von  $W$ , so folgt

$$(f^{-1})_{\mathcal{A}\mathcal{B}} = (f_{\mathcal{B}\mathcal{A}})^{-1}.$$

BEWEIS: Es gilt mit  $\dim V = \dim W = n$

$$E_n = (f^{-1} \circ f)_{\mathcal{A}\mathcal{A}} = (f^{-1})_{\mathcal{A}\mathcal{B}} f_{\mathcal{B}\mathcal{A}}.$$

□

Wir kommen nun zu einer wichtigen Klasse von Abbildungen.

**Definition 6.3** Sei  $V$  ein endlichdimensionaler Vektorraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Eine lineare Abbildung  $f : V \rightarrow V$  heißt orthogonal, wenn gilt:

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \text{für alle } x, y \in V.$$

Für eine orthogonale Abbildung  $f : V \rightarrow V$  gilt

$$\|f(x)\|^2 = \langle f(x), f(x) \rangle = \langle x, x \rangle = \|x\|^2.$$

Insbesondere ist  $f$  injektiv und wegen  $\dim V < \infty$  auch surjektiv, also invertierbar. Weiter ist  $f$  isometrisch (abstandstreu), denn

$$\|f(x) - f(y)\|^2 = \|f(x - y)\|^2 = \|x - y\|^2.$$

Umgekehrt folgt aus isometrisch auch orthogonal, wie man zeigen kann. Sind  $f, g : V \rightarrow V$  orthogonal, so auch  $g \circ f$  und  $f^{-1}$ . Man bezeichnet die Menge der orthogonalen Abbildungen als die orthogonale Gruppe von  $V$ .

Es stellt sich die Frage, ob wir an der Matrixdarstellung der Abbildung  $f$  sehen können, ob  $f$  orthogonal ist. Die Antwort heißt nein, sofern wir eine beliebige Basis verwenden. Nehmen wir aber eine Orthonormalbasis, so ist das anders.

**Satz 6.5** Sei  $V$  ein Euklidischer Vektorraum mit Orthonormalbasis  $\mathcal{A} = \{v_1, \dots, v_n\}$ . Für eine lineare Abbildung  $f : V \rightarrow V$  mit Matrix  $F$  bezüglich  $\mathcal{A}$  sind folgende Aussagen äquivalent:

- (1)  $f$  ist orthogonal.
- (2) Die Spalten von  $F$  sind eine Orthonormalbasis bezüglich des Standardskalarprodukts.

BEWEIS: Sei  $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$  das Standardskalarprodukt. Für  $v = \sum_{i=1}^n x_i v_i$ ,  $w = \sum_{i=1}^n y_i v_i$  gilt

$$\langle v, w \rangle = \sum_{i,j=1}^n x_i y_j \underbrace{\langle v_i, v_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^n x_i y_i = \langle x, y \rangle_{\mathbb{R}^n}.$$

Nun ist  $(v_i)_{\mathcal{A}} = e_i$  und somit  $f(v_i)_{\mathcal{A}} = F e_i$  nach Satz 6.2, also folgt

$$\langle f(v_i), f(v_j) \rangle = \langle F e_i, F e_j \rangle_{\mathbb{R}^n}.$$

Ist nun  $f$  orthogonal, so erhalten wir

$$\langle Fe_i, Fe_j \rangle_{\mathbb{R}^n} = \langle f(v_i), f(v_j) \rangle = \langle v_i, v_j \rangle = \delta_{ij}.$$

Ist umgekehrt  $\langle Fe_i, Fe_j \rangle_{\mathbb{R}^n} = \delta_{ij}$ , so folgt  $\langle f(v_i), f(v_j) \rangle = \delta_{ij} = \langle v_i, v_j \rangle$ , und hieraus

$$\langle f(v), f(w) \rangle = \sum_{i,j=1}^n x_i y_j \langle f(v_i), f(v_j) \rangle = \sum_{i,j=1}^n x_i y_j \langle v_i, v_j \rangle = \langle v, w \rangle.$$

□

Die Skalarprodukte  $\langle Fe_i, Fe_j \rangle_{\mathbb{R}^n}$  sind genau die Einträge der Matrix  $F^T F$ , das heißt die Eigenschaft aus dem Satz ist gleichbedeutend mit  $F^T F = E_n$ . Daraus folgt weiter  $F^{-1} = F^T$  und insbesondere  $FF^T = E_n$ , das heißt auch die Zeilenvektoren bilden eine Orthonormalbasis.

**Definition 6.4** Die Gruppe der orthogonalen Matrizen  $Q \in \mathbb{R}^{n \times n}$  ist

$$\mathbb{O}(n) = \{Q \in \mathbb{R}^{n \times n} : Q^T Q = E_n\}.$$

Die spezielle orthogonale Gruppe ist

$$\mathbb{S}O(n) = \{Q \in \mathbb{O}(n) : \det(Q) = 1\}.$$

*Bemerkung.* Für  $Q \in \mathbb{O}(n)$  gilt stets  $\det(Q) = \pm 1$ , denn

$$1 = \det(Q^T Q) = \det(Q) \det(Q^T) = |\det(Q)|^2. \quad (6.13)$$

**Beispiel 6.6** Sei  $Q \in \mathbb{O}(2)$ , das heißt die Spalten  $Qe_1, Qe_2$  sind eine Orthonormalbasis. Mit geeignetem  $\varphi \in [0, 2\pi)$  gilt dann

$$Qe_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \quad \text{und} \quad Qe_2 = \begin{pmatrix} \mp \sin \varphi \\ \pm \cos \varphi \end{pmatrix},$$

Das obere Vorzeichen bedeutet  $\det(Q) = +1$ , und  $Q$  ist eine Drehmatrix. Das untere Vorzeichen bedeutet  $\det(Q) = -1$ ; wir kommen auf diesen Fall bald zurück.

**Beispiel 6.7 (Eulersche Winkel)** Wir wollen  $Q \in \mathbb{S}O(3)$  als Produkt von Drehmatrizen der folgenden Form darstellen:

$$D_3(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad D_1(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$$

Sei  $\mathcal{E} = \{e_1, e_2, e_3\}$  die Standardbasis und  $\mathcal{E}''' = \{Qe_1, Qe_2, Qe_3\}$  die Basis der Bildvektoren. Die Ebenen senkrecht zu  $e_3$  und  $Qe_3$  schneiden sich (im allgemeinen) in einer Geraden  $L$ , die als Knotenlinie bezeichnet wird. Wir wählen eine positiv orientierte Orthonormalbasis  $\mathcal{E}' = \{e'_1, e'_2, e'_3\}$  mit  $e'_3 = e_3$  und  $e'_1 \in L$ . Es gilt dann

$$Qe_3 = -\sin \theta e'_2 + \cos \theta e'_3.$$

Wir können annehmen dass  $\langle Qe_3, e'_2 \rangle \leq 0$ , andernfalls ersetzen wir  $e'_1, e'_2$  durch  $-e'_1, -e'_2$ . Somit ist  $\theta \in [0, \pi]$ . Definiere nun die positiv orientierte Basis  $\mathcal{E}'' = \{e''_1, e''_2, e''_3\}$  mit  $e''_1 = e'_1$  und  $e''_3 = Qe_3$ , das heißt  $\text{id}_{\mathcal{E}'\mathcal{E}''} = D_1(\theta)$ . Wegen  $e'_3 = e_3$  und  $e'''_3 = e''_3$  gilt (vgl. Beispiel 6.6)

$$\text{id}_{\mathcal{E}\mathcal{E}'} = D_3(\varphi) \quad \text{und} \quad \text{id}_{\mathcal{E}''\mathcal{E}'''} = D_3(\psi) \quad \text{mit} \quad \varphi, \psi \in [0, 2\pi).$$

Wir erhalten die Darstellung von  $Q$  als Produkt

$$Q = \text{id}_{\mathcal{E}\mathcal{E}'''} = \text{id}_{\mathcal{E}\mathcal{E}'} \text{id}_{\mathcal{E}'\mathcal{E}''} \text{id}_{\mathcal{E}''\mathcal{E}'''} = D_3(\varphi)D_1(\theta)D_3(\psi).$$

Wir werden demnächst zeigen, dass jedes  $Q \in \text{SO}(3)$  eine Drehung im eigentlich Sinn ist, das heißt es gibt eine Drehachse die invariant bleibt, und auf der dazu senkrechten Ebene bildet  $Q$  als zweidimensionale Drehung ab.

Zum Schluss des Abschnitts wollen wir untersuchen, wie sich die Abbildungsmatrix transformiert, wenn wir zu einer anderen Basis und damit zu anderen Koordinaten übergehen.

**Definition 6.5** Sei  $\mathcal{A} = \{v_1, \dots, v_n\}$  eine Basis des Vektorraums  $V$ , und  $\mathcal{B} = \{w_1, \dots, w_n\}$  sei eine neue Basis. Die Matrix

$$C = \text{id}_{\mathcal{A}\mathcal{B}} \in \mathbb{R}^{n \times n}$$

heißt Transformationsmatrix des Basiswechsels von  $\mathcal{A}$  nach  $\mathcal{B}$ .

**Beispiel 6.8** In der  $j$ -ten Spalte der Transformationsmatrix stehen die Koordinaten des neuen Basisvektors  $w_j$  bezüglich der alten Basis  $v_1, \dots, v_n$ . Sei zum Beispiel  $\mathcal{A} = \{e_1, e_2\}$  die Standardbasis des  $\mathbb{R}^2$  und  $\mathcal{B} = \{w_1, w_2\}$  mit

$$w_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

Dann lautet die Transformationsmatrix

$$\text{id}_{\mathcal{A}\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

**Satz 6.6** Sei  $\mathcal{A} = \{v_1, \dots, v_n\}$  eine Basis des Vektorraums  $V$ , und  $\mathcal{B} = \{w_1, \dots, w_n\}$  sei eine neue Basis. Sei  $C = \text{id}_{\mathcal{A}\mathcal{B}}$  die zugehörige Transformationsmatrix. Dann gilt:

- (1)  $x_{\mathcal{B}} = C^{-1}x_{\mathcal{A}}$  für alle  $x \in V$ .
- (2) Hat  $f : V \rightarrow V$  bezüglich  $\mathcal{A}, \mathcal{B}$  die Matrizen  $A, B$ , so gilt  $B = C^{-1}AC$ .

BEWEIS: Aus Satz 6.2, Satz 6.3 und Folgerung 6.4 folgt

$$\begin{aligned} x_{\mathcal{B}} &= \text{id}_{\mathcal{B}\mathcal{A}} x_{\mathcal{A}} = C^{-1}x_{\mathcal{A}} \\ B &= f_{\mathcal{B}\mathcal{B}} = \text{id}_{\mathcal{B}\mathcal{A}} f_{\mathcal{A}\mathcal{A}} \text{id}_{\mathcal{A}\mathcal{B}} = C^{-1}AC. \end{aligned}$$

□

**Beispiel 6.9** Wir beschäftigen uns weiter mit Beispiel 6.4. Die Transformationsmatrix des Basiswechsels von  $\mathcal{E} = \{e_1, e_2\}$  nach  $\mathcal{A} = \{v_1, v_2\}$  lautet

$$\text{id}_{\mathcal{E}\mathcal{A}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Nach Satz 6.6 hat die Abbildung  $f$  bezüglich der Basis  $\mathcal{A}$  somit die Matrix

$$f_{\mathcal{A}\mathcal{A}} = \left(\frac{1}{\sqrt{2}}\right)^{-1} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Dieses Ergebnis hatten wir schon oben ausgerechnet.

**Definition 6.6 (Ähnlichkeit von Matrizen)** Zwei Matrizen  $A, B \in \mathbb{R}^{n \times n}$  heißen *ähnlich*, wenn eine invertierbare Matrix  $C \in \mathbb{R}^{n \times n}$  existiert mit

$$B = C^{-1}AC.$$

Um eine gegebene lineare Abbildung  $f : V \rightarrow V$  zu verstehen, ist eine Basis zu finden, in der die Matrixdarstellung von  $f$  möglichst einfach ist. Dies ist gleichbedeutend damit, zu einer gegebenen Matrix  $A$  eine ähnliche Matrix  $B$  zu finden, die möglichst einfach ist. In interessanten Fällen wird es gelingen, die Matrix  $B$  als Diagonalmatrix zu bestimmen. Im Bezug auf den Wunsch nach Einfachheit ist dies das Optimum. Im allgemeinen kann  $B$  nicht als Diagonalmatrix gewählt werden, jedoch kommen andere Normalformen in Betracht.

## 7 Eigenwerte und Normalformen

Wir beginnen mit folgendem zentralen Begriff.

**Definition 7.1**  $\lambda \in \mathbb{R}$  heißt *Eigenwert* von  $F \in \mathbb{R}^{n \times n}$ , wenn es ein  $v \in \mathbb{R}^n$ ,  $v \neq \mathbf{0}$ , gibt mit

$$Fv = \lambda v.$$

Ein solches  $v$  heißt *Eigenvektor* von  $F$  zum Eigenwert  $\lambda$ .

Der Zusatz  $v \neq \mathbf{0}$  ist wichtig, denn für jedes  $\lambda \in \mathbb{R}$  gilt  $A \cdot \mathbf{0} = \lambda \cdot \mathbf{0}$ . Ohne die Bedingung wäre jedes  $\lambda \in \mathbb{R}$  ein Eigenwert von  $A$ , und das wäre sinnlos.

**Beispiel 7.1** Für  $F = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$  gilt

$$\begin{aligned} Fv_1 &= 3v_1 & \text{für } v_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ Fv_2 &= v_2 & \text{für } v_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \end{aligned}$$

Also sind  $v_1$  und  $v_2$  Eigenvektoren von  $F$  mit den Eigenwerten  $\lambda_1 = 3$  und  $\lambda_2 = 1$ .

Es stellt sich die Frage, ob jede quadratische Matrix Eigenwerte hat und wie wir diese ggf. ausrechnen können.

**Definition 7.2** Das charakteristische Polynom einer Matrix  $F \in \mathbb{R}^{n \times n}$  ist

$$\chi_F(\lambda) = \det(\lambda E_n - F).$$

Die Diagonalelemente von  $\lambda E_n - F$  enthalten jeweils ein  $\lambda$ , die übrigen Einträge enthalten kein  $\lambda$ . Die Determinante ist eine Summe von Produkten von jeweils  $n$  Einträgen der Matrix. Deshalb ist  $\chi_F(\lambda)$  ein Polynom in  $\lambda$  vom Grad höchstens  $n$ . Genauer sehen wir

$$\chi_F(\lambda) = (\lambda - F_{11}) \cdot \dots \cdot (\lambda - F_{nn}) + \sum_{\sigma \neq \text{id}} \varepsilon(\sigma) (\lambda \delta_{\sigma_1 1} - F_{\sigma_1 1}) \cdot \dots \cdot (\lambda \delta_{\sigma_n n} - F_{\sigma_n n}).$$

Für eine Permutation  $\sigma \neq \text{id}$  gilt  $\sigma_i \neq i$  für mindestens zwei  $i$ , die hintere Summe liefert also höchstens Potenzen  $\lambda^i$  mit  $i \leq n - 2$ . Wegen  $\chi_F(0) = \det(-F) = (-1)^n \det(F)$  ergibt sich

$$\chi_F(\lambda) = \lambda^n - \text{tr}(F)\lambda^{n-1} + \dots + (-1)^n \det(F).$$

**Satz 7.1** Die Eigenwerte von  $F \in \mathbb{R}^{n \times n}$  sind genau die Nullstellen des charakteristischen Polynoms  $\chi_F(\lambda)$ .

BEWEIS: Sei  $\lambda \in \mathbb{R}$  mit  $\chi_F(\lambda) = 0$ , also

$$\det(F - \lambda E_n) = (-1)^n \det(\lambda E_n - F) = 0.$$

Nach Satz 5.3 ist  $F - \lambda E_n$  nicht invertierbar. Nach der Dimensionsformel, Satz 2.1, hat  $F - \lambda E_n$  dann einen nichttrivialen Kern, das heißt es gibt ein  $v \in V$ ,  $v \neq 0$ , mit  $(F - \lambda E_n)v = 0$  beziehungsweise  $Fv = \lambda v$ . Somit ist  $\lambda$  ein Eigenwert. Durch Zurückverfolgen des Arguments ergibt sich die umgekehrte Aussage.  $\square$

**Beispiel 7.2** Betrachte eine Matrix  $Q \in \mathbb{O}(2)$ , also

$$Q = \begin{pmatrix} \cos \theta & \mp \sin \theta \\ \sin \theta & \pm \cos \theta \end{pmatrix} \quad \text{mit } \theta \in [0, 2\pi).$$

Im Fall des oberen Vorzeichens, also  $\det(Q) = 1$  bzw.  $Q \in \mathbb{SO}(2)$ , ergibt sich

$$\chi_Q(\lambda) = \det \begin{pmatrix} \lambda - \cos \theta & \sin \theta \\ -\sin \theta & \lambda - \cos \theta \end{pmatrix} = \lambda^2 - 2\lambda \cos \theta + 1 = (\lambda - \cos \theta)^2 + \sin^2 \theta.$$

Für  $\theta \neq 0, \pi$  hat  $Q$  somit keinen Eigenwert. Für  $\theta = 0$  ist  $\lambda = 1$  einziger Eigenwert; es gilt dann  $Q = E_2$  und jeder Vektor  $v \in \mathbb{R}^2$ ,  $v \neq 0$ , ist Eigenvektor. Analog ist  $\lambda = -1$  einziger Eigenwert im Fall  $\theta = \pi$ , es ist  $Q = -E_2$  und wieder sind alle  $v \in \mathbb{R}^2$ ,  $v \neq 0$ , Eigenvektoren.

Im anderen Fall, also  $\det(Q) = -1$  bzw.  $Q \in \mathbb{O}(2) \setminus \mathbb{SO}(2)$ , gilt dagegen

$$\chi_Q(\lambda) = \det \begin{pmatrix} \lambda - \cos \theta & -\sin \theta \\ -\sin \theta & \lambda + \cos \theta \end{pmatrix} = \lambda^2 - 1.$$

Somit hat  $Q$  die beiden Eigenwerte  $\lambda_{1,2} = \pm 1$ . Um für  $\lambda = 1$  einen zugehörigen Eigenvektor  $v = (x, y)$  zu bestimmen, lösen wir das lineare Gleichungssystem  $Qv_1 = v_1$  oder

$$\begin{pmatrix} \cos \theta - 1 & \sin \theta \\ \sin \theta & -\cos \theta - 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Mit  $\cos \theta = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}$  und  $\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}$  erhalten wir

$$\begin{pmatrix} -2 \sin^2 \frac{\theta}{2} & 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} & -2 \cos^2 \frac{\theta}{2} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir erhalten eine Lösung  $v_1 = (\cos \frac{\theta}{2}, \sin \frac{\theta}{2}) \neq 0$ . Analog bestimmen wir einen Eigenvektor  $v_2 \neq 0$  zu  $\lambda = -1$  durch Lösen des Gleichungssystems

$$\begin{pmatrix} \cos \theta + 1 & \sin \theta \\ \sin \theta & -\cos \theta + 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

bzw. nach Umformung

$$\begin{pmatrix} 2 \cos^2 \frac{\theta}{2} & 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} & 2 \sin^2 \frac{\theta}{2} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir können  $v_2 = (-\sin \frac{\theta}{2}, \cos \frac{\theta}{2})$  wählen, damit ist  $\mathcal{A} = \{v_1, v_2\}$  sogar eine Orthonormalbasis. Bezüglich  $\mathcal{A}$  hat  $Q$  die Matrix

$$Q_{\mathcal{A}\mathcal{A}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

mit anderen Worten ist  $Q$  die Spiegelung an der Geraden  $\mathbb{R} \cdot v_1$ . Insgesamt können wir sagen, dass die Matrizen  $Q \in \mathbb{O}(2)$  entweder Drehmatrizen oder Geradenspiegelungen sind, je nach Vorzeichen der Determinante.

Im Beispiel haben wir erst die Eigenwerte als Nullstellen des charakteristischen Polynoms ausgerechnet, dann haben wir Eigenvektoren durch Lösen des zugehörigen linearen Gleichungssystems bestimmt. Dies ist allgemein ein guter Ansatz, um die Eigenwerte und Eigenvektoren zu bestimmen.

**Beispiel 7.3** Das charakteristische Polynom einer  $2 \times 2$ -Matrix  $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  ist

$$\det(\lambda E_2 - F) = \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - \operatorname{tr}(F)\lambda + \det(F).$$

$F$  hat einen Eigenwert  $\lambda \in \mathbb{R}$  genau wenn

$$0 \leq (\operatorname{tr} F)^2 - 4 \det(F) = (a - d)^2 + 4bc.$$

**Beispiel 7.4** Jede Matrix  $F \in \mathbb{R}^{3 \times 3}$  hat mindestens einen Eigenwert  $\lambda \in \mathbb{R}$ . Denn  $\chi_F(\lambda)$  ist ein Polynom dritten Grades, also gilt

$$\chi_F(\lambda) \rightarrow \pm\infty \quad \text{für } \lambda \rightarrow \pm\infty.$$

Nach dem Zwischenwertsatz gibt es dann ein  $\lambda \in \mathbb{R}$  mit  $\chi_F(\lambda) = 0$ .

**Beispiel 7.5** Sei  $Q \in \mathbb{O}(3)$ , also  $Q$  orthogonal. Dann erhält  $Q : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  das Standardskalarprodukt:

$$\langle Qx, Qy \rangle = \langle Q^T Qx, y \rangle = \langle E_3 x, y \rangle = \langle x, y \rangle.$$

Insbesondere gilt  $|Qx| = |x|$  für alle  $x \in \mathbb{R}^3$ . Nach dem vorigen Beispiel hat  $Q$  (mindestens) einen Eigenwert  $\lambda \in \mathbb{R}$ . Ist  $v \neq 0$  zugehöriger Eigenvektor, so folgt

$$|\lambda| |v| = |\lambda v| = |Qv| = |v|,$$

also  $\lambda = \pm 1$ . Daraus folgt

$$Q(\lambda v) = \lambda Qv = \lambda^2 v = v \quad \text{bzw.} \quad Q^{-1}v = \lambda v.$$

Da  $Q^T = Q^{-1}$  erhalten wir für  $w \in \mathbb{R}^3$  mit  $\langle v, w \rangle = 0$

$$\langle Qw, v \rangle = \langle w, Q^T v \rangle = \langle w, Q^{-1}v \rangle = \langle w, \lambda v \rangle = 0.$$

Somit bildet  $Q$  die Ebene  $E$  senkrecht zu  $v$  in sich ab. Wir können  $|v| = 1$  annehmen, und wählen eine positiv orientierte Orthonormalbasis  $\mathcal{A} = \{v_1, v_2, v_3\}$  mit  $v_3 = v$ . Da  $Q$  orthogonal, ist die Matrix der Abbildung  $Q|_E : E \rightarrow E$  bezüglich der Orthonormalbasis  $\{v_1, v_2\}$  in  $\mathbb{O}(2)$ , siehe Satz 6.5. Es gibt nun zwei Fälle:

**Fall 1:**  $Q|_E$  ist eine Drehmatrix, das heißt bezüglich  $v_1, v_2, v_3$  hat  $Q$  die Matrix

$$Q_{\mathcal{A}\mathcal{A}} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{mit } \theta \in [0, 2\pi).$$

Es gilt  $\det(Q) = \lambda$ . Im Fall  $\lambda = 1$  beschreibt  $Q$  eine Drehung um die Achse  $v_3$  mit dem orientierten Drehwinkel  $\theta$ . Ist  $\lambda = -1$ , so folgt auf diese Drehung noch eine Spiegelung an der Ebene  $E = \text{Span}\{v_1, v_2\}$ .

**Fall 2:** Die Matrix von  $Q|_E$  beschreibt eine Geradenspiegelung. In diesem Fall können wir die Vektoren  $v_1, v_2$  so wählen, dass  $Qv_1 = v_1$  und  $Qv_2 = -v_2$  ist, also hat  $Q$  insgesamt die Matrix

$$Q_{\mathcal{A}\mathcal{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Es gilt  $\det(Q) = -\lambda$ . Im Fall  $\lambda = -1$  beschreibt  $Q$  eine Drehung um die Achse  $v_1$  mit Drehwinkel  $\pi$ . Im Fall  $\lambda = 1$  erhalten wir die Spiegelung an der Ebene, die von  $v_1, v_3$  aufgespannt wird. Gehen wir für  $\lambda = -1$  zu der ebenfalls positiv orientierten Basis  $\{v_2, v_3, v_1\}$  über, so erhalten wir die Matrix aus Fall 1 mit  $\theta = \pi$  und  $\lambda = 1$ . Analog ergibt sich für  $\lambda = 1$  mit der Basis  $\{v_3, v_1, v_2\}$  die Matrix aus Fall 1 mit  $\lambda = -1$  und  $\theta = 0$ .

Insgesamt gibt es also immer eine positiv orientierte Orthonormalbasis, so dass  $Q$  eine Matrix hat wie in Fall 1. Insbesondere hat  $Q \in \text{SO}(3)$  immer den Eigenwert  $\lambda = 1$ , der zugehörige Eigenvektor gibt die Drehachse an.

**Definition 7.3** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt diagonalisierbar, wenn  $A$  ähnlich zu einer Diagonalmatrix ist, das heißt es gibt eine invertierbare Matrix  $C \in \mathbb{R}^{n \times n}$  mit

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} =: \Lambda.$$

Eine quadratische Matrix  $A$  ist genau dann diagonalisierbar, wenn sie eine Basis aus Eigenvektoren hat. Genauer gilt:

**Satz 7.2 (zur Diagonalisierbarkeit)** Für  $A \in \mathbb{R}^{n \times n}$ , eine Basis  $v_1, \dots, v_n$  des  $\mathbb{R}^n$  und  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  sind folgende Aussagen äquivalent:

- (1) Es gilt  $Av_j = \lambda_j v_j$  für  $j = 1, \dots, n$ .
- (2) Es gilt  $C^{-1}AC = \Lambda$ , wobei  $C \in \mathbb{R}^{n \times n}$  die Matrix mit den Spalten  $v_j$  ist und  $\Lambda$  die Diagonalmatrix mit den Einträgen  $\lambda_1, \dots, \lambda_n$ .

BEWEIS: Multiplizieren wir die Gleichung  $C^{-1}AC = \Lambda$  von links mit  $C$ , so folgt

$$AC = C\Lambda = C \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Die  $j$ -te Spalte von  $AC$  ist  $Av_j$ , die  $j$ -te Spalte von  $C\Lambda$  ist  $\lambda_j v_j$ . □

Sind  $A, B \in \mathbb{R}^{n \times n}$  ähnlich, so haben sie dasselbe charakteristische Polynom. Denn aus  $B = C^{-1}AC$  folgt (vgl. Übungsaufgabe)

$$\chi_A(\lambda) = \det(C^{-1}(\lambda E_n - A)C) = \det(\lambda E_n - B) = \chi_B(\lambda).$$

Ist daher  $C^{-1}AC = \Lambda$  für  $\Lambda$  diagonal mit Einträgen  $\lambda_1, \dots, \lambda_n$ , so folgt

$$\chi_A(\lambda) = \xi_\Lambda(\lambda) = \det(\lambda E_n - \Lambda) = (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_n).$$

Damit  $A \in \mathbb{R}^{n \times n}$  diagonalisierbar ist, muss also das charakteristische Polynom in reelle Linearfaktoren zerfallen. Wir haben schon Beispiele (etwa  $2 \times 2$ -Drehmatrizen), wo das nicht gegeben ist. Der folgende Satz ist das Hauptresultat zur Diagonalisierbarkeit von reellen Matrizen.

**Satz 7.3 (Hauptachsentransformation)** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch, also  $A^T = A$ . Dann ist  $A$  diagonalisierbar, genauer gibt es eine Orthonormalbasis des  $\mathbb{R}^n$ , die aus Eigenvektoren von  $A$  besteht.

Der Beweis des Satzes ist auch im Hinblick auf Verallgemeinerungen, zum Beispiel bei Eigenwertproblemen für gewöhnliche Differentialgleichungen oder die Schrödingergleichung, von Interesse (und nicht nur das Resultat).

BEWEIS: Sei  $\langle \cdot, \cdot \rangle$  das Standardskalarprodukt. Betrachte die quadratische Form

$$q : \mathbb{R}^n \rightarrow \mathbb{R}, q(x) = \langle Ax, x \rangle = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Die Funktion  $q(x)$  ist stetig, auf der Sphäre  $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$  nimmt sie ihr Minimum an. Es gibt also ein  $v_1 \in \mathbb{R}^n$  mit  $|v_1| = 1$  und

$$q(v_1) = \min\{q(v) : |v| = 1\} =: \lambda_1.$$

Wir zeigen:  $v_1$  ist Eigenvektor von  $A$  zum Eigenwert  $\lambda_1$ . Ist  $v \in \mathbb{R}^n$  mit  $|v| = 1$  und  $\langle v_1, v \rangle = 0$ , so folgt aus der Minimumeigenschaft

$$\begin{aligned} 0 &= \frac{d}{dt} q((\cos t)v_1 + (\sin t)v)|_{t=0} \\ &= \frac{d}{dt} (\cos^2 t \langle Av_1, v_1 \rangle + 2 \sin t \cos t \langle Av_1, v \rangle + \sin^2 t \langle Av, v \rangle)|_{t=0} \\ &= 2 \langle Av_1, v \rangle. \end{aligned}$$

Es gilt also  $Av_1 \perp v$  für alle  $v \perp v_1$ , und damit  $Av_1 = \lambda v_1$  für ein  $\lambda \in \mathbb{R}$ . Bilden wir das Skalarprodukt mit  $v_1$ , so folgt

$$\lambda = \langle Av_1, v_1 \rangle = q(v_1) = \lambda_1.$$

Die Konstruktion der weiteren Eigenvektoren und -werte geschieht induktiv nach demselben Muster. Seien bereits  $v_1, \dots, v_{k-1}$  orthonormale Eigenvektoren zu den Eigenwerten  $\lambda_1, \dots, \lambda_{k-1}$  mit  $k \leq n$ . Es gibt dann  $v_k \in \mathbb{R}^n$  mit  $|v_k| = 1$  und  $v_k \perp v_1, \dots, v_{k-1}$ , so dass

$$q(v_k) = \min\{q(v) : |v| = 1, v \perp v_1, \dots, v_{k-1}\} =: \lambda_k.$$

Aus der Minimumeigenschaft erhalten wir wie oben

$$\langle Av_k, v \rangle = 0 \quad \text{für alle } v \perp v_1, \dots, v_k.$$

Also gilt  $Av_k \in \text{Span}\{v_1, \dots, v_k\}$ , aber andererseits wegen  $A$  symmetrisch

$$\langle Av_k, v_i \rangle = \langle v_k, Av_i \rangle = \lambda_i \langle v_k, v_i \rangle = 0 \quad \text{für } i = 1, \dots, k-1.$$

Es folgt  $Av_k = \lambda v_k$  für ein  $\lambda \in \mathbb{R}$ , und  $\lambda = \langle Av_k, v_k \rangle = q(v_k) = \lambda_k$ . □

Entwickeln wir  $x \in \mathbb{R}^n$  in die Eigenvektor-Basis  $v_1, \dots, v_n$ , also  $x = \sum_{i=1}^n \xi_i v_i$ , so wird die quadratische Form als Summe von Quadraten dargestellt:

$$q(x) = \left\langle A \cdot \sum_{i=1}^n \xi_i v_i, \sum_{i=1}^n \xi_i v_i \right\rangle = \sum_{i,j=1}^n \xi_i \xi_j \underbrace{\langle Av_i, v_j \rangle}_{=\lambda_i \delta_{ij}} = \sum_{i=1}^n \lambda_i \xi_i^2.$$

Für  $n = 2$  wollen wir die Mengen  $M_c = \{\xi \in \mathbb{R}^2 : q(\mathbf{x}) = c\}$  beschreiben. Wir nehmen dazu an, dass  $\{v_1, v_2\}$  die Standardbasis ist und dass  $\lambda_1 \leq \lambda_2$  gilt. Weiter können wir  $\lambda_2 > 0$  voraussetzen, sonst gehen wir zu  $-q(x)$  über. Es ergeben sich drei Fälle:

$\lambda_1 > 0$ : Im Nullpunkt hat  $q(x)$  ein globales, isoliertes Minimum und für  $c > 0$  ist  $M_c$  eine achsensymmetrische Ellipse mit Scheiteln in  $(\pm\sqrt{c/\lambda_1}, 0)$  und  $(0, \pm\sqrt{c/\lambda_2})$ .

$\lambda_1 = 0$ : Auch hier ist im Nullpunkt ein globales Minimum, allerdings ist  $M_0$  die gesamte  $x_1$ -Achse; für  $c > 0$  besteht  $M_c$  aus den parallelen Geraden  $x_2 = \pm\sqrt{c/\lambda_2}$ .

$\lambda_1 < 0$ :  $M_0$  ist Vereinigung der beiden Ursprungsgeraden  $x_2 = \pm\sqrt{-\lambda_1/\lambda_2} x_1$ . Für  $c > 0$  ist  $M_c$  eine nach oben und unten geöffnete Hyperbel mit Scheiteln  $(0, \pm\sqrt{c/\lambda_2})$  und  $M_0$  als Asymptotenlinien. Für  $c < 0$  erhalten wir ebenfalls eine Hyperbel mit Asymptoten

$M_0$ , die aber nach links und rechts geöffnet ist und die Scheitel  $(\pm\sqrt{c/\lambda_1}, 0)$  hat.

Betrachten wir in den drei Fällen die zugehörigen Graphen im  $\mathbb{R}^3$ , so haben wir für  $\lambda_1 > 0$  anschaulich eine Mulde, für  $\lambda_1 = 0$  einen Hohlweg und für  $\lambda_1 < 0$  einen Sattel. Aus der Mulde wird für  $-q(x)$  eine Kuppe. Es ist instruktiv, diese Beschreibung an expliziten Beispielen zu verifizieren und auch ein Bild der zugehörigen Graphen in  $\mathbb{R}^3$  anzufertigen. Man nennt eine quadratische Form nicht degeneriert, wenn die Eigenwerte  $\lambda_i$  alle ungleich Null sind. In diesem Fall bezeichnet man die Anzahl der negativen Eigenwerte als den Index des kritischen Punkts. Die drei Fälle oben sind also der Reihe nach Index Null, degeneriert und Index Eins, die Kuppe ist Index zwei.

Wir bringen noch eine Konsequenz der Hauptachsentransformation. Diese besagt zum Beispiel, dass unter einer invertierbaren linearen Abbildung die Einheitskugel immer auf ein Ellipsoid abgebildet wird.

**Folgerung 7.1 (Polarzerlegung)** Sei  $A \in \mathbb{R}^{n \times n}$  invertierbar. Dann gibt es eine Diagonalmatrix  $\Lambda$  mit Einträgen  $\lambda_i > 0$  und  $Q_1, Q_2 \in \mathbb{O}(n)$ , so dass  $A = Q_1 \Lambda Q_2$ .

BEWEIS:  $A^T A$  ist symmetrisch und hat positive Eigenwerte, denn aus  $A^T A v = \lambda v$  folgt

$$\lambda |v|^2 = \langle A^T A v, v \rangle = |A v|^2 > 0 \quad \text{für } v \neq 0.$$

Nach Satz 7.3 gibt es ein  $C \in \mathbb{O}(n)$  und eine Diagonalmatrix  $\Lambda$  mit  $\lambda_i > 0$ , so dass gilt:

$$A^T A = C \Lambda^2 C^{-1}.$$

$R = C \Lambda C^{-1}$  ist symmetrisch mit  $R^2 = A^T A$ , und  $Q = A R^{-1}$  ist orthogonal wegen

$$Q^T Q = (R^{-1})^T A^T A R^{-1} = R^{-1} R^2 R^{-1} = E_n.$$

Es folgt somit  $A = Q R = Q C \Lambda C^{-1} = Q_1 \Lambda Q_2$  mit  $Q_1 = Q C$ ,  $Q_2 = C^{-1} \in \mathbb{O}(n)$ . □

Wir haben die lineare Algebra über  $\mathbb{R}$  entwickelt, um eine möglichst einfache Darstellung zu erhalten. Damit sind wir aber an dieser Stelle am Ende angekommen. Unser Hauptproblem ist, dass eine lineare Abbildung über  $\mathbb{R}$  keine Eigenwerte haben muss, und dann können wir keine Vereinfachung der Matrixdarstellung erreichen. Dies ist aber möglich, wenn wir die komplexen Zahlen ins Spiel bringen. Unsere Ergebnisse in der linearen Algebra gelten ausnahmslos auch dann, wenn wir Vektorräume mit komplexen Skalaren und komplex-lineare Abbildungen betrachten. Exemplarisch seien folgende Begriffe und Ergebnisse genannt:

- (1) Existenz einer Basis und Dimension
- (2) Darstellung  $\mathbb{C}$ -linearer Abbildungen durch Matrizen in  $\mathbb{C}^{n \times n}$
- (3) Dimensionsformel
- (4) Zeilenrang = Spaltenrang
- (5) Lösungskriterien und -verfahren für lineare Gleichungssysteme
- (6) Determinantenmultiplikationssatz, Invertierbarkeit

(7) Kürzungsregel für Basen, Basiswechsel

(8) Eigenwerte sind Nullstellen des charakteristischen Polynoms

Nur bei Aussagen, die mit dem Skalarprodukt zu tun haben, ist Vorsicht geboten: statt des Skalarprodukts hat man über  $\mathbb{C}$  ein hermitesches Produkt, und im Satz über die Diagonalisierbarkeit muss man statt symmetrischer Matrizen hermitesche Matrizen betrachten, also  $\bar{A}^T = A$ . Aber darauf wollen wir hier nicht weiter eingehen.

Nach dem Fundamentalsatz der Algebra hat jedes Polynom über  $\mathbb{C}$  eine Nullstelle. Für das charakteristische Polynom  $\chi_A(\lambda)$  ergibt sich, dass jede Matrix  $A \in \mathbb{C}^{n \times n}$  mindestens einen Eigenwert  $\lambda \in \mathbb{C}$  hat. Zum Beispiel hat die  $2 \times 2$ -Drehmatrix über  $\mathbb{C}$  die Eigenwerte und Eigenvektoren

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} = (\cos \theta \mp i \sin \theta) \begin{pmatrix} 1 \\ \pm i \end{pmatrix}.$$

Wie jedes Polynom zerfällt  $\chi_A(\lambda)$  über  $\mathbb{C}$  sogar in Linearfaktoren, und dies könnte die Hoffnung wecken, dass eine Matrix  $A \in \mathbb{C}^{n \times n}$  über  $\mathbb{C}$  sogar diagonalisierbar ist. Das ist korrekt, wenn das Polynom  $n$  verschiedene Nullstellen hat, denn dann gehört zu jeder Nullstelle ein Eigenvektor und zusammen bilden diese die gesuchte Basis. Im Fall von mehrfachen Nullstellen ist die Aussage aber im allgemeinen falsch, ein Gegenbeispiel liefern die sogenannten Jordanschen Blockmatrizen:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_0 \end{pmatrix} \in \mathbb{C}^{n \times n}.$$

Die Matrix hat nur den Eigenwert  $\lambda_0$ , denn es gilt (obere Dreiecksmatrix)

$$\chi_{J_n(\lambda_0)}(\lambda) = \det(\lambda E_n - J_n(\lambda_0)) = (\lambda - \lambda_0)^n.$$

Der Ansatz  $J_n(\lambda_0)x = \lambda_0 x$  für  $x \in \mathbb{R}^n$  ergibt aber

$$\lambda_0 x_i + x_{i+1} = \lambda_0 x_i \quad \text{für } i = 1, \dots, n-1,$$

das heißt  $x_{i+1} = 0$  für  $i = 1, \dots, n-1$ . Somit ist  $e_1$  bis auf Skalierung der einzige Eigenvektor, insbesondere ist die Matrix nicht diagonalisierbar. Diese Situation ist typisch. Wir geben ohne Beweis folgendes Resultat an.

**Satz 7.4 (Jordansche Normalform)** *Jede Matrix  $A \in \mathbb{C}^{n \times n}$  ist über  $\mathbb{C}$  ähnlich zu einer Matrix, die aus Jordanblöcken  $J_{n_i}(\lambda_i)$ ,  $i = 1, \dots, r$ , besteht.*

Jede Matrix  $A \in \mathbb{R}^{n \times n}$  kann als Matrix in  $\mathbb{C}^{n \times n}$  aufgefasst werden, und hat daher über  $\mathbb{C}$  eine solche Jordansche Normalform. Dies spielt eine Rolle für das Studium von gewöhnlichen Differentialgleichungen mit konstanten Koeffizienten.

# Kapitel 6

## Differentialrechnung für Funktionen in mehreren Variablen

### 1 Topologische Grundbegriffe

den Begriff der Stetigkeit auf Funktionen, die auf Teilmengen des  $\mathbb{R}^n$  gegeben sind. Vorab führen wir offene und abgeschlossene Mengen ein. Eine Funktion einer Variablen ist üblicherweise auf einem Intervall definiert, entsprechend könnten wir Funktionen von mehreren Variablen nur auf Rechtecken oder Quadern betrachten. Das wäre aber zu einschränkend, eine kurze Diskussion der topologischen Grundbegriffe ist notwendig.

Bereits bekannt ist die Euklidische Norm auf dem  $\mathbb{R}^n$ , also

$$|x| = \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad \text{für } x = (x_1, \dots, x_n) \in \mathbb{R}^n. \quad (1.1)$$

Es gilt  $|x| \geq 0$ ,  $|\lambda x| = |\lambda| |x|$  für  $\lambda \in \mathbb{R}$ , und die Dreiecksungleichung

$$|x + y| \leq |x| + |y| \quad \text{für alle } x, y \in \mathbb{R}^n.$$

Dies hatten wir mit der Ungleichung von Cauchy-Schwarz hergeleitet. Die offene Kugel mit Mittelpunkt  $x \in \mathbb{R}^n$  und Radius  $r > 0$  ist die Menge

$$B_r(x) = \{y \in \mathbb{R}^n : |y - x| < r\}.$$

**Definition 1.1 (offene Menge)** Eine Menge  $D \subset \mathbb{R}^n$  heißt offen, wenn es zu jedem  $x \in D$  ein  $\varepsilon > 0$  gibt mit  $B_\varepsilon(x) \subset D$ .

**Beispiel 1.1** Die Kugel  $B_r(x_0)$  ist offen im Sinn der Definition 1.1. Denn aus der Dreiecksungleichung folgt (Bild)

$$|x - x_0| < r \quad \Rightarrow \quad B_\varepsilon(x) \subset B_r(x_0) \quad \text{für } \varepsilon = r - |x - x_0| > 0.$$

Die leere Menge und  $\mathbb{R}^n$  sind offen, und jede Vereinigung von offenen Mengen ist offen. Bei Durchschnitten muss man aufpassen, im allgemeinen sind nur endliche Durchschnitte wieder offen.

**Definition 1.2 (Konvergenz)** Eine Folge  $x_k \in \mathbb{R}^n$  konvergiert gegen  $x \in \mathbb{R}^n$ , falls gilt:

$$|x_k - x| \rightarrow 0 \quad \text{mit } k \rightarrow \infty.$$

Mit anderen Worten: zu jedem  $\varepsilon > 0$  gibt es ein  $K \in \mathbb{N}$  mit  $x_k \in B_\varepsilon(x)$  für alle  $k \geq K$ .

Die folgende Tatsache wird sehr oft benutzt: eine Folge  $x_k \in \mathbb{R}^n$  konvergiert genau dann gegen  $x \in \mathbb{R}^n$ , wenn sie komponentenweise konvergiert:

$$\lim_{k \rightarrow \infty} x_k = x \quad \Leftrightarrow \quad \lim_{k \rightarrow \infty} (x_k)^i = x^i \quad \text{für } i = 1, \dots, n.$$

Dies ergibt sich aus den Ungleichungen

$$|(x_k)^i - x^i| \leq |x_k - x| \leq \sqrt{n} \max_{1 \leq i \leq n} |(x_k)^i - x^i|.$$

Ich schreibe die Koordinaten hier oben nur um Verwechslung mit dem Folgenindex zu vermeiden.

**Definition 1.3 (abgeschlossene Menge)** Eine Menge  $D \subset \mathbb{R}^n$  heißt abgeschlossen, wenn folgende Implikation gilt:

$$x_k \in D, \quad x_k \rightarrow x \quad \Rightarrow \quad x \in D.$$

Es ist naheliegend, die Begriffe *offen* und *abgeschlossen* als Gegensätze anzusehen. Das ist nicht ganz richtig, vielmehr gilt:

- Es gibt viele Mengen  $D \subset \mathbb{R}^n$ , die weder offen noch abgeschlossen sind, zum Beispiel halboffene Intervalle in  $\mathbb{R}$ .
- Die leere Menge und der ganze  $\mathbb{R}^n$  sind sowohl offen als auch abgeschlossen.
- Eine Menge  $D \subset \mathbb{R}^n$  ist genau dann offen, wenn  $\mathbb{R}^n \setminus D$  abgeschlossen ist.

**Definition 1.4 (Rand und Abschluss einer Menge)** Ein Punkt  $x \in \mathbb{R}^n$  heißt *Randpunkt* von  $D \subset \mathbb{R}^n$ , wenn es in jeder Kugel  $B_\varepsilon(x)$  einen Punkt aus  $D$  und einen Punkt nicht aus  $D$  gibt. Die Menge der Randpunkte wird mit  $\partial D$  bezeichnet, der Abschluss von  $D$  ist die Menge  $\overline{D} = D \cup \partial D$ .

**Beispiel 1.2** Für  $B_r(x_0) = \{x \in \mathbb{R}^n : |x - x_0| < r\}$  behaupten wir

$$\partial B_r(x_0) = \{x \in \mathbb{R}^n : |x - x_0| = r\}.$$

Denn ist  $|x - x_0| = r$ , so liegen in  $B_\varepsilon(x)$  Punkte aus  $B_r(x_0)$  (zum Beispiel  $x_0 + t(x - x_0)$  mit  $t \nearrow 1$ ), sowie Punkte nicht aus  $B_r(x_0)$  (zum Beispiel  $x$  selbst). Somit ist  $x$  ein Randpunkt. Für  $|x - x_0| < r$  hatten wir dagegen ein  $\varepsilon > 0$  gefunden mit  $B_\varepsilon(x) \subset B_r(x_0)$ , das heißt  $x$  ist kein Randpunkt. Entsprechend gilt

$$|x - x_0| > r \quad \Rightarrow \quad B_\varepsilon(x) \subset \mathbb{R}^n \setminus B_r(x_0) \quad \text{für } \varepsilon = |x - x_0| - r > 0.$$

Auch diese Punkte sind also keine Randpunkte, die Behauptung ist gezeigt. Es folgt weiter

$$\overline{B_r(x_0)} = \{x \in \mathbb{R}^n : |x - x_0| \leq r\}.$$

**Definition 1.5 (Grenzwert und Stetigkeit)** Sei  $f : D \rightarrow \mathbb{R}$  wobei  $D \subset \mathbb{R}^n$ .

(1)  $f(x)$  konvergiert für  $x \rightarrow x_0$  gegen  $a \in \mathbb{R}$ , falls es zu jedem  $\varepsilon > 0$  ein  $\delta > 0$  gibt, so dass

$$|f(x) - a| < \varepsilon \quad \text{für alle } x \in D \text{ mit } 0 < |x - x_0| < \delta.$$

Notation:  $\lim_{x \rightarrow x_0} f(x) = a$ .

(2)  $f(x)$  heißt stetig in  $x_0 \in D$ , falls  $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ .

(3)  $f(x)$  heißt stetig auf  $D$ , wenn  $f(x)$  in allen Punkten  $x \in D$  stetig ist.

Diese Definitionen sind analog zum eindimensionalen Fall, insbesondere wiederholen wir die dort gegebenen Hinweise (siehe Kapitel 2.5):

- Für die Existenz des Grenzwerts ist es egal, ob  $f(x)$  im Punkt  $x_0$  definiert ist bzw. welchen Funktionswert die Funktion dort hat.
- Die Definition des Grenzwerts ist nicht sinnvoll, wenn  $x_0$  isolierter Punkt von  $D$  ist, das heißt es gibt ein  $\varepsilon > 0$ , so dass  $x_0$  einziger Punkt aus  $D$  in  $B_\varepsilon(x_0)$ .
- Summe, Produkt und Quotient (außerhalb der Nullstellen des Nenners) von stetigen Funktionen sind wieder stetig. Ebenfalls stetig ist die Verkettung mit einer stetigen Funktion auf  $\mathbb{R}$ .

**Beispiel 1.3** Für jedes  $i = 1, \dots, n$  ist die Koordinatenfunktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = x_i$ , stetig auf  $\mathbb{R}^n$ . Denn es gilt

$$|f(y) - f(x)| = |y_i - x_i| < \varepsilon \quad \text{für } |y - x| < \varepsilon.$$

Damit sind auch Polynomfunktionen stetig, also Funktionen der Form

$$P(x) = \sum_{i_1, \dots, i_n=1}^N a_{i_1 \dots i_n} x_1^{i_1} \cdot \dots \cdot x_n^{i_n}.$$

**Beispiel 1.4** Sei  $a \in \mathbb{R}^n$  ein fester Punkt. Dann ist die Abstandsfunktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = |x - a|$ , stetig. Denn für  $x, y \in \mathbb{R}^n$  gilt mit der Dreiecksungleichung

$$f(y) = |y - a| \leq |y - x| + |x - a| = f(x) + |y - x|,$$

also  $f(y) - f(x) \leq |y - x|$ . Vertauschen wir  $x, y$ , so folgt

$$|f(y) - f(x)| \leq |y - x| < \varepsilon \quad \text{für } |y - x| < \varepsilon.$$

**Definition 1.6 (Beschränktheit)** Eine Menge  $D \subset \mathbb{R}^n$  heißt beschränkt, wenn es ein  $K \geq 0$  gibt mit  $|x| \leq K$  für alle  $x \in D$ .

**Satz 1.1 (Existenz von Extremstellen)** Sei  $D \subset \mathbb{R}^n$  abgeschlossen und beschränkt, und  $f : D \rightarrow \mathbb{R}$  sei stetig. Dann gibt es  $x_0, x_1 \in D$  mit

$$f(x_0) \leq f(x) \leq f(x_1) \quad \text{für alle } x \in D.$$

Insbesondere ist die Funktion  $f$  beschränkt.

BEWEIS: (Skizze) Um die Minimalstelle zu finden, wählen wir Punkte  $x_k \in D$  mit

$$\lim_{k \rightarrow \infty} f(x_k) = \inf_{x \in D} f(x).$$

Eine solche Folge nennt man eine Minimalfolge. Angenommen die  $x_k$  konvergieren gegen ein  $x_0 \in \mathbb{R}^n$ . Da  $D$  abgeschlossen ist, folgt dann  $x_0 \in D$  und weiter wegen der Stetigkeit von  $f(x)$

$$f(x_0) = \lim_{k \rightarrow \infty} f(x_k) = \inf_{x \in D} f(x).$$

Somit ist die gesuchte Minimalstelle gefunden. Im allgemeinen muss die Folge  $x_k$  nicht konvergieren, zum Beispiel könnte  $f(x)$  mehrere Minimalstellen haben. Aber es reicht, wenn wir eine konvergente Teilfolge  $x_{k_1}, x_{k_2}, \dots$  mit  $k_1 < k_2 < \dots$  finden, denn diese ist auch eine Minimalfolge, und das Argument von oben ist gültig.

Eine beschränkte Folge im  $\mathbb{R}^n$  hat nun immer eine solche konvergente Teilfolge. Im Fall  $n = 1$  sei dazu  $I_0$  ein Intervall, das alle Folgenglieder enthält. Dann wählen wir eine Teilfolge, die ganz in einer der Hälften von  $I_0$  liegt. Sei  $I_1$  diese Hälfte, wähle dann eine Teilfolge, die ganz in einer Hälfte von  $I_1$  liegt. Wir erhalten sukzessive Teilfolgen, so dass die  $m$ -te Teilfolge ganz in  $I_m$  liegt, wobei  $I_0 \supset I_1 \supset \dots$  und  $|I_m| = 2^{-m}|I_0| \rightarrow 0$ . Allerdings muss keine dieser Teilfolgen konvergieren. Der Trick ist folgendes Diagonalargument von Bolzano und Weierstrass: betrachte die Folge, deren  $m$ -tes Glied das  $m$ -te Glied der  $m$ -ten Teilfolge ist. Dies ist die gewünschte konvergente Teilfolge, denn für jedes  $m$  liegt die Folge ab Nummer  $m$  ganz in  $I_m$ . In Dimension  $n \geq 1$  wendet man dieses Argument auf jede der Koordinaten an.  $\square$

**Beispiel 1.5** Sei  $K \subset \mathbb{R}^n$  kompakt. Dann gibt es zu jedem  $a \in \mathbb{R}^n$  einen nächsten Punkt  $x_0 \in K$ , das heißt es gilt

$$|x_0 - a| \leq |x - a| \quad \text{für alle } x \in K.$$

Der Punkt  $x_0$  ist nicht notwendig eindeutig, betrachte etwa  $K = \{1, -1\} \subset \mathbb{R}$  und  $a = 0$ .

**Beispiel 1.6** Betrachte die quadratische Form

$$q : \mathbb{R}^n \rightarrow \mathbb{R}, \quad q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Wie oben festgestellt ist  $q(x)$  als Polynomfunktion stetig auf  $\mathbb{R}^n$ . Die Standardsphäre  $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$  ist beschränkt (wähle  $K = 1$ ) und abgeschlossen: die Funktion  $f(x) = |x|$  ist stetig, also gilt

$$x_k \in \mathbb{S}^{n-1}, \quad x_k \rightarrow x \quad \Rightarrow \quad |x| = \lim_{k \rightarrow \infty} |x_k| = 1.$$

Somit ist  $x \in \mathbb{S}^{n-1}$ . Aus Satz 1.1 folgt, dass  $q(x)$  eine Minimalstelle auf  $\mathbb{S}^{n-1}$  hat. Dies hatten wir bei der Hauptachsentransformation, siehe Satz 7.3, benutzt.

## 2 Die partiellen Ableitungen

Wir kommen nun zur Differentiation von Funktionen im  $\mathbb{R}^n$ . Um für diese Ableitungen zu definieren, ist die einfachste und vielfach beste Idee, alle Variablen bis auf  $x_j$  als konstant aufzufassen und die resultierende Funktion der einen Variablen  $x_j$  wie üblich zu differenzieren. Auf diese Weise ergeben sich für  $j = 1, \dots, n$  die  $n$  partiellen Ableitungen der Funktion.

**Definition 2.1** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $f : \Omega \rightarrow \mathbb{R}$ . Die partielle Ableitung von  $f$  nach  $x_j$  an der Stelle  $x \in \Omega$  ist (falls existent)

$$\begin{aligned} \frac{\partial f}{\partial x_j}(x) &= \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_j + t, \dots, x_n) - f(x_1, \dots, x_j, \dots, x_n)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x + te_j) - f(x)}{t}. \end{aligned}$$

Dies ist genau die Ableitung der Funktion  $x_j \mapsto f(x_1, \dots, x_j, \dots, x_n)$ . Ebenfalls übliche Bezeichnungen sind  $\partial_j f(x)$  und  $f_{x_j}(x)$ .

Das Symbol  $\partial$  (statt  $d$ ) weist darauf hin, dass es sich um eine partielle Ableitung nach einer der Variablen handelt. Der folgende Satz formuliert Aussagen der eindimensionalen Analysis im jetzigen Kontext.

**Satz 2.1 (Ableitungsregeln)** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $x \in \Omega$ . Die Existenz der partiellen Ableitungen  $\partial_j f(x)$  und  $\partial_j g(x)$  sei vorausgesetzt. Dann gelten folgende Aussagen:

(a) Linearität: für  $f, g : \Omega \rightarrow \mathbb{R}$  und  $\alpha, \beta \in \mathbb{R}$  gilt

$$\partial_j(\alpha f + \beta g)(x) = \alpha \partial_j f(x) + \beta \partial_j g(x).$$

(b) Produktregel: für  $f, g : \Omega \rightarrow \mathbb{R}$  gilt

$$\partial_j(fg)(x) = (\partial_j f)(x)g(x) + f(x)(\partial_j g)(x).$$

(d) Quotientenregel: für  $f, g : \Omega \rightarrow \mathbb{R}$  mit  $g(x) \neq 0$  gilt

$$\partial_j \left( \frac{f}{g} \right) (x) = \frac{(\partial_j f)(x)g(x) - f(x)(\partial_j g)(x)}{g(x)^2}.$$

(e) Kettenregel: ist  $f : \Omega \rightarrow I \subset \mathbb{R}$  und  $\varphi : I \rightarrow \mathbb{R}$  differenzierbar in  $f(x)$ , so folgt

$$\partial_j(\varphi \circ f)(x) = \varphi'(f(x))\partial_j f(x).$$

**Beispiel 2.1** Wir betrachten die Euklidische Abstandsfunktion vom Nullpunkt

$$r : \mathbb{R}^n \rightarrow \mathbb{R}, r(x) = |x| = \sqrt{x_1^2 + \dots + x_n^2}.$$

In  $x \neq 0$  existieren die partiellen Ableitungen, und zwar gilt mit der Kettenregel

$$\partial_j r(x) = \frac{2x_j}{2\sqrt{x_1^2 + \dots + x_n^2}} = \frac{x_j}{r} \quad \text{für } r = |x|.$$

Im Nullpunkt sind die partiellen Ableitungen dagegen nicht definiert, denn  $r(0 + te_i) = |t|$  ist in  $t = 0$  nicht differenzierbar. Die Funktion  $\partial_j r$  ist in  $x \neq 0$  ihrerseits partiell differenzierbar, und wir erhalten die zweiten partiellen Ableitungen

$$\partial_i(\partial_j r)(x) = \frac{(\partial_i x_j)r - x_j \partial_i r}{r^2} = \frac{1}{r} \left( \delta_{ij} - \frac{x_i x_j}{r^2} \right).$$

Ist  $\varphi : (0, \infty) \rightarrow \mathbb{R}$  zweimal differenzierbar, so berechnen wir weiter für  $f = \varphi \circ r : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\begin{aligned} \partial_j f(x) &= \varphi'(r) \partial_j r = \varphi'(r) \frac{x_j}{r}, \\ \partial_i(\partial_j f)(x) &= \varphi''(r) \partial_i r \partial_j r + \varphi'(r) \partial_i(\partial_j r) = \varphi''(r) \frac{x_i x_j}{r^2} + \frac{\varphi'(r)}{r} \left( \delta_{ij} - \frac{x_i x_j}{r^2} \right). \end{aligned}$$

Der Operator  $\Delta = \sum_{i=1}^n \partial_i^2$  heißt Laplaceoperator, die Lösungen der Gleichung  $\Delta f = 0$  heißen harmonische Funktionen. Wir können jetzt die rotationssymmetrischen harmonischen Funktionen ausrechnen, und zwar erhalten wir

$$0 \stackrel{!}{=} \Delta f(x) = \varphi''(r) + \frac{n-1}{r} \varphi'(r) = r^{1-n} (r^{n-1} \varphi'(r))'.$$

Diese Gleichung hat die Lösungen, mit Integrationskonstanten  $a, b \in \mathbb{R}$ ,

$$\varphi(r) = \begin{cases} a \frac{r^{2-n}}{2-n} + b & \text{für } n \geq 3 \\ a \log r + b & \text{für } n = 2. \end{cases}$$

Für  $n = 3$  ist  $f$  das Newtonsche Gravitationspotential bzw. Coulombpotential.

Im Beispiel traten zweite partielle Ableitungen auf. Ist die Ableitungsfunktion  $\partial_j f : \Omega \rightarrow \mathbb{R}$  definiert und wiederum nach  $x_i$  partiell differenzierbar, so setzen wir

$$\partial_i \partial_j f(x) := \partial_i(\partial_j f)(x) \quad (\text{alternative Notation } \frac{\partial^2 f}{\partial x_i \partial x_j}(x)). \quad (2.2)$$

Entsprechend werden partielle Ableitungen beliebiger Ordnung gebildet. Die folgenden Klassen von Funktionen spielen in der Analysis eine wichtige Rolle. Wir werden sehen, dass die Eigenschaft der Stetigkeit der partiellen Ableitungen in vielen Anwendungen wesentlich ist.

**Definition 2.2 ( $C^k$ -Räume)** Sei  $\Omega \subset \mathbb{R}^n$  offen und  $k \in \mathbb{N}_0 \cup \{\infty\}$ . Wir bezeichnen mit  $C^k(\Omega)$  die Menge aller  $k$ -mal stetig differenzierbaren Funktionen auf  $\Omega$ , das heißt alle partiellen Ableitungen  $\partial_{i_1} \dots \partial_{i_j} f$  der Ordnung  $0 \leq j \leq k$  (bzw.  $j < \infty$  im Fall  $k = \infty$ ) sind definiert und stetig auf  $\Omega$ .

Wir wollen nun zeigen, dass die Operatoren  $\partial_i$  und  $\partial_j$  auf  $C^2$ -Funktionen vertauschen. Beispiele zeigen, dass dies nicht allein aus der Existenz der Ableitungen  $\partial_i \partial_j f$  und  $\partial_j \partial_i f$  folgt, die Stetigkeit dieser Ableitungen ist wesentlich.

**Satz 2.2 (Symmetrie der zweiten Ableitung)** Für  $f \in C^2(\Omega)$  gilt

$$\partial_i \partial_j f = \partial_j \partial_i f \quad \text{auf } \Omega \quad \text{für alle } 1 \leq i, j \leq n.$$

BEWEIS: Es reicht, eine Funktion von zwei Variablen zu betrachten. Bilde für  $f = f(x, y)$  nacheinander die Differenzenquotienten nach  $y$  und  $x$ :

$$\begin{aligned}\delta_2^h f(x, y) &= \frac{f(x, y+h) - f(x, y)}{h}, \\ \delta_1^h(\delta_2^h f)(x, y) &= \frac{\delta_2^h f(x+h, y) - \delta_2^h f(x, y)}{h} \\ &= \frac{f(x+h, y+h) - f(x+h, y) - f(x, y+h) + f(x, y)}{h^2}.\end{aligned}$$

Das Ergebnis ist symmetrisch bzgl.  $x$  und  $y$ , wir haben also

$$\delta_1^h(\delta_2^h f)(x, y) = \delta_2^h(\delta_1^h f)(x, y).$$

Mit dem Mittelwertsatz können wir einen Differenzenquotienten durch die Ableitung an einer Zwischenstelle ersetzen. Wir machen das links erst bezüglich  $x$  und dann bzgl.  $y$ , und erhalten  $\xi, \eta$  mit  $|\xi - x|, |\eta - y| \leq |h|$ , so dass gilt:

$$\begin{aligned}\delta_1^h(\delta_2^h f)(x, y) &= \partial_1(\delta_2^h f)(\xi, y) \\ &= \frac{1}{h}(\partial_1 f(\xi, y+h) - \partial_1 f(\xi, y)) \\ &= \partial_2(\partial_1 f)(\xi, \eta).\end{aligned}$$

Analog können wir rechts umformen, es folgt mit anderen Zwischenstellen  $\xi', \eta'$

$$\partial_2(\partial_1 f)(\xi, \eta) = \partial_1(\partial_2 f)(\xi', \eta').$$

Mit  $h \rightarrow 0$  folgt wegen der Stetigkeit  $\partial_2(\partial_1 f)(x, y) = \partial_1(\partial_2 f)(x, y)$ . □

Durch Induktion ergibt sich allgemeiner die

**Folgerung 2.1** Für  $f \in C^k(\Omega)$  vertauschen die partiellen Ableitungen bis zur Ordnung  $k$ .

Eine geringfügige Verallgemeinerung der partiellen Ableitung ist die Richtungsableitung. Sie hat den Vorteil, dass sie nicht von der Wahl des Koordinatensystems explizit abhängt.

**Definition 2.3 (Richtungsableitung)**  $\Omega \subset \mathbb{R}^n$  offen und  $f : \Omega \rightarrow \mathbb{R}$ . Für  $x \in \Omega$  und  $v \in \mathbb{R}^n$  heißt der Grenzwert (falls existent)

$$\partial_v f(x) = \lim_{t \rightarrow 0} \frac{f(x+tv) - f(x)}{t}$$

Richtungsableitung von  $f$  an der Stelle  $x$  in Richtung  $v$ . Dies ist die gewöhnliche Ableitung der Funktion  $t \mapsto f(x+tv)$  an der Stelle  $t=0$ .

**Beispiel 2.2** Die Richtungsableitung von  $r(x) = |x|$  in  $x \in \mathbb{R}^n \setminus \{0\}$  in Richtung  $v \in \mathbb{R}^n$  ist

$$\partial_v r(x) = \frac{d}{dt} \sqrt{|x|^2 + 2t\langle x, v \rangle + |v|^2} \Big|_{t=0} = \left\langle \frac{x}{|x|}, v \right\rangle.$$

Aus Gründen der Übersichtlichkeit haben wir die partiellen Ableitungen erst nur für skalare Funktionen, also Funktionen mit Werten in  $\mathbb{R}$ , betrachtet. Die Verallgemeinerung auf vektorwertige Funktionen ist aber einfach, die Ableitung wird komponentenweise gebildet. Wir betrachten das mal im Spezialfall der parametrisierten Kurven.

Eine stetige Abbildung  $\gamma : I \rightarrow \mathbb{R}^m$ ,  $I$  Intervall, nennt man eine parametrisierte Kurve. Es ist hilfreich, sich  $\gamma(t)$  als Ortsvektor eines Punkts vorzustellen, der sich abhängig von der Zeit  $t \in I$  bewegt. Der Geschwindigkeitsvektor (oder die Ableitung) von  $\gamma(t)$  ist (falls existent)

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_m(t) \end{pmatrix}.$$

Die Bogenlänge von  $\gamma \in C^1(I, \mathbb{R}^m)$  ist definiert als das Integral

$$L(\gamma) = \int_a^b |\gamma'(t)| dt. \quad (2.3)$$

Zur Motivation dieser Definition betrachte eine Unterteilung  $a = t_0 < \dots < t_N = b$ . Dann sollte die Länge von  $\gamma$  in etwa die Länge des zugehörigen Polygonzugs sein (Bild):

$$L(\gamma) \approx \sum_{k=1}^N |\gamma(t_k) - \gamma(t_{k-1})|.$$

Da  $\gamma'(t)$  auf jedem Teilintervall nahezu konstant ist, gilt (*Weg = Geschwindigkeit  $\times$  Zeit*)

$$|\gamma(t_k) - \gamma(t_{k-1})| \approx |\gamma'(t_{k-1})|(t_k - t_{k-1}) \approx \int_{t_{k-1}}^{t_k} |\gamma'(t)| dt.$$

Summieren wir für  $k = 1, \dots, N$ , so ergibt sich (näherungsweise) die Formel (2.3).

**Beispiel 2.3** Die Schraubenlinie ist die parametrisierte Kurve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^3, \gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ ht \end{pmatrix}.$$

Die Bahn der Kurve liegt auf dem Zylinder  $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = r^2\}$ . Tangentenvektor sowie Beschleunigungsvektor lauten

$$\gamma'(t) = \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix} \quad \text{und} \quad \gamma''(t) = \begin{pmatrix} -r \cos t \\ -r \sin t \\ 0 \end{pmatrix}.$$

Bei einem Umlauf wächst die dritte Komponente von  $\gamma(t)$  um den Wert  $2\pi h$ , die sogenannte Ganghöhe der Schraubenlinie. Die Bogenlänge eines Umlaufs ist

$$L(\gamma|_{[0, 2\pi]}) = \int_0^{2\pi} |\gamma'(t)| dt = \int_0^{2\pi} \sqrt{r^2 + h^2} dt = 2\pi \sqrt{r^2 + h^2}.$$

Bei Abrollen des Zylinders wird aus der Schraubenlinie eine Gerade, das betrachtete Stück entspricht der Hypotenuse im rechtwinkligen Dreieck mit Katheten  $2\pi r$  und  $2\pi h$ .

Für eine vektorwertige Funktion von mehreren Variablen werden die partiellen Ableitungen ebenfalls komponentenweise gebildet. Mit  $f : \Omega \rightarrow \mathbb{R}^m$ ,  $\Omega \subset \mathbb{R}^n$ , gilt also

$$\partial_j f(x) = \begin{pmatrix} \partial_j f_1(x) \\ \vdots \\ \partial_j f_m(x) \end{pmatrix} \quad \text{für } j = 1, \dots, n.$$

Im Fall  $n = 2, m = 3$  kann man sich vorstellen, dass  $f$  eine Fläche im  $\mathbb{R}^3$  parametrisiert.

**Beispiel 2.4** Sei  $\gamma : I \rightarrow \mathbb{R}^2$ ,  $\gamma(s) = (x(s), r(s))$ , eine Kurve mit  $r(s) > 0$  für alle  $s \in I$ . Durch Rotation um die  $x$ -Achse erhalten wir dann die Rotationsfläche

$$f : I \times \mathbb{R} \rightarrow \mathbb{R}^3, f(s, \theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} x(s) \\ r(s) \\ 0 \end{pmatrix} = \begin{pmatrix} x(s) \\ r(s) \cos \theta \\ r(s) \sin \theta \end{pmatrix}.$$

Die partiellen Ableitungen von  $f(s, \theta)$  lauten

$$\frac{\partial f}{\partial s}(s, \theta) = \begin{pmatrix} x'(s) \\ r'(s) \cos \theta \\ r'(s) \sin \theta \end{pmatrix} \quad \text{und} \quad \frac{\partial f}{\partial \theta}(s, \theta) = \begin{pmatrix} 0 \\ -r(s) \sin \theta \\ r(s) \cos \theta \end{pmatrix}.$$

Die Kurve  $s \mapsto f(s, 0) = (\gamma(s), 0)$  hat den Geschwindigkeitsvektor  $\frac{\partial f}{\partial s}(s, 0) = (\gamma'(s), 0)$ . Die Kurve  $s \mapsto f(s, \theta)$  ergibt sich aus  $f(s, 0)$  durch Drehung um den Winkel  $\theta$ , dementsprechend ist  $\frac{\partial f}{\partial s}(s, \theta)$  der gedrehte Geschwindigkeitsvektor. Die Kurve  $\theta \mapsto f(s, \theta)$  beschreibt einen Kreis mit Radius  $r(s)$  in der Ebene  $x = x(s)$ . Der Vektor  $\frac{\partial f}{\partial \theta}(s, \theta)$  ist der Geschwindigkeitsvektor dieses Kreises.

### 3 Die Jacobimatrix

**Definition 3.1 (Jacobimatrix)** Sei  $f \in C^1(\Omega, \mathbb{R}^m)$  wobei  $\Omega \subset \mathbb{R}^n$ . Die Jacobimatrix von  $f$  im Punkt  $x \in \Omega$  ist

$$Df(x) = \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix} \in \mathbb{R}^{m \times n}. \quad (3.4)$$

In der  $i$ -ten Zeile stehen die Ableitungen der  $i$ -ten Komponente von  $f$ , in der  $j$ -ten Spalte stehen die Ableitungen von  $f$  nach der Variablen  $x_j$ .

**Beispiel 3.1** Zweidimensionale Polarkoordinaten sind definiert durch

$$\phi : (0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{0\}, \phi(r, \theta) = (r \cos \theta, r \sin \theta).$$

Die Jacobimatrix dieser Abbildung lautet

$$D\phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

**Beispiel 3.2** Betrachte für  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$  die affine Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $f(x) = Ax + b$ . Dann gilt  $f_i(x) = \sum_{j=1}^n a_{ij}x_j + b_i$  und somit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij} \quad \text{bzw.} \quad Df(x) = A \quad \text{für alle } x \in \mathbb{R}^n.$$

**Satz 3.1 (von der linearen Approximation)** Ist  $f \in C^1(\Omega, \mathbb{R}^m)$ , so gilt im Punkt  $x_0$  mit  $A = Df(x_0)$  die Approximationseigenschaft

$$\lim_{x \rightarrow x_0} \frac{f(x) - (f(x_0) + A(x - x_0))}{|x - x_0|} = 0. \quad (3.5)$$

BEWEIS: Die Konvergenz ist gleichbedeutend mit Konvergenz der Komponenten, daher können wir  $m = 1$  annehmen. Sei zuerst  $x_0 = 0$  und  $f(0) = 0$ ,  $Df(0) = 0$ . Dann gilt mit dem Mittelwertsatz

$$\begin{aligned} f(x) &= \sum_{j=1}^n (f(x_1, \dots, x_{j-1}, x_j, 0, \dots, 0) - f(x_1, \dots, x_{j-1}, 0, \dots, 0)) \\ &= \sum_{j=1}^n (\partial_j f)(x_1, \dots, x_{j-1}, \xi_j, 0, \dots, 0) x_j. \end{aligned}$$

Also gilt die Abschätzung

$$\frac{|f(x)|}{|x|} \leq \sum_{j=1}^n |(\partial_j f)(x_1, \dots, x_{j-1}, \xi_j, 0, \dots, 0)| \rightarrow 0 \quad \text{mit } x \rightarrow 0.$$

Sind  $x_0$ ,  $f(x_0)$  und  $A = Df(x_0)$  beliebig, so betrachte die  $C^1$ -Funktion

$$g(y) = f(x_0 + y) - (f(x_0) + Ay) = f(x_0 + y) - f(x_0) - \sum_{k=1}^n \partial_k f(x_0) y_k.$$

Es gilt dann  $g(0) = 0$  sowie  $Dg(0) = 0$ , und somit

$$\frac{f(x) - (f(x_0) + A(x - x_0))}{|x - x_0|} = \frac{g(x - x_0)}{|x - x_0|} \rightarrow 0 \quad \text{mit } x \rightarrow x_0.$$

□

Der Satz besagt, dass das Verhalten von  $f(x)$  für  $x \rightarrow x_0$  in erster Ordnung durch die affine Funktion  $f(x_0) + A(x - x_0)$  bestimmt wird. Um zu sagen, dass eine Funktion  $f(x)$  mit  $x \rightarrow x_0$  schneller verschwindet bzw. schwächer wächst als eine Funktion  $g(x)$ , wird oft das Landausche  $o$ -Symbol benutzt:

$$f(x) = o(g(x)) \quad \text{für } x \rightarrow x_0 \quad \Leftrightarrow \quad \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} \rightarrow 0.$$

Die Aussage des Satzes lautet damit

$$f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|) \quad \text{für } x \rightarrow x_0.$$

Das ist ziemlich griffig, aber man muss die Bedeutung des  $o$ -Symbols im Kopf behalten. Eine erste Konsequenz des Satzes ist, dass wir die Richtungsableitung mit den partiellen Ableitungen ausrechnen können.

**Folgerung 3.1** Die Richtungsableitung von  $f \in C^1(\Omega)$  in Richtung  $v \in \mathbb{R}^n$  ist

$$\partial_v f(x) = Df(x)v = \sum_{j=1}^n \partial_j f(x_0)v_j.$$

BEWEIS: Im Fall  $v = 0$  stimmt die Aussage. Für  $v \neq 0$  rechnen wir

$$\frac{f(x+tv) - f(x)}{t} - Df(x)v = \frac{f(x+tv) - (f(x) - Df(x)(tv))}{t} \rightarrow 0 \quad \text{mit } t \rightarrow 0.$$

□

**Definition 3.2 (Gradient)** Der Gradient einer Funktion  $f \in C^1(\Omega)$  im Punkt  $x \in \Omega$  ist

$$\text{grad } f(x) = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \in \mathbb{R}^n.$$

Ist  $\text{grad } f(x) = 0$ , so heißt  $x$  kritischer Punkt von  $f$ . Ist  $x$  nicht kritisch, so ist die Richtung von  $\text{grad } f(x)$  genau die, in der  $f(x)$  am stärksten ansteigt. Denn für  $v \in \mathbb{R}^n$  mit  $|v| = 1$  gilt nach Folgerung 3.1 und Cauchy-Schwarz

$$\partial_v f(x) = \langle \text{grad } f(x), v \rangle \leq |\text{grad } f(x)|. \quad (3.6)$$

Gleichheit tritt genau dann ein, wenn der Einheitsvektor  $v$  in Richtung von  $\text{grad } f(x)$  weist.

Beachten Sie, dass der Gradient nur für skalare Funktionen definiert wird. Für diese ist nach unserer Definition die Jacobimatrix  $Df(x) \in \mathbb{R}^{1 \times n}$ , also der Zeilenvektor mit Einträgen  $\partial_j f(x)$ . Der Gradient ist de facto dasselbe, nur haben wir ihn als Spaltenvektor geschrieben, das heißt als Element von  $\mathbb{R}^{n \times 1}$ . In den Standardkoordinaten ist der Unterschied rein formal, bei Wechseln des Koordinatensystems spielt er aber eine Rolle, weil sich die Jacobimatrix anders transformiert als der Gradient. Das werden wir sehen.

**Beispiel 3.3** Der Gradient der Funktion  $f(x) = \varphi(r)$  mit  $r(x) = |x|$  ist nach Beispiel 2.1

$$\text{grad } f(x) = \varphi'(r) \frac{x}{r} \quad \text{für } x \neq 0.$$

**Beispiel 3.4** Sei  $q(x) = \langle Ax, x \rangle = \sum_{i,k=1}^n a_{ik}x_i x_k$  quadratische Form der symmetrischen Matrix  $A \in \mathbb{R}^{n \times n}$ . Dann gilt

$$\partial_j q(x) = \sum_{i,k=1}^n a_{ik}(\delta_{ij}x_k + x_i\delta_{jk}) = \sum_{k=1}^n a_{jk}x_k + \sum_{i=1}^n a_{ij}x_i = 2 \sum_{k=1}^n a_{jk}x_k.$$

Das bedeutet  $\text{grad } q(x) = 2Ax$ .

Wir kommen jetzt zur mehrdimensionalen Kettenregel. Wir betrachten erst den wichtigen Spezialfall, wenn eine Funktion längs einer Kurve ausgewertet wird.

**Satz 3.2 (Kettenregel längs Kurven)** Sei  $f \in C^1(\Omega)$  und  $x = x(t)$ ,  $t \in (a, b)$ , sei eine  $C^1$ -Kurve in  $\Omega$  mit  $x(t_0) = x_0$ . Dann gilt

$$\frac{d}{dt}f(x(t))|_{t=t_0} = Df(x_0)x'(t_0) = \sum_{j=1}^n (\partial_j f)(x_0)x'_j(t_0).$$

BEWEIS: Nach Satz 3.1 gilt

$$\varepsilon(x) := \frac{f(x) - (f(x_0) + Df(x_0)(x - x_0))}{|x - x_0|} \rightarrow 0 \quad \text{mit } x \rightarrow x_0.$$

Wir berechnen nun

$$\begin{aligned} \frac{f(x(t)) - f(x_0)}{t - t_0} - Df(x_0)x'(t_0) &= \frac{f(x(t)) - (f(x_0) + Df(x_0)(x(t) - x(t_0)))}{t - t_0} \\ &\quad + Df(x_0)\left(\frac{x(t) - x(t_0)}{t - t_0} - x'(t_0)\right) \\ &= \frac{\varepsilon(x(t))|x(t) - x_0|}{t - t_0} + Df(x_0)\left(\frac{x(t) - x(t_0)}{t - t_0} - x'(t_0)\right). \end{aligned}$$

Da  $\lim_{t \rightarrow t_0} \frac{x(t) - x(t_0)}{t - t_0} = x'(t_0)$ , geht die rechte Seite gegen Null und der Satz ist bewiesen.  $\square$   
Ist  $f(x(t))$  konstant, so folgt

$$0 = \frac{d}{dt}f(x(t)) = Df(x(t))x'(t) = \langle \text{grad } f(x(t)), x'(t) \rangle.$$

Die Mengen  $\{x \in \Omega : f(x) = \text{const.}\}$  sind die Niveaulinien ( $n = 2$ ) oder Niveauflächen ( $n = 3$ ) der Funktion  $f$ . Die Aussage bedeutet, der Gradient von  $f$  steht senkrecht auf diese Linien oder Flächen.

Wir kommen nun zur allgemeinen Fassung der Kettenregel. Sind  $f(x) = Ax$  und  $g(y) = By$  lineare Abbildungen mit  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{p \times m}$ , so gilt  $g(f(x)) = BAx$ , und diese Funktion hat die Jacobimatrix  $BA \in \mathbb{R}^{p \times n}$ , siehe Beispiel 3.2. Daraus ergibt sich, dass allgemein die Jacobimatrix einer Verkettung das Produkt der Jacobimatrizen an den jeweiligen Stellen sein muss.

**Satz 3.3 (Kettenregel)** Seien  $f \in C^1(U, \mathbb{R}^m)$  und  $g \in C^1(V, \mathbb{R}^p)$  mit  $U \subset \mathbb{R}^n$ ,  $V \subset \mathbb{R}^m$ . Dann ist  $g \circ f \in C^1(U, \mathbb{R}^p)$  und es gilt die Kettenregel

$$D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0).$$

Mit andern Worten gilt mit  $y_0 = f(x_0)$

$$\frac{\partial(g \circ f)_i}{\partial x_k}(x_0) = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j}(y_0) \frac{\partial f_j}{\partial x_k}(x_0) \quad \text{für } 1 \leq i \leq p, 1 \leq k \leq n.$$

BEWEIS: Mit  $y(t) = f(x_0 + te_k)$  gilt  $y(0) = f(x_0) = y_0$ , und es folgt aus der Kettenregel längs Kurven, Satz 3.2,

$$\frac{\partial(g \circ f)_i}{\partial x_k}(x_0) = \frac{d}{dt}g_i(y(t))|_{t=0} = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j}(y_0)y'_j(0) = \sum_{j=1}^m \frac{\partial g_i}{\partial y_j}(y_0) \frac{\partial f_j}{\partial x_k}(x_0).$$

□

Oft wird die Kettenregel so beschrieben: um die Ableitung der  $i$ -ten Komponente von  $g \circ f$  nach  $x_k$  zu berechnen, wird  $g$  nach allen Variablen  $y_j = f_j(x)$  abgeleitet, und diese werden nach  $x_k$  nachdifferenziert. Die Interpretation der Ableitung als Matrix bzw. lineare Abbildung ermöglicht oft ein gutes geometrisches Verständnis. Allerdings ist beim Umgang mit Zeilen- und Spaltenvektoren Vorsicht geboten, zum Beispiel kann bei der Produktregel Verwirrung entstehen, wenn eine der beteiligten Funktionen vektorwertig ist. Erst recht wird die Notation kompliziert, wenn zweite oder höhere Ableitungen zu bilden sind. Im Zweifelsfall sollte man auf die partiellen Ableitungen zurückgreifen.

## 4 Erste Anwendungen der Ableitung

Um aus Eigenschaften der Ableitung auf das Verhalten einer Funktion zu schließen, haben wir bei Funktionen einer Variablen zwei Hilfsmittel, den Mittelwertsatz und den Hauptsatz der Differential- und Integralrechnung. Diese können wir auch für Funktionen mehrerer Variabler nutzen, indem wir die Funktionen längs Kurven  $x = x(t)$  auswerten.

**Lemma 4.1** Sei  $f \in C^1(\Omega)$ , wobei  $\Omega \subset \mathbb{R}^n$ , und  $x_0, x_1 \in \Omega$ . Ist  $x : [0, 1] \rightarrow \Omega$  eine  $C^1$ -Verbindungskurve von  $x_0$  nach  $x_1$ , das heißt  $x(0) = x_0$  und  $x(1) = x_1$ , so gilt

$$f(x_1) - f(x_0) = \int_a^b Df(x(t))x'(t) dt. \quad (4.1)$$

BEWEIS: Die Kettenregel für Kurven, Satz 3.2, besagt

$$\frac{d}{dt}f(x(t)) = Df(x(t))x'(t).$$

Durch Integration folgt die Behauptung. □

**Definition 4.1** Eine offene Menge  $\Omega \subset \mathbb{R}^n$  heißt *wegweise zusammenhängend*, wenn es zu je zwei Punkten  $x_0, x_1 \in \Omega$  eine  $C^1$ -Verbindungskurve von  $x_0$  nach  $x_1$  in  $\Omega$  gibt.

**Satz 4.1 (Konstanzsatz)** Sei  $\Omega \subset \mathbb{R}^n$  offen und *wegweise zusammenhängend*. Dann gilt:

$$Df(x) = 0 \quad \text{für alle } x \in \Omega \quad \Rightarrow \quad f \text{ ist konstant.}$$

BEWEIS: Wähle zu  $x_0, x_1$  eine  $C^1$ -Verbindungskurve  $x : [0, 1] \rightarrow \Omega$ . Aus Lemma 4.1 folgt

$$f(x_1) - f(x_0) = \int_a^b Df(x(t))x'(t) dt = 0.$$

Da  $x_0, x_1$  beliebig, ist die Funktion konstant. □

Wählen wir als Verbindungskurven Strecken, so können wir auch quantitative Aussagen machen. Allerdings liegt die Verbindungsstrecke zwischen zwei Punkten  $x_0, x_1 \in \Omega$  nicht unbedingt in  $\Omega$ , das müssen wir extra garantieren.

**Definition 4.2** Eine offene Menge  $\Omega \subset \mathbb{R}^n$  heißt *konvex*, falls gilt:

$$x_0, x_1 \in \Omega \quad \Rightarrow \quad (1-t)x_0 + tx_1 \in \Omega \quad \text{für alle } t \in [0, 1].$$

**Satz 4.2 (Schrankensatz)** Sei  $\Omega \subset \mathbb{R}^n$  offen und konvex, und  $f \in C^1(\Omega)$ . Es gelte  $|Df(x)| \leq L < \infty$  für alle  $x \in \Omega$ . Dann folgt

$$|f(x_1) - f(x_0)| \leq L|x_1 - x_0| \quad \text{für alle } x_0, x_1 \in \Omega.$$

BEWEIS: Für  $x(t) = (1-t)x_0 + tx_1$  gilt  $x'(t) = x_1 - x_0$ , also folgt aus Lemma 4.1

$$\begin{aligned} |f(x_1) - f(x_0)| &= \left| \int_0^1 Df(x(t))(x_1 - x_0) dt \right| \\ &\leq \int_0^1 |Df(x(t))(x_1 - x_0)| dt \\ &\leq \int_0^1 |Df(x(t))| |x_1 - x_0| dt \quad (\text{Cauchy-Schwarz}) \\ &\leq L|x_1 - x_0|. \end{aligned}$$

□

Als nächstes diskutieren wir lokale Extrema von Funktionen mehrerer Variabler.

**Definition 4.3** Ein Punkt  $x \in \Omega$  mit  $Df(x) = 0$  heißt kritischer Punkt von  $f$ .

Eine quadratische Form  $q(x) = \langle Ax, x \rangle$  mit  $A \in \mathbb{R}^{n \times n}$  symmetrisch hat den Gradienten  $\text{grad } q(x) = 2Ax$ , also ist der Nullpunkt kritischer Punkt von  $q$ . In der linearen Algebra hatten wir gezeigt, dass  $q(x)$  bezüglich einer geeigneten Orthonormalbasis  $v_1, \dots, v_n$  eine Summe von Quadraten ist:

$$q(x) = \sum_{i=1}^n \lambda_i \xi_i^2 \quad \text{für } x = \sum_{i=1}^n \xi_i v_i.$$

Der Nullpunkt ist genau dann ein Minimum von  $q(x)$ , wenn die Eigenwerte  $\lambda_i$  von  $A$  alle nichtnegativ sind. Es handelt sich um ein striktes Minimum, wenn die  $\lambda_i$  alle positiv sind. Für Maxima gilt genau das Entsprechende, nur mit nichtnegativen bzw. positiven Eigenwerten. Für quadratische Funktionen einer Variablen ist ein kritischer Punkt automatisch ein Extremum. Das gilt nicht bei Funktionen mehrerer Variabler, zum Beispiel hat die quadratische Form  $q(x_1, x_2) = x_1^2 - x_2^2$  im Nullpunkt weder ein Minimum noch ein Maximum, auch nicht lokal.

**Definition 4.4** Eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt positiv definit (bzw. positiv semidefinit), falls gilt

$$\langle Ax, x \rangle > 0 \quad (\text{bzw. } \langle Ax, x \rangle \geq 0) \quad \text{für alle } x \in \mathbb{R}^n, x \neq 0.$$

Entsprechend für negativ (semi-)definit.

Bei dem Begriff der *Monotonie* lassen wir den Gleichheitsfall zu, und sprechen ausdrücklich von *streng monoton*, wenn wir die Gleichheit ausschließen wollen. Hier ist der Sprachgebrauch anders, *definit* bedeutet automatisch die strikte Ungleichung. In manchen Büchern werden die Symbole  $A > 0$  bzw.  $A \geq 0$  benutzt. Das Beispiel oben zeigt allerdings, dass für eine symmetrische Matrix weder  $A \geq 0$  noch  $A \leq 0$  gelten muss. Matrizen mit sowohl positiven als auch negativen Eigenwerten heißen indefinit.

**Definition 4.5** Sei  $f \in C^2(\Omega)$  mit  $\Omega \subset \mathbb{R}^n$  offen. Die zweite Ableitung oder Hesse-Matrix von  $f$  im Punkt  $x \in \Omega$  ist

$$D^2 f(x) = (\partial_{ij}^2 f(x)) \in \mathbb{R}^{n \times n}.$$

Die Symmetrie der Hesse-Matrix wurde in Satz 2.2 gezeigt.

**Satz 4.3 (notwendige Bedingungen für Extrema)** Sei  $\Omega \subset \mathbb{R}^n$  offen, und  $f \in C^2(\Omega)$  habe in  $x \in \Omega$  ein Minimum (Maximum). Dann gilt:

- (1)  $Df(x) = 0$ , also ist  $x$  kritischer Punkt.
- (1)  $D^2 f(x)$  ist positiv semidefinit (negativ semidefinit).

BEWEIS: Für  $v \in \mathbb{R}^n$  berechnen wir mit der Kettenregel

$$\begin{aligned} \frac{d}{dt} f(x + tv) &= \sum_{j=1}^n \partial_j f(x + tv) v_j = Df(x + tv)v \\ \frac{d^2}{dt^2} f(x + tv)|_{t=0} &= \sum_{i,j=1}^n \partial_{ij}^2 f(x) v_i v_j = \langle D^2 f(x)v, v \rangle. \end{aligned}$$

Die Funktion  $t \mapsto f(x + tv)$  hat bei  $t = 0$  ein Minimum, also folgt

$$0 = \frac{d}{dt} f(x + tv)|_{t=0} = Df(x)v \quad \text{für alle } v \in \mathbb{R}^n.$$

Dies bedeutet  $Df(x) = 0$ . Weiter folgt aus dem eindimensionalen Fall

$$0 \leq \frac{d^2}{dt^2} f(x + tv)|_{t=0} = \langle D^2 f(x)v, v \rangle \quad \text{für alle } v \in \mathbb{R}^n.$$

Somit ist  $D^2 f(x)$  positiv semidefinit. □

Sei nun  $Df(x) = 0$ , also  $x \in \Omega$  kritischer Punkt der Funktion  $f$ . Was sind *hinreichende Bedingungen* dafür, dass  $f$  am Punkt  $x \in \Omega$  zum Beispiel ein Minimum hat?

**Satz 4.4 (Hinreichende Bedingung für lokales Minimum)** Sei  $x \in \Omega$  kritischer Punkt der Funktion  $f \in C^2(\Omega)$ . Ist  $D^2 f(x)$  positiv definit, so hat  $f$  an der Stelle  $x$  ein lokales, isoliertes Minimum.

Um diese Aussage zu zeigen, brauchen wir eine lokale Entwicklung von  $f$ .

**Satz 4.5 (Taylor-Approximation)** Für  $f \in C^2(\Omega)$  sei  $q(x)$  das Taylorpolynom zweiter Ordnung im Entwicklungspunkt  $x_0 \in \Omega$ , das heißt

$$q(x) = f(x_0) + Df(x_0)(x - x_0) + \frac{1}{2} D^2 f(x_0)(x - x_0, x - x_0).$$

Dann folgt

$$\lim_{x \rightarrow x_0} \frac{f(x) - q(x)}{|x - x_0|^2} = 0.$$

BEWEIS: Wir betrachten die eindimensionale Funktion

$$\varphi : [0, 1] \rightarrow \mathbb{R}, \varphi(t) = f(x_0 + tv) \quad \text{mit } v = x - x_0.$$

Die Ableitungen der Funktion  $\varphi$  lauten

$$\varphi'(t) = Df(x_0 + tv)v \quad \text{und} \quad \varphi''(t) = D^2f(x_0 + tv)(v, v).$$

Nach der Taylorschen Formel, siehe Satz 4.1, gibt es ein  $\tau \in [0, 1]$  mit

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(\tau),$$

beziehungsweise mit  $\xi = (1 - \tau)x_0 + \tau x$

$$\begin{aligned} f(x) &= f(x_0) + Df(x_0)(x - x_0) + \frac{1}{2}D^2f(\xi)(x - x_0, x - x_0) \\ &= q(x) + \frac{1}{2}(D^2f(\xi) - D^2f(x_0))(x - x_0, x - x_0). \end{aligned}$$

Es folgt

$$\frac{f(x) - q(x)}{|x - x_0|^2} = \frac{1}{2}(D^2f(\xi) - D^2f(x_0))\left(\frac{x - x_0}{|x - x_0|}, \frac{x - x_0}{|x - x_0|}\right).$$

Mit  $x \rightarrow x_0$  geht  $D^2f(\xi)$  gegen  $D^2f(x_0)$ , also geht die rechte Seite gegen Null.  $\square$

In der Landau-Notation bedeutet das  $f(x) - q(x) = o(|x - x_0|^2)$ , oder

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + \frac{1}{2}D^2f(x_0)(x - x_0, x - x_0) + o(|x - x_0|^2) \quad \text{für } x \rightarrow x_0.$$

Allgemeiner kann für eine  $C^k$ -Funktion eine Taylorentwicklung der Ordnung  $k$  hergeleitet werden, allerdings erfordert das einen gewissen Notationsaufwand.

Wir zeigen jetzt das hinreichende Kriterium von Satz 4.4. Sei  $Df(x_0) = 0$ , also  $x_0$  kritischer Punkt, und  $D^2f(x_0)$  sei positiv definit. Wir wenden die Hauptachsentransformation auf die symmetrische Matrix  $D^2f(x_0)$  bzw. die zugehörige quadratische Form an. Genauer wurde in Satz 7.3 gezeigt: es gibt einen Vektor  $v_1 \in \mathbb{R}^n$  mit  $|v_1| = 1$ , so dass gilt:

$$D^2f(x_0)(v_1, v_1) = \inf\{D^2f(x_0)(v, v) : |v| = 1\} =: \lambda_1.$$

Da  $D^2f(x_0)$  positiv definit, ist  $\lambda_1 > 0$  und damit gilt für alle  $v \neq 0$

$$D^2f(x_0)(v, v) = D^2f(x_0)\left(\frac{v}{|v|}, \frac{v}{|v|}\right) |v|^2 \geq \lambda_1 |v|^2.$$

Es folgt wegen  $Df(x_0) = 0$

$$\begin{aligned} f(x) - f(x_0) &= \frac{1}{2}D^2f(x_0)(x - x_0, x - x_0) + f(x) - q(x) \\ &\geq \left(\frac{\lambda_1}{2} + \frac{f(x) - q(x)}{|x - x_0|^2}\right) |x - x_0|^2 \\ &\geq \frac{\lambda_1}{4} |x - x_0|^2 > 0, \end{aligned}$$

für  $x \neq x_0$  hinreichend nahe bei  $x_0$ .

## 5 Parameterabhängige Integrale

In diesem Abschnitt geht es um Integrale, die von einem Parameter abhängen, und die Differentiation unter dem Integralzeichen. Solche Integrale treten in vielen Anwendungen auf, als Beispiel sei die Fouriertransformierte genannt:

$$\hat{f}(p) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ipx} dx.$$

Der Ehrlichkeit halber sei gesagt, dass wir dieses Beispiel hier nicht behandeln werden: wir betrachten nur Integrale auf einem endlichen Intervall, um nicht mit zusätzlichen Konvergenzproblemen kämpfen zu müssen. Die Differentiation unter dem Integralzeichen ist auch bei vielen uneigentlichen Integralen möglich, für hinreichende Kriterien verweisen wir auf die Literatur. Unsere Version mit endlichen Intervallen hat bereits interessante Anwendungen, unter anderem werden wir die Euler-Lagrange Gleichung für eindimensionale Variationsintegrale herleiten.

Sei  $D \subset \mathbb{R}$  ein offenes (Parameter-)Intervall und  $I = [a, b]$  ein kompaktes Intervall. Für eine gegebene Funktion  $f : D \times I \rightarrow \mathbb{R}$ ,  $f = f(x, y)$ , betrachten wir die neue Funktion

$$\phi : D \rightarrow \mathbb{R}, \phi(x) = \int_I f(x, y) dy. \quad (5.2)$$

Diese Funktion nennt man ein parameterabhängiges Integral. Damit  $\phi$  wohldefiniert ist, müssen die Integrale existieren. Das ist garantiert, wenn für alle  $x \in D$  die Funktion  $f(x, \cdot) : I \rightarrow \mathbb{R}$ ,  $y \mapsto f(x, y)$ , stetig ist. Wir interessieren uns für die Stetigkeit und die Ableitung der Funktion  $\phi(x)$ . Hierfür brauchen wir folgende Aussage.

**Satz 5.1 (gleichmäßige Stetigkeit)** Sei  $f : K \rightarrow \mathbb{R}$  stetig, wobei  $K \subset \mathbb{R}^n$  abgeschlossen und beschränkt. Dann ist  $f$  gleichmäßig stetig, das heißt es gilt

$$\text{osc}_f(\delta) = \sup_{x, x' \in K, |x-x'| < \delta} |f(x) - f(x')| \rightarrow 0 \quad \text{mit } \delta \searrow 0.$$

Die Funktion  $\text{osc}_f(\delta)$  heißt Oszillationsfunktion von  $f$ .

Dies kann durch Widerspruch gezeigt werden. Da die Information eher technischer Natur ist, wollen wir hier darauf verzichten. Stattdessen kommen wir gleich zum ersten Resultat.

**Satz 5.2 (Stetigkeit von Parameterintegralen)** Sei  $f \in C^0(D \times I)$ ,  $f = f(x, y)$ , wobei  $D \subset \mathbb{R}^n$  offen und  $I = [a, b]$ . Dann ist die Funktion

$$\phi : D \rightarrow \mathbb{R}, \phi(x) = \int_I f(x, y) dy,$$

stetig.

BEWEIS: Sei erst  $D$  abgeschlossen und beschränkt. Dann gilt für  $x, x' \in D$  mit  $|x - x'| < \delta$

$$|\phi(x') - \phi(x)| = \left| \int_I (f(x', y) - f(x, y)) dy \right| \leq (b - a) \text{osc}_f(\delta).$$

Die rechte Seite geht mit  $\delta \searrow 0$  gegen Null. Ist  $D$  offen, so können wir  $D$  durch abgeschlossene und beschränkte Intervalle ausschöpfen, also ist  $\phi(x)$  stetig auf  $D$ .  $\square$

Wir gehen direkt weiter zur Differenzierbarkeit und Berechnung der Ableitung.

**Satz 5.3 (Differentiation unter dem Integral)** Sei  $D \subset \mathbb{R}$  offen und  $I = [a, b]$ . Für  $f \in C^0(D \times I)$  setze  $\phi : D \rightarrow \mathbb{R}$ ,  $\phi(x) = \int_I f(x, y) dy$ . Ist  $\frac{\partial f}{\partial x} \in C^0(D \times I)$ , so folgt

$$\frac{d\phi}{dx}(x) = \int_I \frac{\partial f}{\partial x}(x, y) dy,$$

und es gilt  $\phi \in C^1(D)$ .

BEWEIS: Wir nehmen wieder  $D$  abgeschlossen und beschränkt an. Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\frac{f(x+h, y) - f(x, y)}{h} = \frac{1}{h} \int_0^1 \frac{d}{dt} f(x+th, y) ds = \int_0^1 \frac{\partial f}{\partial x}(x+th, y) dt.$$

Berechne nun

$$\begin{aligned} \left| \frac{\phi(x+h) - \phi(x)}{h} - \int_I \frac{\partial f}{\partial x}(x, y) dy \right| &= \left| \int_I \left( \frac{f(x+h, y) - f(x, y)}{h} - \frac{\partial f}{\partial x}(x, y) \right) dy \right| \\ &= \left| \int_I \int_0^1 \left( \frac{\partial f}{\partial x}(x+th, y) - \frac{\partial f}{\partial x}(x, y) \right) dt dy \right| \\ &\leq \int_I \int_0^1 \left| \frac{\partial f}{\partial x}(x+th, y) - \frac{\partial f}{\partial x}(x, y) \right| dt dy \\ &\leq (b-a) \operatorname{osc}_{\frac{\partial f}{\partial x}}(|h|). \end{aligned}$$

Mit  $h \rightarrow 0$  geht die rechte Seite gegen Null, also ist die Formel gezeigt. Die Stetigkeit der Ableitung folgt aus Satz 5.2.  $\square$

Als Warnung betrachten wir für  $f : [a, b] \rightarrow \mathbb{R}$  stetig das Integral

$$\phi(x) = \frac{1}{2} \int_a^b \operatorname{sign}(x-y) f(y) dy.$$

Es ist  $\frac{\partial}{\partial x} \operatorname{sign}(x-y) = \operatorname{sign}'(x-y) = 0$ , außer wenn  $x = y$ . Würden wir diese Ausnahmestelle ignorieren und naiv unter dem Integral differenzieren, so wäre

$$\phi'(x) = \frac{1}{2} \int_a^b \operatorname{sign}'(x-y) f(y) dy = 0 \quad \text{für } x \in (a, b).$$

Das ist aber falsch, durch Aufspalten des Integrals sieht man

$$\phi'(x) = \frac{1}{2} \frac{d}{dx} \left( \int_a^x f(y) dy - \int_x^b f(y) dy \right) = f(x).$$

Formal gilt also die Gleichung

$$f(x) = \int_a^b \operatorname{sign}'(x-y) f(y) dy \quad \text{für } x \in (a, b).$$

Das wird oft in der Form  $\operatorname{sign}'(x-y) = \delta(x-y)$  notiert. Dabei soll  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  eine Funktion sein mit  $\delta(x) = 0$  für  $x \neq 0$ , andererseits soll aber gelten

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \quad \text{sofern } f(x) = 0 \text{ für } |x| \text{ groß.}$$

Eine solche Funktion  $\delta(x)$  kann es aber nicht geben. Um der Notation Sinn zu verleihen, muss man den Raum der Funktionen verlassen und Distributionen einführen. Die Notation kann nützlich sein, wenn man diesen Hintergrund versteht.

Eine nützliche Anwendung der Parameterableitung ist die Vertauschbarkeit der Integrationsreihenfolge in iterierten Integralen.

**Satz 5.4 (Fubini)** Seien  $I = [a, b]$ ,  $J = [\alpha, \beta]$  kompakte Intervalle. Dann gilt

$$\int_{\alpha}^{\beta} \left( \int_a^b f(x, y) dx \right) dy = \int_a^b \left( \int_{\alpha}^{\beta} f(x, y) dy \right) dx \quad \text{für } f \in C^0(I \times J).$$

BEWEIS: Wir betrachten die Funktionen  $\phi, \psi : [a, b] \rightarrow \mathbb{R}$  mit

$$\phi(x) = \int_{\alpha}^{\beta} \left( \int_a^x f(\xi, y) d\xi \right) dy \quad \text{und} \quad \psi(x) = \int_a^x \left( \int_{\alpha}^{\beta} f(\xi, y) dy \right) d\xi.$$

Nach Satz 5.2 sind  $y \mapsto \int_a^x f(\xi, y) d\xi$  sowie  $\xi \mapsto \int_{\alpha}^{\beta} f(\xi, y) dy$  stetig, und damit beide Seiten wohldefiniert mit  $\phi(a) = \psi(a) = 0$ . Wir zeigen  $\phi'(x) = \psi'(x)$  für alle  $x \in I$ , woraus die Behauptung  $\phi(b) = \psi(b)$  folgt. Der Hauptsatz der Differential- und Integralrechnung liefert

$$\psi'(x) = \int_{\alpha}^{\beta} f(x, y) dy.$$

Weiter hat die Funktion  $F(x, y) = \int_a^x f(\xi, y) d\xi$  die partielle Ableitung  $\frac{\partial F}{\partial x} = f \in C^0(I \times J)$ , und aus Satz 5.3 folgt

$$\phi'(x) = \int_{\alpha}^{\beta} \frac{\partial F}{\partial x}(x, y) dy = \int_{\alpha}^{\beta} f(x, y) dy.$$

□

Jetzt kommen wir zur Variationsrechnung. Auf der Menge aller Kurven  $x : I = [a, b] \rightarrow \Omega$ ,  $x = x(t)$ , betrachten wir das Funktional

$$\mathcal{F} : C^1(I, \Omega) \rightarrow \mathbb{R}, \quad \mathcal{F}(x) = \int_I f(t, x(t), x'(t)) dt.$$

$\mathcal{F}$  ist keine Funktion von endlich vielen Variablen, sondern ist auf dem Raum  $C^1(I, \Omega)$  definiert, darum spricht man auch von einem *Funktional* statt einer Funktion. Die Bezeichnung *Variationsintegral* ist ebenfalls üblich. Die Funktion

$$f : I \times \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad f = f(t, x, v),$$

heißt Lagrangefunktion. Durch sie ist das Funktional  $\mathcal{F}$  definiert.

**Beispiel 5.1** Die Bogenlänge von  $x \in C^1(I, \Omega)$  ist

$$\mathcal{F}(x) = \int_I |x'(t)| dt.$$

Für die Bogenlänge gilt also  $f(t, x, v) = |v|$ .

**Beispiel 5.2** Ist der Boden von unterschiedlicher Qualität, so hängt die Laufzeit nicht nur von der Länge des Wegs ab. Dies modelliert das Integral

$$\mathcal{F}(c) = \int_I \omega(x(t)) |x'(t)| dt.$$

Hier ist also  $f(t, x, v) = \omega(x)|v|$ . Dasselbe Integral tritt auch bei der Lichtausbreitung in Medien mit nichtkonstanter optischer Dichte auf.

**Beispiel 5.3** Es sei  $x = x(t)$  die Bahn eines Teilchens der Masse  $m$  in einem Kraftfeld mit Potential  $V : \Omega \rightarrow \mathbb{R}$ . Zur Zeit  $t$  hat das Teilchen dann die kinetische Energie  $\frac{m}{2}|x'(t)|^2$  und die potentielle Energie  $V(x(t))$ . Das sogenannte Wirkungsintegral ist

$$\mathcal{F}(x) = \int_I \left( \frac{m}{2} |x'(t)|^2 - V(x(t)) \right) dt.$$

Die zugehörige Lagrangefunktion ist  $f(t, x, v) = \frac{m}{2}|v|^2 - V(x)$ .

Besonderes Interesse gilt natürlich den optimalen Kurven oder Bahnen, die das Funktional  $\mathcal{F}(x)$  minimieren oder maximieren gegenüber allen Vergleichskurven. Unser Ziel ist es, diese Kurven zu charakterisieren. Wie bei der Diskussion der Extrema zeigen wir allerdings nur eine notwendige Bedingung. Dazu betrachten wir eine Schar (oder: *Variation*) von Kurven

$$x \in C^2((-\varepsilon_0, \varepsilon_0) \times I, \mathbb{R}^n), \quad x = x(\varepsilon, t) \quad \text{mit } x(0, t) = x(t). \quad (5.3)$$

Weiter definieren wir das Vektorfeld der Variation durch

$$\varphi : I \rightarrow \mathbb{R}^n, \quad \varphi(t) = \frac{\partial x}{\partial \varepsilon}(0, t).$$

**Lemma 5.1 (Erste Variation)** Sei  $f \in C^2(I \times \Omega \times \mathbb{R}^n)$ ,  $f = f(t, x, v)$ , wobei  $I = [a, b]$  und  $\Omega \subset \mathbb{R}^n$  offen. Definiere für  $x = x(\varepsilon, t)$  wie in (5.3) die Funktion

$$\phi(\varepsilon) = \mathcal{F}(x(\varepsilon, \cdot)) = \int_I f(t, x(\varepsilon, t), \frac{\partial x}{\partial t}(\varepsilon, t)) dt.$$

Dann gilt die Formel

$$\frac{d\phi}{d\varepsilon} \Big|_{\varepsilon=0} = \int_I \sum_{i=1}^n L_f(x)_i \varphi_i dt + \left[ \sum_{i=1}^n \frac{\partial f}{\partial v_i}(t, x, x') \varphi_i \right]_{t=a}^{t=b}. \quad (5.4)$$

Dabei ist  $L_f(x) : I \rightarrow \mathbb{R}^n$  gegeben durch

$$L_f(x)_i = \frac{\partial f}{\partial x_i}(t, x, x') - \frac{d}{dt} \left[ \frac{\partial f}{\partial v_i}(t, x, x') \right].$$

BEWEIS: Differentiation unter dem Integral und die Kettenregel liefern

$$\frac{d\phi}{d\varepsilon} \Big|_{\varepsilon=0} = \int_I \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i}(t, x, x') \frac{\partial x_i}{\partial \varepsilon}(0, t) + \frac{\partial f}{\partial v_i}(t, x, x') \frac{\partial^2 x_i}{\partial \varepsilon \partial t}(0, t) \right) dt.$$

Nun gilt

$$\frac{\partial x}{\partial \varepsilon}(0, t) = \varphi(t) \quad \text{und} \quad \frac{\partial^2 x}{\partial \varepsilon \partial t}(0, t) = \frac{\partial^2 x}{\partial t \partial \varepsilon}(0, t) = \varphi'(t).$$

Durch partielle Integration des hinteren Terms folgt nun

$$\frac{d\phi}{d\varepsilon}\Big|_{\varepsilon=0} = \int_I \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i}(t, x, x') - \frac{d}{dt} \left[ \frac{\partial f}{\partial v_i}(t, x, x') \right] \right) \varphi_i + \left[ \sum_{i=1}^n \frac{\partial f}{\partial v_i}(t, x, x') \varphi_i \right]_{t=a}^{t=b}.$$

Damit ist das Lemma bewiesen.  $\square$

Im folgenden Satz besagt die Voraussetzung, dass die Richtungsableitung des Funktionals  $\mathcal{F}(x)$  an der Stelle  $x$  in Richtung von  $\varphi$  gleich Null ist, für alle  $\varphi$  wie beschrieben. Mit andern Worten können wir  $x$  als kritischen Punkt des Funktionals  $\mathcal{F}(x)$  betrachten.

**Satz 5.5 (Euler-Lagrange-Gleichung)** Sei  $\mathcal{F}(x) = \int_a^b f(t, x, x')$  wie in Lemma 5.1, und

$$\frac{d}{d\varepsilon} \mathcal{F}(x + \varepsilon\varphi)\Big|_{\varepsilon=0} = 0 \quad \text{für alle } \varphi \in C^2(I, \mathbb{R}^n) \text{ mit } \varphi(a) = \varphi(b) = 0.$$

Dann ist  $x(t)$  eine Lösung der Euler-Lagrange-Gleichungen

$$L_f(x) = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i}(t, x, x') - \frac{d}{dt} \frac{\partial f}{\partial v_i}(t, x, x') = 0 \quad \text{für } i = 1, \dots, n.$$

BEWEIS: Wähle  $\varphi = \eta e_i$  mit  $\eta \in C^2(I)$  reellwertig,  $\eta(a) = \eta(b) = 0$ . Die Randterme in Lemma 5.1 verschwinden, da  $\varphi(a) = \varphi(b) = 0$ . Also gilt für alle diese  $\eta$

$$\int_I L_f(x)_i \eta dt = 0 \quad \text{für alle } \eta \in C^2(I) \text{ mit } \eta(a) = \eta(b) = 0.$$

Hieraus folgt  $L_f(x)_i = 0$ .  $\square$

Der Beweis für den letzten Schluss ist nicht schwierig, wir lassen ihn dennoch aus. Stattdessen interpretieren wir die Aussage mit dem Integral-Skalarprodukt

$$\int_I L_f(x)_i \eta dt = \langle L_f(x)_i, \eta \rangle.$$

Wir haben die Information, dass  $L_f(x)_i$  auf alle  $\eta$  senkrecht steht. Der Schluss besagt, dass dies nur die Nullfunktion sein kann.

**Beispiel 5.4** Für die Bogenlänge gilt  $f(t, x, v) = |v|$ , also  $\frac{\partial f}{\partial x_i} = 0$  und  $\frac{\partial f}{\partial v_i} = \frac{v_i}{|v|}$  für  $v \neq 0$ . Damit ist die Euler-Lagrange-Gleichung

$$L_f(x) = -\frac{d}{dt} \frac{x'}{|x|} = 0.$$

Wir müssen dabei  $x'(t) \neq 0$  für alle  $t \in I$  voraussetzen. Die Gleichung sagt aus, dass der Einheitstangentenvektor von  $x(t)$  konstant ist. Dies bedeutet, dass  $x(t)$  eine gerade Strecke parametrisiert.

**Beispiel 5.5** Für das Wirkungsintegral  $f(t, x, v) = \frac{m}{2}|v|^2 - V(x)$  haben wir  $\frac{\partial f}{\partial v_i} = mv_i$  und

$$\frac{\partial f}{\partial x_i}(t, x, v) = -\frac{\partial V}{\partial x_i}(x) =: F_i(x).$$

Das Vektorfeld  $F(x) = -\text{grad } V(x)$  ist das Kraftfeld des Potentials, das Minuszeichen ist in der Physik üblich. Die Euler-Lagrange Gleichung lautet nun

$$L_f(x) = F(x) - \frac{d}{dt}(mx') = F(x) - mx'' = 0.$$

Dies sind die Newtonschen Bewegungsgleichungen.

## 6 Kurvenintegrale

**Definition 6.1** Auf der offenen Menge  $\Omega \subset \mathbb{R}^n$  sei ein stetiges Vektorfeld  $F : \Omega \rightarrow \mathbb{R}^n$  gegeben. Das Kurvenintegral von  $F$  längs einer  $C^1$ -Kurve  $\gamma : [a, b] \rightarrow \Omega$  ist

$$\int_{\gamma} F \cdot \vec{ds} := \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt.$$

In der Physik ist zum Beispiel  $F$  das Gravitationsfeld, und das Kurvenintegral ergibt die Arbeit, die beim Transport einer Masse längs einer Kurve innerhalb des Feldes verrichtet wird. Wir schreiben hier  $F \cdot \gamma'$  für das Standard-Skalarprodukt, das ist die übliche Notation. Der Ausdruck  $\vec{ds}$  wird als vektorielles Längenelement bezeichnet, jedoch ist das nur eine formale Notation; das Integral ist durch die rechte Seite definiert. Immerhin ist die Merkregel nützlich, dass  $\vec{ds}$  durch  $\gamma'(t) dt$  zu ersetzen ist.

**Beispiel 6.1 (Gravitationsfeld)** Das Gravitationsfeld der Sonne ist gegeben durch

$$F : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^3, F(x) = -\frac{Cx}{|x|^3} \quad \text{mit } C > 0.$$

Längs einer Bahn  $x(t)$ ,  $t \in [a, b]$ , wird folgende Arbeit verrichtet:

$$E(\gamma) = - \int_a^b \frac{Cx(t)}{|x(t)|^3} \cdot x'(t) dt = \int_a^b \frac{d}{dt} \frac{C}{|x(t)|} = \left( \frac{C}{|x(b)|} - \frac{C}{|x(a)|} \right).$$

**Beispiel 6.2 (Winkelvektorfeld I)** Betrachte das Vektorfeld

$$W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, W(x, y) = \left( -\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

Ist  $\gamma(t) = r(t)(\cos \theta(t), \sin \theta(t))$ ,  $t \in [a, b]$ , mit  $C^1$ -Funktionen  $r(t) > 0$  und  $\theta(t)$ , so gilt

$$\int_{\gamma} W \cdot \vec{ds} = \int_a^b \frac{1}{r} \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \cdot \left( r' \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + r\theta' \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \right) dt = \theta(b) - \theta(a). \quad (6.5)$$

Also liefert das Kurvenintegral den umlaufenen Winkel.

Wir wollen einige Eigenschaften des Kurvenintegrals zusammenstellen. Direkt aus der Definition folgt die Linearität

$$\int_{\gamma} (\lambda_1 F_1 + \lambda_2 F_2) \cdot \vec{ds} = \lambda_1 \int_{\gamma} F_1 \cdot \vec{ds} + \lambda_2 \int_{\gamma} F_2 \cdot \vec{ds}. \quad (6.6)$$

Als zweites haben wir die Additivität unter Zerlegungen: ist  $a = t_0 < t_1 < \dots < t_N = b$ , so gilt für  $\gamma_i = \gamma|_{[t_{i-1}, t_i]}$

$$\int_{\gamma} F \cdot \vec{ds} = \sum_{i=1}^N \int_{\gamma_i} F \cdot \vec{ds}. \quad (6.7)$$

Insbesondere ist das Kurvenintegral auch für stückweise  $C^1$ -Kurven definiert, zum Beispiel wenn  $C^1$ -Kurven aneinander gehängt werden.

Die dritte Eigenschaft ist die Invarianz unter Umparametrisierungen. Sei  $\gamma : [a, b] \rightarrow \Omega$ ,  $\gamma = \gamma(t)$ , eine gegebene  $C^1$ -Kurve. Durch die Substitution  $t = \tau(u)$  mit  $\tau : [\alpha, \beta] \rightarrow [a, b]$  bijektiv erhalten wir die Kurve

$$\tilde{\gamma} : [\alpha, \beta] \rightarrow \Omega, \quad \tilde{\gamma}(u) = \gamma(\tau(u)).$$

Wir berechnen nun mit der Kettenregel und Substitution  $\tau(u) = t$

$$\begin{aligned} \int_{\tilde{\gamma}} F \cdot \vec{ds} &= \int_{\alpha}^{\beta} F(\tilde{\gamma}(u)) \cdot \tilde{\gamma}'(u) du \\ &= \int_{\alpha}^{\beta} F(\gamma(\tau(u))) \cdot \gamma'(\tau(u)) \tau'(u) du \\ &= \int_{\tau(\alpha)}^{\tau(\beta)} F(\gamma(t)) \cdot \gamma'(t) dt. \end{aligned}$$

$\tau$  heißt orientierungserhaltend, wenn  $\tau(\alpha) = a$  und  $\tau(\beta) = b$ , andernfalls heißt  $\tau$  orientierungsumkehrend. Es gilt also

$$\int_{\gamma \circ \tau} F \cdot \vec{ds} = \pm \int_{\gamma} F \cdot \vec{ds}, \quad \text{falls } \tau \text{ orientierungserhaltend (-umkehrend)}. \quad (6.8)$$

Die Physiker nennen das Gravitationsfeld konservativ, weil der Energieerhaltungssatz gilt. Die verrichtete Arbeit zum Beispiel bei Transport einer Masse von der Mensa zum Schauinsland entspricht genau der zugewonnenen Lageenergie, und diese kommt beim Herunterrollen auch wieder heraus, theoretisch jedenfalls. Der Begriff des konservativen Feldes ist auch in der Mathematik interessant.

**Definition 6.2** Sei  $\Omega \subset \mathbb{R}^n$  offen. Ein Vektorfeld  $F \in C^0(\Omega, \mathbb{R}^n)$  heißt Gradientenfeld (bzw. konservativ), wenn es eine Funktion  $\varphi \in C^1(\Omega)$  gibt mit

$$\text{grad } \varphi = F.$$

Diese Funktion heißt Potentialfunktion von  $F$ .

Ist  $\Omega \subset \mathbb{R}^n$  wegwiese zusammenhängend, so unterscheiden sich zwei Potentialfunktionen  $\varphi_{1,2}$  von  $F$  nur um eine additive Konstante, denn es gilt

$$\text{grad}(\varphi_1 - \varphi_2) = F - F = 0.$$

Die Kenntnis einer Potentialfunktion ist sehr nützlich, denn damit kann das Kurvenintegral mühelos berechnet werden. Ist  $\varphi \in C^1(\Omega)$  Potentialfunktion von  $F(x)$ , so gilt für eine  $C^1$ -Kurve  $\gamma : [a, b] \rightarrow \Omega$  nach der Kettenregel

$$\begin{aligned} \int_{\gamma} F \cdot \vec{ds} &= \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \text{grad} \varphi(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_a^b \frac{d}{dt} \varphi(\gamma(t)) dt = \varphi(\gamma(b)) - \varphi(\gamma(a)). \end{aligned} \quad (6.9)$$

Für alle Kurven mit demselben Anfangs- und Endpunkt liefert das Kurvenintegral also denselben Wert. Die Frage ist nun, wann diese günstige Situation vorliegt, mit anderen Worten:

*Welche Vektorfelder  $F : \Omega \rightarrow \mathbb{R}^n$  haben eine Potentialfunktion?*

Für  $n = 1$  ist eine Potentialfunktion nichts anderes als eine Stammfunktion, also  $\varphi'(x) = F(x)$ . Nach dem Hauptsatz der Differential- und Integralrechnung können wir eine solche Stammfunktion immer finden. Nicht so für  $n \geq 2$ !

**Satz 6.1 (Rotationsfreiheit)** Für ein Gradientenfeld  $F \in C^1(\Omega, \mathbb{R}^n)$  gilt

$$\partial_i F_j = \partial_j F_i \quad \text{für alle } i, j = 1, \dots, n.$$

BEWEIS: Ist  $F = \text{grad} \varphi$ , so folgt  $\varphi \in C^2(\Omega)$  und wegen der Vertauschbarkeit der partiellen Ableitungen, Satz 2.2, gilt

$$\partial_i F_j = \partial_i \partial_j \varphi = \partial_j \partial_i \varphi = \partial_j F_i.$$

□

*Bemerkung.* Für  $n = 3$  lässt sich diese Bedingung schreiben als

$$\text{rot } F = (\partial_2 F_3 - \partial_3 F_2, \partial_3 F_1 - \partial_1 F_3, \partial_1 F_2 - \partial_2 F_1) = 0.$$

**Beispiel 6.3**  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $F(x, y) = (-y, x)$ , hat keine Stammfunktion, denn es ist

$$\partial_1 F_2 = 1, \text{ aber } \partial_2 F_1 = -1.$$

**Beispiel 6.4 (Winkelvektorfeld II)** Für das Vektorfeld

$$W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, W(x, y) = \left( -\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

berechnen wir

$$\partial_1 W_2 = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \partial_2 W_1,$$

das heißt die notwendige Bedingung aus Satz 6.1 ist erfüllt. Nach Beispiel 6.2 gilt jedoch für die Kurven  $\gamma_k : [0, 2\pi] \rightarrow \mathbb{R}^2 \setminus \{0\}$ ,  $\gamma_k(t) = r(\cos kt, \sin kt)$ ,

$$\int_{\gamma_k} W \cdot \vec{ds} = 2\pi k \quad (\neq 0 \text{ für } k \in \mathbb{Z} \setminus \{0\}).$$

Hätte  $W$  eine Potentialfunktion  $\varphi$  auf  $\mathbb{R}^2 \setminus \{0\}$ , so müsste gelten

$$\int_{\gamma_k} W \cdot \vec{ds} = \varphi(\gamma(2\pi k)) - \varphi(\gamma(0)) = 0.$$

Das Beispiel zeigt, dass die Rotationsfreiheit aus Satz 6.1 nicht immer ausreicht, um die Existenz eines Potentials zu garantieren. Wir werden sehen, dass es zusätzlich auf die Geometrie des Gebiets  $\Omega$  ankommt. Zunächst untersuchen wir, wie sich das Integral längs einer Schar von Kurven verhält. Statt *Schar* ist der modernere Ausdruck *Homotopie* üblich.

**Satz 6.2 (Homotopieformel)** Sei  $F \in C^1(\Omega, \mathbb{R}^n)$  und  $\gamma \in C^2([a, b] \times [0, 1], \Omega)$ ,  $\gamma = \gamma(s, t)$ . Dann gilt

$$\begin{aligned} \int_{\gamma(\cdot, 1)} F \cdot \vec{ds} - \int_{\gamma(\cdot, 0)} F \cdot \vec{ds} &= \int_{\gamma(b, \cdot)} F \cdot \vec{ds} - \int_{\gamma(a, \cdot)} F \cdot \vec{ds} \\ &\quad - \int_0^1 \int_a^b \sum_{i,j=1}^n (\partial_i F_j - \partial_j F_i)(\gamma) \frac{\partial \gamma^i}{\partial s} \frac{\partial \gamma^j}{\partial t} ds dt. \end{aligned}$$

BEWEIS: Das Kurvenintegral ist ein Variationsintegral mit der Euler-Lagrange Funktion

$$f(x, v) = F(x) \cdot v = \sum_{i=1}^n F_i(x) v_i.$$

Wir berechnen den zugehörigen Euler-Lagrange Operator, wobei  $' = \frac{\partial}{\partial s}$ .

$$\begin{aligned} L_f(\gamma)_j &= \frac{\partial f}{\partial x_j}(\gamma, \gamma') - \frac{d}{ds} \left[ \frac{\partial f}{\partial v_j}(\gamma, \gamma') \right] \\ &= \sum_{i=1}^n \frac{\partial F_i}{\partial x_j}(\gamma) \gamma'_i - \frac{d}{ds} F_j(\gamma) \\ &= \sum_{i=1}^n \left( \frac{\partial F_i}{\partial x_j}(\gamma) - \frac{\partial F_j}{\partial x_i}(\gamma) \right) \gamma'_i. \end{aligned}$$

Einsetzen des Ergebnisses in die Formel für die erste Variation ergibt

$$\begin{aligned} \frac{d}{dt} \int_{\gamma(\cdot, t)} F \cdot \vec{ds} &= \left[ \sum_{i=1}^n F_i(\gamma) \frac{\partial \gamma^i}{\partial t} \right]_{s=a}^{s=b} \\ &\quad + \int_a^b \sum_{i,j=1}^n \left( \frac{\partial F_i}{\partial x_j}(\gamma) - \frac{\partial F_j}{\partial x_i}(\gamma) \right) \frac{\partial \gamma^i}{\partial s} \frac{\partial \gamma^j}{\partial t} ds. \end{aligned}$$

Die Behauptung folgt nun durch Integration bezüglich  $t \in [0, 1]$ . □

**Folgerung 6.1 (Homotopieinvarianz)** Sei  $\partial_i F_j = \partial_j F_i$  für  $i, j = 1, \dots, n$  in Satz 6.2, und  $\gamma(a, t) = p$  und  $\gamma(b, t) = q$  für alle  $t \in [0, 1]$  (fester Anfangs- und Endpunkt). Dann folgt

$$\int_{\gamma(\cdot, 1)} F \cdot \vec{ds} = \int_{\gamma(\cdot, 0)} F \cdot \vec{ds}.$$

BEWEIS: Aufgrund der Voraussetzungen ist die rechte Seite in Satz 6.2 gleich Null.  $\square$

Wir kommen nun zurück auf die Geometrie des Gebiets.  $\Omega \subset \mathbb{R}^n$  heißt *einfach zusammenhängend*, wenn  $\Omega$  zusammenhängend ist und wenn gilt: sind  $\gamma_{0,1} : [a, b] \rightarrow \Omega$  zwei stetige Kurven von  $p$  nach  $q$ , so gibt es eine stetige Schar von Kurven  $\gamma(\cdot, t) : [a, b] \rightarrow \Omega$ ,  $t \in [0, 1]$ , von  $p$  nach  $q$  mit  $\gamma(\cdot, 0) = \gamma_0$  und  $\gamma(\cdot, 1) = \gamma_1$ . Es ist einfach zu sehen, dass ein konvexes Gebiet einfach zusammenhängend ist.

**Lemma 6.1** Sei  $\Omega \subset \mathbb{R}^n$  einfach zusammenhängend, und  $F \in C^1(\Omega, \mathbb{R}^n)$  erfülle  $\partial_i F_j = \partial_j F_i$ . Sind  $\gamma_0, \gamma_1 : [a, b] \rightarrow \Omega$  Kurven mit dem gleichem Anfangs- und Endpunkt, so gilt

$$\int_{\gamma_1} F \cdot \vec{ds} = \int_{\gamma_0} F \cdot \vec{ds}.$$

BEWEIS: Nach Voraussetzung gibt es eine Schar von Kurven  $\gamma(\cdot, t)$  wie oben beschrieben. Aus Folgerung 6.1 folgt dann die Behauptung. Es gibt dabei ein technisches Problem mit der Regularität der Homotopie, wir brauchen  $C^2$ . Aber das kann geregelt werden.  $\square$

**Satz 6.3 (Existenz eines Potentials)** Sei  $\Omega \subset \mathbb{R}^n$  einfach zusammenhängendes Gebiet. Erfüllt  $F \in C^1(\Omega, \mathbb{R}^n)$  die Integrabilitätsbedingung

$$\partial_i F_j = \partial_j F_i \quad \text{für alle } i, j = 1, \dots, n,$$

so hat  $F$  eine Potentialfunktion  $\varphi \in C^2(\Omega)$ .

BEWEIS: Sei  $x_0 \in \Omega$  fest. Zu  $x \in \Omega$  wählen wir eine Kurve  $\gamma_x \in C^2([0, 1], \Omega)$  mit  $\gamma_x(0) = x_0$  und  $\gamma_x(1) = x$ , und setzen

$$\varphi(x) = \int_{\gamma_x} F \cdot \vec{ds}.$$

Nach Lemma 6.1 hängt der Wert  $\varphi(x)$  nicht von der Wahl der Verbindungskurve  $\gamma_x$  ab, die Funktion  $\varphi$  ist definiert. Sei  $x \in \Omega$  und  $h \in \mathbb{R}$  klein. Wir erhalten eine Kurve von  $x_0$  nach  $x + he_j$ , indem wir  $\gamma_x$  mit der Kurve  $c(t) = x + the_j$ ,  $t \in [0, 1]$ , zusammensetzen. Es folgt

$$\frac{\varphi(x + he_j) - \varphi(x)}{h} = \int_c F \cdot \vec{ds} = \int_0^1 F(x + the_j) \cdot e_j dt \rightarrow F_j(x) \quad \text{mit } h \rightarrow 0.$$

Also gilt  $\partial_j \varphi = F_j$  für  $j = 1, \dots, n$ .  $\square$

**Beispiel 6.5 (Winkelvektorfeld III)** Nach Beispiel 6.2 erfüllt das Vektorfeld

$$W : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, W(x, y) = \left( -\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

die notwendige Bedingung  $\partial_1 W_2 = \partial_2 W_1$ . In Beispiel 6.4 haben wir gesehen, dass  $W$  aber keine Potentialfunktion auf  $\mathbb{R}^2 \setminus \{0\}$  hat. Somit ist  $\mathbb{R}^2 \setminus \{0\}$  nicht einfach zusammenhängend. Aber zum Beispiel  $H = \{(x, y) \in \mathbb{R}^2 : y > 0\}$  ist konvex und damit einfach zusammenhängend, auf  $H$  muss es eine Potentialfunktion geben. Tatsächlich können wir sie angeben, und zwar

$$\varphi : H \rightarrow \mathbb{R}, \varphi(x, y) = \arccos \frac{x}{x^2 + y^2}.$$

Beim Kurvenintegral werden Vektorfelder integriert, aber natürlich kann man auch skalare Funktionen längs Kurven integrieren. Wir stellen uns vor, dass  $\gamma : [a, b] \rightarrow \Gamma \subset \mathbb{R}^3$  eine Parametrisierung des Kurvenstücks  $\Gamma \subset \mathbb{R}^3$  ist. Für eine gegebene Funktion  $\varrho : \Gamma \rightarrow \mathbb{R}$  setzen wir dann

$$\int_{\Gamma} \varrho ds := \int_a^b \varrho(\gamma(t)) |\gamma'(t)| dt. \quad (6.10)$$

Für  $\varrho = 1$  erhalten wir das Bogenlängenintegral, und wir interpretieren  $ds = |\gamma'(t)| dt$  als Bogenlängenelement von  $\gamma$ . Die Funktion  $\varrho(x)$  könnte zum Beispiel eine Massendichte oder Ladungsdichte (pro Längeneinheit) sein, das Integral ist dann die Gesamtmasse oder -ladung von  $\Gamma$ . Ist  $\tilde{\gamma}(u) = \gamma(\tau(u))$  eine orientierungserhaltende Umparametrisierung,  $\tau : [\alpha, \beta] \rightarrow [a, b]$ , so ist  $\tau'(u) > 0$  und es folgt

$$\begin{aligned} \int_{\tilde{\gamma}} \varrho ds &= \int_{\alpha}^{\beta} \varrho(\tilde{\gamma}(u)) |\tilde{\gamma}'(u)| du \\ &= \int_{\alpha}^{\beta} \varrho(\gamma(\tau(u))) |\gamma'(\tau(u))| |\tau'(u)| du \\ &= \int_a^b \varrho(\gamma(t)) |\gamma'(t)| dt. \end{aligned}$$

Ist  $t = \tau(u)$  orientierungsumkehrend, so drehen sich die Integrationsgrenzen um, andererseits ist  $\tau'(u) < 0$  und somit  $|\tau'(u)| du = -\tau'(u) du = -dt$ . Im Effekt bleibt das Integral also auch dann gleich, anders als beim Kurvenintegral eines Vektorfelds.

Sei nun auf  $\Gamma$  eine Orientierung und damit eine Einheitstangente  $\vec{T} : \Gamma \rightarrow \mathbb{R}^3$ ,  $|\vec{T}(x)| = 1$ , gegeben. Sei  $\gamma : [a, b] \rightarrow \Gamma$  eine Parametrisierung mit der gegebenen Orientierung. Wir können dann das Kurvenintegral des Vektorfelds  $F$  wie folgt umschreiben:

$$\int_{\gamma} F \cdot \vec{ds} = \int_a^b F(\gamma(t)) \cdot \frac{\gamma'(t)}{|\gamma'(t)|} |\gamma'(t)| dt = \int_a^b F(\gamma(t)) \cdot \vec{T}(\gamma(t)) |\gamma'(t)| dt = \int_{\Gamma} F \cdot \vec{T} ds.$$

Wir sehen deutlicher, dass im Kurvenintegral eines Vektorfelds  $F$  die tangentielle Komponente von  $F$  integriert wird.

Mit diesen Vorbereitungen wollen wir nun die Homotopieformel aus Satz 6.2 in etwas anderem Licht betrachten, und zwar im dreidimensionalen Fall. Wir stellen uns vor, dass die Abbildung  $\gamma(s, t)$  ein Flächenstück parametrisiert:

$$\gamma : [a, b] \times [0, 1] \rightarrow \Sigma \subset \mathbb{R}^3, \quad \gamma = \gamma(s, t).$$

Der Rand  $\partial\Sigma$  von  $\Sigma$  wird durch die vier Kurven  $\gamma(\cdot, 0)$ ,  $\gamma(b, \cdot)$ ,  $\gamma(\cdot, 1)$  und  $\gamma(a, \cdot)$  parametrisiert. Durchlaufen wir den Rand des Rechtecks entgegen dem Uhrzeigersinn, so werden die ersten beiden Kurven orientierungstreu durchlaufen, bei den letzten beiden dreht sich die Orientierung um. Wir erhalten folgende Darstellung:

$$\oint_{\partial\Sigma} F \cdot \vec{T} ds = \int_{\gamma(\cdot, 0)} F \cdot \vec{ds} + \int_{\gamma(b, \cdot)} F \cdot \vec{ds} - \int_{\gamma(\cdot, 1)} F \cdot \vec{ds} - \int_{\gamma(a, \cdot)} F \cdot \vec{ds}. \quad (6.11)$$

Der Kreis beim Integralzeichen symbolisiert, dass über eine geschlossene Kurve integriert wird. Nun kommen wir zum Flächenintegral. Der Flächeninhalt eines Parallelogramms, das

von zwei Vektoren  $a, b \in \mathbb{R}^3$  aufgespannt wird, ist  $|a \times b|$ , wobei  $\times$  das Kreuzprodukt bedeutet. Die Abbildung  $\gamma(s, t)$  bildet ein Quadrat mit Ecken  $(s, t)$  und  $(s + \Delta s, t + \Delta t)$  näherungsweise auf ein Parallelogramm ab, das von den Vektoren  $(\partial_s \gamma) \Delta s$  sowie  $(\partial_t \gamma) \Delta t$  aufgespannt wird. Das Flächenelement sollte daher wie folgt gegeben sein:

$$dA = |\partial_s \gamma \times \partial_t \gamma| ds dt.$$

Das Integral einer Funktion  $\varphi : \Sigma \rightarrow \mathbb{R}$  über die Fläche wird daher wie folgt definiert:

$$\int_{\Sigma} \varphi dA = \int_0^1 \int_a^b \varphi(\gamma(s, t)) |\partial_s \gamma \times \partial_t \gamma| ds dt.$$

Wieder könnte  $\varphi$  eine Massen- oder Ladungsdichte sein, jetzt aber pro Flächeneinheit. Auf die Frage, ob die Definition unabhängig ist von der Parametrisierung, kommen wir später zurück. Die Einheitsnormale der Fläche lautet nun

$$\vec{N} = \frac{\partial_s \gamma \times \partial_t \gamma}{|\partial_s \gamma \times \partial_t \gamma|}.$$

In der folgenden Rechnung verwenden wir den  $\varepsilon_{ijk}$ -Tensor aus Aufgabe 4, Serie1.

$$\begin{aligned} \text{rot } F \cdot (\partial_s \gamma \times \partial_t \gamma) &= \sum_{i,j,k=1}^3 \varepsilon_{ijk} \partial_i F_j e_k \cdot \sum_{l,m,n=1}^3 \varepsilon_{lmn} \partial_s \gamma^l \partial_t \gamma^m e_n \\ &= \sum_{i,j,k,l,m=1}^3 \varepsilon_{ijk} \varepsilon_{lmk} \partial_i F_j \partial_s \gamma^l \partial_t \gamma^m \\ &= \sum_{i,j,l,m=1}^3 (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \partial_i F_j \partial_s \gamma^l \partial_t \gamma^m \\ &= \sum_{i,j=1}^3 \partial_i F_j \partial_s \gamma^i \partial_t \gamma^j - \sum_{i,j=1}^3 \partial_i F_j \partial_s \gamma^j \partial_t \gamma^i \\ &= \sum_{i,j=1}^3 (\partial_i F_j - \partial_j F_i) \partial_s \gamma^i \partial_t \gamma^j. \end{aligned}$$

Wir erhalten

$$\begin{aligned} \int_{\Sigma} \text{rot } F \cdot \vec{N} dA &= \int_0^1 \int_a^b \text{rot } F \cdot (\partial_s \gamma \times \partial_t \gamma) ds dt \\ &= \int_0^1 \int_a^b \sum_{i,j=1}^3 (\partial_i F_j - \partial_j F_i) \partial_s \gamma^i \partial_t \gamma^j ds dt. \end{aligned} \tag{6.12}$$

Setzen wir (6.11) und (6.12) in die Homotopieformel aus Satz 6.2 ein, so folgt insgesamt

**Satz 6.4 (von Stokes)** *Sei  $\Sigma \subset \mathbb{R}^3$  ein Flächenstück mit Normale  $\vec{N}$ . Sei  $\vec{T}$  die Einheits-tangente auf  $\partial \Sigma$ , so dass  $\vec{N} \times \vec{T}$  ins Innere des Flächenstücks weist. Dann gilt für jedes  $C^1$ -Vektorfeld  $F$  auf  $\mathbb{R}^3$*

$$\int_{\Sigma} \text{rot } F \cdot \vec{N} dA = \oint_{\partial \Sigma} F \cdot \vec{T} ds.$$

Sei  $\Sigma \subset \mathbb{R}^2$  und  $\vec{N} = e_3$  als Normale gewählt. Die Wahl der Tangente  $\vec{T}$  besagt dann, dass bei Durchlaufen des Randes das Gebiet auf der linken Seite liegt, das heißt  $\partial\Sigma$  wird gegen den Uhrzeigersinn durchlaufen. Das war auch unsere Wahl für  $[a, b] \times [0, 1]$ . Wir haben den Satz bewiesen für Flächenstücke, die auf einem Rechteck parametrisiert sind. Allgemeinere Flächen kann man unter Umständen durch Zerlegen behandeln. Für geschlossene Flächen sollten sich die Randterme gegenseitig wegheben, und es sollte folgen

$$\oint_{\Sigma} \operatorname{rot} F \cdot \vec{N} \, dA = 0.$$

Unsere Herleitung des Satzes von Stokes ist insofern rigoros, als die Integrale über die Parametrisierung genau definiert sind und damit die Formel hergeleitet wird. Dagegen haben wir den Begriff des Flächenstücks  $\Sigma$  und seines Randes  $\partial\Sigma$  nicht präzisiert.

## 7 Lokale Auflösung von Gleichungen

Für die Lösbarkeit von linearen Gleichungen haben wir in den linearen Algebra Kriterien entwickelt. Eine nichtlineare Gleichung

$$g(x, y) = 0$$

kann nur in Ausnahmefällen *explizit*, also durch algebraische Umformungen, gelöst werden. Hier stellen wir uns vor, dass  $x$  die Rolle eines Parameters spielt und die Lösung als Funktion  $y = u(x)$  dieses Parameters gefunden werden soll. Da keine Formel für die Lösung angegeben wird, spricht man von einer implizit gegebenen Funktion oder kurz einer impliziten Funktion. Wir betrachten genauer die Situation, dass schon eine konkrete Lösung  $g(x_0, y_0) = 0$  bekannt ist, und wollen die Lösungsmenge nahe bei  $(x_0, y_0)$  analysieren.

**Beispiel 7.1** Betrachte die quadratische Gleichung

$$g(x, y) = y^2 - xy + 1 = 0.$$

Für welche Parameter  $x$  ist die Gleichung lösbar, und wie hängt die Lösung von  $x$  ab? In diesem Fall können wir das voll beantworten:

- Für  $x \in (-2, 2)$  gibt es keine Lösung.
- Für  $x = \pm 2$  gibt es genau eine Lösung, nämlich  $y = \pm 1$ .
- Für  $|x| > 2$  gibt es genau zwei Lösungen  $y = u_{\pm}(x)$ , mit  $u_{\pm}(x) = \frac{x \pm \sqrt{x^2 - 4}}{2}$ .

Nehmen wir etwa die Lösung  $(x_0, y_0) = (3, \frac{3+\sqrt{5}}{2})$ . Für  $|x - x_0| < \delta$  gibt es genau eine Lösung mit  $|y - y_0| < \varepsilon$ , nämlich  $y = u_+(x)$ . Wir dürfen  $\delta > 0$  nicht zu groß wählen, denn für  $x \in (-2, 2)$  gibt es gar keine Lösung. Und wir dürfen  $\varepsilon > 0$  nicht zu groß wählen, sonst haben wir auch den Punkt  $(3, \frac{3-\sqrt{5}}{2})$  und die Eindeutigkeit ist verletzt. Ungünstig wird die Sache für  $(x_0, y_0) = (2, 1)$ . Die Existenz gilt für kein Intervall  $(x_0 - \delta, x_0 + \delta)$ , und die Eindeutigkeit ist in  $(y_0 - \delta, y_0 + \delta)$  ebenfalls nie erfüllt. Beachten Sie

$$\frac{\partial g}{\partial y}(x, y) = 2y - x.$$

Auf der Menge  $\{(x, y) : g(x, y) = 0\}$  ist  $\frac{\partial g}{\partial y}$  nur in den Punkten  $(2, 1)$  und in  $(-2, -1)$  gleich Null. Es wäre zu vermuten, dass diese partielle Ableitung eine Rolle spielt.

**Satz 7.1 (implizite Funktionen)** Sei  $g = g(x, y)$  stetig differenzierbar auf der offenen Menge  $D \subset \mathbb{R}^2$ , und sei  $(x_0, y_0) \in D$  mit  $g(x_0, y_0) = 0$ . Ist  $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$ , so gibt es  $\delta, \varepsilon > 0$  mit folgender Eigenschaft:

- (1) Zu jedem  $x \in (x_0 - \delta, x_0 + \delta)$  gibt es genau ein  $y \in (y_0 - \varepsilon, y_0 + \varepsilon)$  mit  $g(x, y) = 0$ .  
Damit ist die implizite Funktion  $y = u(x)$  definiert.
- (2)  $u : (x_0 - \delta, x_0 + \delta) \rightarrow (y_0 - \varepsilon, y_0 + \varepsilon)$  ist in  $x_0$  differenzierbar mit Ableitung

$$u'(x_0) = - \frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}.$$

BEWEIS: Wir nehmen  $\frac{\partial g}{\partial y}(x_0, y_0) > 0$  an. Da  $\frac{\partial g}{\partial y}$  stetig, gibt es ein  $\varepsilon > 0$  mit

$$\frac{\partial g}{\partial y}(x, y) > 0 \quad \text{auf } [x_0 - \varepsilon, x_0 + \varepsilon] \times [y_0 - \varepsilon, y_0 + \varepsilon].$$

Es folgt  $g(x_0, y_0 - \varepsilon) < 0 < g(x_0, y_0 + \varepsilon)$ . Da  $g$  stetig, gibt es ein  $\delta \in (0, \varepsilon]$  mit

$$\begin{aligned} g(x, y_0 - \varepsilon) &< 0 & \text{für } x \in [x_0 - \delta, x_0 + \delta] \\ g(x, y_0 + \varepsilon) &> 0 & \text{für } x \in [x_0 - \delta, x_0 + \delta]. \end{aligned}$$

Für  $x \in (x_0 - \delta, x_0 + \delta)$  ist nun  $g(x, \cdot)$  stetig auf  $[y_0 - \varepsilon, y_0 + \varepsilon]$  mit  $g(x, y_0 - \varepsilon) < 0 < g(x, y_0 + \varepsilon)$ . Nach dem Zwischenwertsatz gibt es ein  $y \in (y_0 - \varepsilon, y_0 + \varepsilon)$  mit  $g(x, y) = 0$ . Da  $g(x, \cdot)$  streng monoton wachsend auf  $[y_0 - \varepsilon, y_0 + \varepsilon]$ , ist  $u(x) := y$  eindeutig bestimmt.

Wäre  $u(x)$  nicht stetig in  $x_0$ , so gibt es  $x_n \rightarrow x_0$  mit  $u(x_n) \rightarrow y_1 \neq y_0$ , also

$$g(x_0, y_1) = \lim_{n \rightarrow \infty} g(x_n, u(x_n)) = 0.$$

Das widerspricht aber der Eindeutigkeit, also gilt  $u(x) \rightarrow u(x_0)$ . Um  $u'(x_0)$  zu bestimmen, verwenden wir den Mittelwertsatz: mit geeigneten Zwischenstellen  $\xi, \eta$  gilt

$$\begin{aligned} 0 &= g(x, u(x)) - g(x_0, u(x_0)) \\ &= g(x, u(x)) - g(x_0, u(x)) + g(x_0, u(x)) - g(x_0, u(x_0)) \\ &= \frac{\partial g}{\partial x}(\xi, u(x))(x - x_0) + \frac{\partial g}{\partial y}(x_0, \eta)(u(x) - u(x_0)), \end{aligned}$$

beziehungsweise nach Umformung

$$\frac{u(x) - u(x_0)}{x - x_0} = - \frac{\frac{\partial g}{\partial x}(\xi, g(x))}{\frac{\partial g}{\partial y}(x_0, \eta)}.$$

Mit  $x \rightarrow x_0$  geht  $\xi \rightarrow x_0$  sowie  $u(x), \eta \rightarrow u(x_0)$ , und die Behauptung folgt. □

Die Formel für  $u'(x_0)$  kann immer mit der Kettenregel wieder hergeleitet werden:

$$g(x, u(x)) = 0 \quad \Rightarrow \quad 0 = \frac{d}{dx} g(x, u(x)) = \frac{\partial g}{\partial x}(x, u(x)) + \frac{\partial g}{\partial y}(x, u(x))u'(x).$$

Durch Auflösen folgt dann für alle  $x \in (x_0 - \delta, x_0 + \delta)$  die Ableitung der impliziten Funktion

$$u'(x) = \frac{\frac{\partial g}{\partial x}(x, u(x))}{\frac{\partial g}{\partial y}(x, u(x))}. \quad (7.13)$$

Dies gilt tatsächlich auf dem ganzen Intervall, denn der Satz ist auch im Punkt  $(x, u(x))$  statt  $(x_0, y_0)$  anwendbar.

**Beispiel 7.2** Betrachte die Gleichung

$$g(x, y) = y^n - xy + 1 = 0 \quad \text{mit } n \in \mathbb{N} \text{ ungerade.}$$

Hierfür gibt es keine allgemeine Lösungsformel. Es gilt  $g(0, -1) = 0$ ,

$$\frac{\partial g}{\partial x}(0, -1) = 1 \quad \text{und} \quad \frac{\partial g}{\partial y}(0, -1) = n.$$

Nach dem impliziten Funktionensatz gibt es für  $x$  nahe bei Null eine eindeutige Lösung  $y = u(x)$  nahe bei  $-1$ , und es gilt

$$u'(0) = -\frac{\frac{\partial g}{\partial x}(0, -1)}{\frac{\partial g}{\partial y}(0, -1)} = -\frac{1}{n}.$$

Wir haben also die Näherung  $u(x) \approx -1 - \frac{x}{n}$  für die Nullstelle.

Ein Beweis der höherdimensionalen Version des Satzes über implizite Funktionen würde hier zu weit führen, aber wir wollen diese zumindest formulieren. Im Anschluss werden wir eine geometrische Interpretation des Satzes geben. Betrachte also die  $k$  Gleichungen

$$g(x, y) = 0 \quad \text{mit } (x, y) \in D \subset \mathbb{R}^m \times \mathbb{R}^k \text{ und } g : D \rightarrow \mathbb{R}^k.$$

Wir wollen nach  $y$  auflösen und setzen voraus, dass die Zahl der Gleichungen mit der Zahl der Unbekannten übereinstimmt, was ja bestimmt sinnvoll ist. Bezeichne mit  $D_x g(x_0, y_0) \in \mathbb{R}^{k \times m}$  beziehungsweise  $D_y g(x_0, y_0) \in \mathbb{R}^{k \times k}$  die Jacobimatrizen bezüglich der Variablen  $x \in \mathbb{R}^m$  bzw.  $y \in \mathbb{R}^k$ . Es sei eine Lösung  $(x_0, y_0)$  der Gleichung gegeben. Entwickeln wir  $f(x, y)$  bei  $(x_0, y_0)$  und ignorieren Terme höherer Ordnung, so erhalten wir das lineare Gleichungssystem

$$0 = g(x, y) \approx A(x - x_0) + B(y - y_0) \quad \text{mit } A = D_x g(x_0, y_0), B = D_y g(x_0, y_0).$$

Dieses lineare System ist eindeutig nach  $y$  auflösbar genau wenn  $B = D_y f(x_0, y_0) \in \mathbb{R}^{k \times k}$  invertierbar ist, und zwar lautet die Lösung

$$y = y_0 - B^{-1}A(x - x_0).$$

**Satz 7.2 (implizite Funktionen)** Sei  $g : D \rightarrow \mathbb{R}^k$  stetig differenzierbar auf der offenen Menge  $D \subset \mathbb{R}^m \times \mathbb{R}^k$ , und sei  $(x_0, y_0) \in D$  mit  $g(x_0, y_0) = 0$ . Ist  $D_y g(x_0, y_0)$  invertierbar, so gibt es Umgebungen  $U \subset \mathbb{R}^m$  von  $x_0$  und  $V \subset \mathbb{R}^k$  von  $y_0$  mit folgender Eigenschaft:

- (1) Zu jedem  $x \in U$  gibt es genau ein  $y \in V$  mit  $g(x, y) = 0$ , damit ist die implizite Funktion  $y = u(x)$  definiert.

(2)  $u : U \rightarrow V$  ist stetig differenzierbar mit Ableitung

$$Du(x) = -D_y g(x, u(x))^{-1} D_x g(x, u(x)).$$

Bisher war die Aufteilung in Unbekannte  $y$  und Parameter  $x$  a priori vorgegeben. Das ist nicht immer so, zum Beispiel sind Höhenlinien einfach Nullstellenmengen

$$\Gamma = \{(x, y) \in D : g(x, y) = 0\},$$

die beiden Koordinaten  $(x, y)$  sind völlig gleichberechtigt. Wir können uns dann aussuchen, ob wir nach  $x$  oder nach  $y$  auflösen und die Höhenlinie entweder als Graph  $y = u(x)$  über der  $x$ -Achse oder als Graph  $x = u(y)$  über der  $y$ -Achse schreiben, natürlich im allgemeinen mit einer anderen Funktion  $u$ . Entsprechendes gilt für eine Höhenfläche

$$\Sigma = \{(x, y, z) \in D : g(x, y, z) = 0\}.$$

In der Nähe eines Punkts  $(x_0, y_0, z_0) \in \Sigma$  kann eine Darstellung als Graph  $z = u(x, y)$  existieren, aber auch eine Darstellung über der  $(x, z)$ -Ebene oder der  $(y, z)$ -Ebene. Der Satz über implizite Funktionen liefert lokal diese Darstellungen, wenn die jeweilige Auflösungsbedingung erfüllt ist.

**Beispiel 7.3** Betrachte die zwei-dimensionale Sphäre

$$\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\} \quad \text{mit } g(x, y, z) = x^2 + y^2 + z^2 - 1.$$

Es gilt  $\frac{\partial g}{\partial z}(x_0, y_0, z_0) = 2z_0 \neq 0$  für  $z_0 \neq 0$ . Nach dem Satz über implizite Funktionen kann  $\mathbb{S}^2$  in der Nähe jedes Punkts  $(x_0, y_0, z_0) \in \mathbb{S}^2$  als Graph  $z = u(x, y)$  geschrieben werden, es sei denn  $(x_0, y_0, z_0)$  liegt auf dem Äquator ( $z_0 = 0$ ). Tatsächlich haben wir die Darstellungen

$$z = \sqrt{1 - x^2 - y^2} \quad \text{für } z > 0 \quad \text{bzw.} \quad z = -\sqrt{1 - x^2 - y^2} \quad \text{für } z < 0.$$

In der Umgebung eines Punkts  $(x_0, y_0, 0) \in \mathbb{S}^2$  kann die Sphäre tatsächlich nicht als Graph  $z = u(x, y)$  geschrieben werden. Aber man hat eine lokale Graphendarstellung  $x = u(y, z)$ , außer wenn  $x_0 = 0$ , und  $y = u(x, z)$  wenn nicht  $y_0 = 0$ . Da mindestens eine der drei Koordinaten  $x_0, y_0, z_0$  ungleich Null sein muss wegen  $x_0^2 + y_0^2 + z_0^2 = 1$ , hat man eine lokale Darstellung als Graph für jeden Punkt auf  $\mathbb{S}^2$ .

Wir kommen nun zu Extremwertaufgaben mit einer Nebenbedingung. Folgendes Beispiel ist uns schon begegnet.

**Beispiel 7.4** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrische Matrix mit der quadratischen Form

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, f(x) = \langle Ax, x \rangle = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Wir haben die Funktion  $f(x)$  minimiert unter der Nebenbedingung  $g(x) = |x|^2 = 1$ , also unter Punkten auf der Sphäre  $\mathbb{S}^{n-1}$ .

Im folgenden Satz definiert die Funktion  $g(x)$  bzw. die Menge  $M$  die Nebenbedingung, und  $f(x)$  ist die Zielfunktion, deren Extremwert gesucht wird.

**Satz 7.3 (Lagrange-Multiplikatoren-Regel)** Seien  $f, g \in C^1(\Omega)$  mit  $\Omega \subset \mathbb{R}^n$  offen. Sei  $M = \{x \in \Omega : g(x) = 0\}$  und  $a \in M$  mit

$$f(a) \leq f(x) \quad \text{für alle } x \in M.$$

Ist dann  $\text{grad } g(a) \neq 0$ , so gibt es ein  $\lambda \in \mathbb{R}$  mit  $\text{grad } f(a) = \lambda \text{grad } g(a)$ .

BEWEIS: Sei  $\gamma(t)$  eine Kurve in  $M$  mit  $\gamma(0) = a$ . Dann hat die Funktion  $f(\gamma(t))$  ein Minimum bei  $t = 0$ , also folgt mit  $V = \gamma'(0)$

$$0 = \frac{d}{dt} f(\gamma(t))|_{t=0} = \langle \text{grad } f(\gamma(0)), \gamma'(0) \rangle = \langle \text{grad } f(a), V \rangle.$$

Die Menge aller Ableitungsvektoren  $V \in \mathbb{R}^n$  solcher Kurven  $\gamma(t)$  nennen wir den Tangentialraum  $T_a M$  von  $M$  im Punkt  $a$ . Da  $g(x)$  sogar identisch Null auf  $M$  ist, folgt

$$\text{grad } f(a), \text{grad } g(a) \perp T_a M.$$

Wir zeigen, dass es  $n - 1$  linear unabhängige Vektoren in  $T_a M$  gibt. Dann sind die beiden Gradienten parallel, und wegen  $\text{grad } g(a) \neq 0$  folgt die Behauptung. Durch Umm Nummerieren der Koordinaten können wir dazu annehmen, dass

$$\frac{\partial g}{\partial x_n}(a) \neq 0.$$

Nach dem Satz über implizite Funktionen ist  $M$  in einer Umgebung von  $a$  als Graph einer  $C^1$ -Funktion  $x_n = u(x')$  darstellbar, wobei  $x' = (x_1, \dots, x_{n-1})$ ; speziell gilt  $a = (a', u(a'))$  wegen  $a \in M$ . Wir wählen nun  $\gamma_i(t) = (a' + te_i, u(a' + te_i)) \in M$  für  $i = 1, \dots, n - 1$ . Es folgt  $\gamma_i(0) = a$  und somit

$$T_a M \ni \frac{d}{dt} \gamma_i(t)|_{t=0} = (e_i, \partial_i u(a')).$$

Die Vektoren sind tatsächlich linear unabhängig, denn es gilt

$$\sum_{i=1}^{n-1} \lambda_i (e_i, \partial_i u(a')) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n-1} \lambda_i e_i = 0 \quad \Rightarrow \quad \lambda_i = 0.$$

□

Für eine beliebige Menge  $M \subset \mathbb{R}^n$  ist  $T_a M$  kein Unterraum. Hier ist aber  $M$  die Nullstellenmenge von  $g(x)$  und  $\text{grad } g(a) \neq 0$ . Der implizite Funktionensatz liefert dann, dass  $T_a M$  ein  $(n - 1)$ -dimensionaler Unterraum ist.  $(n - 1)$ -dimensionaler Unterraum ist. Denn betrachten wir allgemeiner die Kurven  $\gamma(t) = (a' + tv, u(a' + tv))$  mit beliebigem  $v \in \mathbb{R}^{n-1}$ , so folgt

$$\{(v, \partial_v u(a')) : v \in \mathbb{R}^{n-1}\} \subset T_a M.$$

Die Menge links ist aber ein  $(n - 1)$ -dimensionaler Unterraum. Andererseits kann es wegen  $T_a M \perp \text{grad } g(a)$  keine weiteren Tangentialvektoren geben.

**Beispiel 7.5** Für eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  betrachten wir nochmals das Problem, die quadratische Form  $f(x) = \langle Ax, x \rangle$  zu minimieren unter der Nebenbedingung  $g(x) = |x|^2 - 1 = 0$ . Wie in Beispiel 1.6 bewiesen, hat das Minimumproblem eine Lösung, das heißt es gibt ein  $x_0 \in \mathbb{S}^{n-1}$  mit  $f(x_0)$  kleinstmöglich. Da  $\text{grad } g(x_0) = 2x_0 \neq 0$  wegen  $|x_0| = 1$ , gibt es nach Satz 7.3 ein  $\lambda \in \mathbb{R}$  mit  $\text{grad } f(x_0) = \lambda \text{grad } g(x_0)$ . Aber  $\text{grad } f(x_0) = 2Ax_0$  nach Beispiel 3.4, also erhalten wir  $Ax_0 = \lambda x_0$ . Dies hatten wir schon in Kapitel 5, Satz 7.3, gesehen.

In konkreten Anwendungen kann es vorkommen, dass die Bedingung  $\text{grad } g(x) \neq 0$  nicht allen Punkten von  $M = \{x \in \Omega : g(x) = 0\}$  erfüllt ist; diese Punkte müssen dann gesondert betrachtet werden.

# Kapitel 7

## Mehrdimensionale Integration

### 1 Prinzip von Cavalieri und Fubini

In diesem Abschnitt erklären wir kurz die Definition des mehrdimensionalen Riemannsches Integrals. Als erstes definieren wir das Integral stetiger Funktionen auf  $n$ -dimensionalen Quadern  $Q$ , und zwar per Induktion nach der Dimension  $n$ . Wir bezeichnen das Volumen von  $Q$ , also das Produkt der Kantenlängen, mit  $|Q|$ . Für  $n = 1$  haben wir das übliche Riemannsches Integral auf einem Intervall  $Q = [a, b]$ .

Für  $n = 2$  ist  $Q = [a_1, b_1] \times [a_2, b_2]$  ein Rechteck, und wir setzen für eine stetige Funktion  $f = f(x, y)$  auf diesem Rechteck

$$\int_Q f(x, y) \, dx dy = \int_{a_2}^{b_2} \left( \int_{a_1}^{b_1} f(x, y) \, dx \right) dy.$$

Das Parameterintegral  $F(y) = \int_{a_1}^{b_1} f(x, y) \, dx$  ist nach Satz 5.2 in Kapitel 5 eine stetige Funktion, und damit ist das iterierte Integral auf der rechten Seite tatsächlich definiert. Für  $n = 3$ , also  $Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$  und eine stetige Funktion  $f = f(x, y, z)$ , setzen wir

$$\int_Q f(x, y, z) \, dx dy dz = \int_{a_3}^{b_3} \left( \int_{Q'} f(x, y, z) \, dx dy \right) dz$$

Dabei ist  $Q' = [a_1, b_1] \times [a_2, b_2]$ . Das innere Integral  $F_1(z) = \int_{Q'} f(x, y, z) \, dx dy$  ist bereits definiert, allerdings müssen wir noch zeigen, dass es stetig von  $z$  abhängt. Es gilt

$$F_1(z) = \int_{a_2}^{b_2} F_2(y, z) \, dy \quad \text{mit} \quad F_2(y, z) = \int_{a_1}^{b_1} f(x, y, z) \, dx.$$

Nach Satz 5.2 in Kapitel 5 ist  $F_2(y, z)$  stetig in beiden Variablen, und dann  $F_1(z)$  ebenfalls stetig. Damit ist die Definition auch für  $n = 3$  gerechtfertigt. Allgemein definieren wir für  $n$  beliebig, also  $Q = Q' \times [a_n, b_n]$  und  $f = f(x', x_n)$  stetig mit  $x' \in Q' \subset \mathbb{R}^{n-1}$ ,  $x_n \in [a_n, b_n]$ ,

$$\int_Q f(x) \, dx = \int_{a_n}^{b_n} \left( \int_{Q'} f(x', x_n) \, dx' \right) dx_n.$$

Per Induktion sieht man, dass das innere Integral stetig von  $x_n$  abhängt und somit das Integral wohldefiniert ist. Wir schreiben ein  $n$ -dimensionales Integral meistens in der Form

$\int_Q f(x) dx$ , wobei  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  für alle Koordinaten steht. In einem spezifischen Integral in  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  schreiben wir auch  $\int_Q f(x, y) dx dy$  oder  $\int_Q f(x, y, z) dx dy dz$ , wobei nun  $x, y, z$  die einzelnen Koordinaten bezeichnet.

Es stellt sich die offensichtliche Frage, ob diese Definition von der Reihenfolge abhängt, in der die einzelnen Variablen integriert werden. Wir zeigen deshalb jetzt eine alternative Darstellung des Integrals als Grenzwert Riemannscher Summen. In diesen Riemannschen Summen können die Koordinaten vertauscht werden, woraus dann die Behauptung folgt. Der Einfachheit halber beschränken wir uns auf Zerlegungen, bei denen jedes Intervall  $[a_i, b_i]$  äquidistant durch  $N$  Unterteilungspunkte zerlegt wird. Außerdem nehmen wir als Stützstellen die Unterteilungspunkte.

**Satz 1.1 (Approximation durch Riemannsche Summen)** *Für jeden  $n$ -dimensionalen Quader  $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$  und  $f : Q \rightarrow \mathbb{R}$  stetig gilt*

$$\int_Q f(x) dx = \lim_{N \rightarrow \infty} \sum_{k_1, \dots, k_n=1}^N f(x_{1,k_1}, \dots, x_{n,k_n}) \Delta x_1 \dots \Delta x_n.$$

Dabei ist  $\Delta x_i = \frac{b_i - a_i}{N}$  und  $x_{i,k} = a_i + k \Delta x_i$ .

BEWEIS: Wir geben den Beweis nur für  $n = 2$ . Für  $n \geq 3$  ist das Argument ganz analog, nur ist der Notationsaufwand wegen der vielen Indizes größer. Wir haben also die Zerlegung

$$\begin{aligned} a_1 = x_0 \leq \dots \leq x_N = b_1 & \quad \text{mit} \quad x_k = a_1 + k \Delta x \\ a_2 = y_0 \leq \dots \leq y_N = b_2 & \quad \text{mit} \quad y_\ell = a_2 + \ell \Delta y, \end{aligned}$$

wobei  $\Delta x = \frac{b_1 - a_1}{N}$ ,  $\Delta y = \frac{b_2 - a_2}{N}$ . Es folgt, wenn  $S_Z(f)$  die obige Riemannsche Summe ist,

$$\begin{aligned} \int_Q f(x, y) dx dy &= \sum_{k, \ell=1}^N \int_{y_{\ell-1}}^{y_\ell} \left( \int_{x_{k-1}}^{x_k} f(x, y) dx \right) dy \\ S_Z(f) &= \sum_{k, \ell=1}^N f(x_k, y_\ell) \Delta x \Delta y = \int_{y_{\ell-1}}^{y_\ell} \left( \int_{x_{k-1}}^{x_k} f(x_k, y_\ell) dx \right) dy. \end{aligned}$$

Wir erinnern nun an den Begriff der Oszillation einer Funktion:

$$\text{osc}(f, \delta) = \sup_{|(x,y)-(x',y')| < \delta} |f(x, y) - f(x', y')|.$$

Für  $N$  groß ist der Durchmesser von  $[x_{k-1}, x_k] \times [y_{\ell-1}, y_\ell]$  kleiner als  $\delta$ , also gilt

$$|f(x, y) - f(x_k, y_\ell)| \leq \text{osc}(f, \delta) \quad \text{für } (x, y) \in [x_{k-1}, x_k] \times [y_{\ell-1}, y_\ell].$$

Mit der Dreiecksungleichung ergibt sich

$$\begin{aligned} \left| \int_Q f(x, y) dx dy - S_Z(f) \right| &\leq \sum_{k, \ell=1}^N \int_{x_{k-1}}^{x_k} \left( \int_{y_{\ell-1}}^{y_\ell} |f(x, y) - f(x_k, y_\ell)| dy \right) dx \\ &\leq \text{osc}(f, \delta) \sum_{k, \ell=1}^N \Delta x \Delta y \\ &\leq \text{osc}(f, \delta) |Q|. \end{aligned}$$

Aber  $\text{osc}(f, \delta) \rightarrow 0$  mit  $\delta \searrow 0$ , siehe Satz 5.1 in Kapitel 5. Damit ist der Satz bewiesen.  $\square$   
 Die Vertauschbarkeit der Integrationsreihenfolge ist direkte Konsequenz. Insbesondere gilt

**Folgerung 1.1 (Fubini)** Sei  $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$  und  $f : Q \rightarrow \mathbb{R}$  stetig. Dann kann das Integral für jedes  $1 \leq i \leq n$  wie folgt berechnet werden:

$$\int_Q f(x) dx = \int_{a_i}^{b_i} \left( \int_{Q'} f(x) dx' \right) dx_i.$$

Dabei ist  $Q' = [a_1, b_1] \times \dots \times \widehat{[a_i, b_i]} \times \dots \times [a_n, b_n]$  und  $dx' = dx_1 \dots \widehat{dx_i} \dots dx_n$ . Das Dach bedeutet, dass dieser Faktor auszulassen ist.

Wir notieren einige Eigenschaften, die direkt aus den Regeln für das eindimensionale Riemannintegral - oder alternativ aus der Darstellung als Grenzwert Riemannscher Summen - folgen.

**Satz 1.2 (Eigenschaften des Integrals)** Auf einem  $n$ -dimensionalen Quader  $Q$  gilt:

- (1) Linearität:  $\int_Q (\lambda f(x) + \mu g(x)) dx = \lambda \int_Q f(x) dx + \mu \int_Q g(x) dx$ .
- (2) Integral einer Konstante: Die konstante Funktion  $f : Q \rightarrow \mathbb{R}$  mit Wert  $f(x) = c$  hat das Integral  $\int_Q f(x) dx = c |Q|$ .
- (3) Monotonie: Aus  $f(x) \leq g(x)$  folgt  $\int_Q f(x) dx \leq \int_Q g(x) dx$ .
- (4) Abschätzung:  $\left| \int_Q f(x) dx \right| \leq \int_Q |f(x)| dx \leq (\sup_Q |f|) |Q|$ .

Nun kommen wir zu allgemeineren Mengen als nur Rechtecken oder Quadern. Eine beliebige Menge im  $\mathbb{R}^n$  approximieren wir von innen und außen durch Vereinigungen von endlich vielen  $n$ -dimensionalen Würfeln. Sei dazu  $\mathcal{Q}$  die Menge aller abgeschlossenen Gitterwürfel  $Q$  mit ganzzahligen Koordinaten und Kantenlänge Eins, und sei  $\mathcal{Q}_k$  die Menge der skalierten Würfel  $2^{-k}Q$  für  $k \in \mathbb{N}_0$ . Setze

$$\begin{aligned} B_k &= \text{Vereinigung der Würfel in } \mathcal{Q}_k, \text{ die im Inneren von } B \text{ enthalten sind,} \\ B^k &= \text{Vereinigung der Würfel in } \mathcal{Q}_k, \text{ die den Abschluss von } B \text{ treffen.} \end{aligned}$$

Beim Übergang von  $k$  zu  $k+1$  kommen bei  $B_k$  Würfel hinzu, bei  $B^k$  fallen welche weg. Somit

$$B_0 \subset B_1 \subset \dots \subset \text{int } B \subset B \subset \overline{B} \subset \dots \subset B^1 \subset B^0.$$

Wegen Monotonie und Beschränktheit existieren somit das innere und äußere Volumen

$$|B|_* := \lim_{k \rightarrow \infty} |B_k| \leq \lim_{k \rightarrow \infty} |B^k| =: |B|^*.$$

**Definition 1.1** Eine beschränkte Menge  $B \subset \mathbb{R}^n$  heißt Jordan-messbar, wenn  $|B|_* = |B|$ . Diese Zahl ist dann das Volumen  $|B|$  von  $B$ .

Die Bedingung ist gleichbedeutend mit  $|\partial B|^* = 0$ . Denn ein Gitterwürfel  $Q \in \mathcal{Q}_k$  trifft genau dann  $\partial B$ , wenn er den Abschluss von  $B$  trifft, aber nicht im Inneren von  $B$  liegt (nachprüfen!). Demnach gilt  $(\partial B)^k = B^k \setminus B_k$  und somit

$$|\partial B|^* = |B|^* - |B|_*.$$

Wir können nun stetige Funktionen auf Jordan-messbaren Mengen integrieren.

**Definition 1.2 (Riemann-Integral II)** Sei  $B \subset \mathbb{R}^n$  Jordan-messbar. Für  $f : B \rightarrow \mathbb{R}$  beschränkt und stetig mit  $f \geq 0$  definieren wir

$$\int_B f(x) dx = \lim_{k \rightarrow \infty} \int_{B_k} f(x) dx.$$

In einem zweiten Schritt setzen wir für eine beliebige beschränkte, stetige Funktion  $f : B \rightarrow \mathbb{R}$

$$\int_B f(x) dx = \int_B f^+(x) dx - \int_B f^-(x) dx.$$

Die Menge  $B_k$  ist die Vereinigung der Gitterwürfel, die im Inneren von  $B$  enthalten sind. Das Integral über  $B_k$  ist damit definiert als Summe der Integrale auf den einzelnen Gitterwürfeln; auf diesen ist die Funktion  $f$  nach Voraussetzung stetig. Beim Übergang von  $B_k$  zu  $B_{k+1}$  kommen Würfel hinzu, damit ist die Folge  $\int_{B_k} f(x) dx$  monoton wachsend für  $f \geq 0$ . Andererseits ist sie nach oben beschränkt durch

$$\int_{B_k} f(x) dx \leq (\sup_B f) |B_k| \leq (\sup_B f) |B|_* < \infty.$$

Wegen Monotonie und Beschränktheit existiert also der Grenzwert der Integrale. Um zu zeigen, dass die Definition sinnvoll ist, müssen wir außerdem nachweisen, dass sie im Fall eines Quaders  $B$  mit unserer ersten Integraldefinition konsistent ist. In diesem Fall sind die  $B_k$  Quader, die in  $B$  enthalten sind. Man kann  $B \setminus B_k$  dann in Quader zerlegen, deren Gesamtvolumen gegen Null geht. Da  $f$  beschränkt ist, gehen die Differenzintegrale gegen Null wie behauptet.

Kurz zum zweiten Teil der Definition: es ist

$$f^+(x) = \begin{cases} f(x) & \text{falls } f(x) \geq 0, \\ 0 & \text{falls } f(x) \leq 0 \end{cases} \quad f^-(x) = \begin{cases} -f(x) & \text{falls } f(x) \leq 0, \\ 0 & \text{falls } f(x) \geq 0 \end{cases}.$$

Insbesondere sind  $f^\pm$  stetig und es gilt  $f = f^+ - f^-$ . Die Definition ist also konsistent mit der Linearität des Integrals.

Wir formulieren nun eine zweite Variante des Satzes von Fubini.

**Satz 1.3 (Cavalieri)** Sei  $B \subset \{(x, z) \in \mathbb{R}^{n-1} \times \mathbb{R} : a \leq z \leq b\}$  Jordan-messbar, und sei  $f : B \rightarrow \mathbb{R}$  stetig. Dann gilt

$$\int_B f(x, z) dx dz = \int_a^b \left( \int_{B_z} f(x, z) dx \right) dz,$$

wobei  $B_z = \{x \in \mathbb{R}^{n-1} : (x, z) \in B\}$  den Querschnitt von  $B$  in Höhe  $z \in [a, b]$  bezeichnet.

**Beispiel 1.1** Der Flächeninhalt  $|D|$  der Kreisscheibe

$$D = \{(x, z) \in \mathbb{R} \times \mathbb{R} : z \in [-1, 1], x^2 + z^2 \leq 1\}$$

ergibt sich wie folgt: es ist  $D_z = [-\sqrt{1-z^2}, \sqrt{1-z^2}]$ , also gilt

$$|D| = \int_{-1}^1 \mathcal{L}(D_z) dz = 2 \int_{-1}^1 \sqrt{1-z^2} dz = 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t dt = \pi.$$

Hier haben wir  $z = \sin t$  substituiert. Analog berechnen wir das Volumen  $V(B)$  der Kugel  $B = \{(x, z) \in \mathbb{R}^2 \times \mathbb{R} : |x|^2 + z^2 \leq 1\}$ . Es ergibt sich

$$|B| = \int_{-1}^1 \mathcal{A}(B_z) dz = \pi \int_{-1}^1 (1 - z^2) dz = \frac{4\pi}{3}.$$

BEWEIS: (*des Cavalierischen Prinzips*) Wir können  $f \geq 0$  annehmen, der allgemeine Fall folgt dann durch Zerlegung  $f = f^+ - f^-$ . Da  $B_k$  Vereinigung endlich vieler Gitterwürfel ist und auf jedem einzelnen Gitterwürfel das Integral genau als iteriertes Integral definiert wurde, erhalten wir jedenfalls

$$\int_{B_k} f(x, z) dx dz = \int_a^b \left( \int_{(B_k)_z} f(x, z) dx \right) dz.$$

Man kann nun zeigen, dass die rechte Seite mit  $k \rightarrow \infty$  konvergiert (für die linke Seite ist das klar), so dass folgt

$$\int_B f(x, z) dx dz = \int_a^b \left( \int_{B_z} f(x, z) dx \right) dz.$$

□

Eine Anwendung des Satzes von Fubini bzw. Cavalieri ist die partielle Integration.

**Lemma 1.1** Sei  $B \subset \mathbb{R}^n$  offen und beschränkt, und  $f \in C^1(B)$  mit  $f = 0$  außerhalb einer abgeschlossenen Menge  $K \subset B$ . Dann gilt

$$\int_B \frac{\partial f}{\partial x_i} dx = 0 \quad \text{für } i = 1, \dots, n.$$

BEWEIS: Wir wählen einen Quader  $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$ , so dass  $B$  in  $Q$  enthalten ist, und setzen  $f(x)$  durch Null fort:

$$f_0 : Q \rightarrow \mathbb{R}, f_0(x) = \begin{cases} f(x) & \text{für } x \in B \\ 0 & \text{für } x \in Q \setminus B. \end{cases}$$

Es folgt  $f \in C^1(Q)$  mit  $\frac{\partial f}{\partial x_i} = 0$  auf  $Q \setminus B$ , und  $f = 0$  auf  $\partial Q$ . Mit dem Satz von Fubini und dem Hauptsatz der Differential- und Integralrechnung folgt, wenn wir  $dx' = dx_1 \dots \widehat{dx_i} \dots dx_n$  und  $Q' = [a_1, b_1] \times \dots \times \widehat{[a_i, b_i]} \times \dots \times [a_n, b_n]$  schreiben,

$$\int_B \frac{\partial f}{\partial x_i} dx = \int_Q \frac{\partial f}{\partial x_i} dx = \int_{Q'} \left( \int_{a_i}^{b_i} \frac{\partial f}{\partial x_i}(x) dx_i \right) dx' = \int_{Q'} \underbrace{[f(x)]_{x_i=a_i}^{x_i=b_i}}_{=0} dx' = 0.$$

□

**Satz 1.4 (Partielle Integration)** Sei  $B \subset \mathbb{R}^n$  offen und beschränkt. Für  $f \in C^1(B)$  und  $X \in C^1(B, \mathbb{R}^n)$  sei  $fX = 0$  außerhalb einer abgeschlossenen Menge  $K \subset B$ . Dann gilt

$$\int_B \langle \text{grad } f, X \rangle dx = - \int_B f \text{div } X dx.$$

BEWEIS: Die Funktion  $fX_i$  erfüllt die Voraussetzung von Lemma 1.1, also gilt

$$0 = \int_B \frac{\partial(fX_i)}{\partial x_i} dx = \int_B \frac{\partial f}{\partial x_i} X_i dx + \int_B f \frac{\partial X_i}{\partial x_i} dx.$$

Summation über  $i = 1, \dots, n$  ergibt die Behauptung.

□

## 2 Koordinatentransformationen

Das vielleicht wichtigste Hilfsmittel zur Berechnung mehrdimensionaler Integrale ist der Transformationsatz. Wir erinnern an folgende Aussage, die wir bei der Diskussion der Determinante gezeigt haben.

**Satz 5.9 (Lineare Transformationsformel)** Für  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  gilt

$$V(A(B)) = |\det(A)| V(B) \quad \text{für alle } B \subset \mathbb{R}^n.$$

Ist  $B$  ein Würfel, so ist  $A(B)$  ein Parallelotop und die Formel folgt aus den Eigenschaften der Determinante durch Scherung. Allgemeine Mengen  $B$  haben wir dann durch Vereinigungen von Gitterwürfeln approximiert, genau wie oben dargestellt. Als Beispiel haben wir das Volumen eines Ellipsoids ausgerechnet, als Bild der Kugel unter einer Diagonalmatrix mit den Halbachsen als Diagonalelementen. Wir suchen nun eine Verallgemeinerung für Abbildungen, die nicht linear sind. Außerdem wollen wir nicht nur Volumina behandeln, sondern beliebige Integrale.

**Definition 2.1** Eine Abbildung  $\phi : U \rightarrow V$  zwischen offenen Mengen  $U, V \subset \mathbb{R}^n$  heißt *Koordinatentransformation (oder Diffeomorphismus)*, wenn  $\phi$  bijektiv ist und  $\phi, \phi^{-1}$  beide  $C^1$ -Abbildungen sind.

Aus  $\phi^{-1}(\phi(x)) = x$  folgt mit der Kettenregel

$$D\phi^{-1}(\phi(x))D\phi(x) = \text{Id} \quad \text{für alle } x \in U.$$

Also ist die Jacobideterminante einer Koordinatentransformation überall ungleich Null.

**Beispiel 2.1** Die Polarkoordinatenabbildung

$$\phi : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}, \quad \phi(r, \theta) = (r \cos \theta, r \sin \theta),$$

ist eine Koordinatentransformation.

**Satz 2.1 (Transformationsformel)** Sei  $\phi : U \rightarrow V$  eine Koordinatentransformation zwischen den offenen Mengen  $U, V \subset \mathbb{R}^n$ . Dann gilt für jede Jordan-messbare Menge  $B \subset U$  und jede stetige Funktion  $f : \phi(B) \rightarrow \mathbb{R}$  die Formel

$$\int_{\phi(B)} f(y) dy = \int_B f(\phi(x)) |\det D\phi(x)| dx.$$

BEWEIS: Wir können  $f \geq 0$  annehmen, sonst zerlegen wir  $f = f^+ - f^-$ . Tatsächlich können wir sogar  $f > 0$  annehmen, indem wir  $f(x)$  durch die Funktionen  $f(x) + \varepsilon$  approximieren. Wir zeigen nun, dass für jeden Würfel  $Q$  im Inneren von  $B$  gilt

$$\int_{\phi(Q)} f(y) dy \leq \int_Q f(\phi(x)) |\det D\phi(x)| dx. \quad (2.1)$$

Wäre das falsch, so gibt es einen Würfel  $Q_0$  und ein  $\theta > 1$  mit

$$\theta \int_{Q_0} f(\phi(x)) |\det D\phi(x)| dx \leq \int_{\phi(Q_0)} f(y) dy.$$

Durch sukzessive Zerlegung in  $2^n$  Teilwürfel finden wir  $Q_0 \supset Q_1 \supset Q_2 \supset \dots$  mit

$$\theta \int_{Q_k} f(\phi(x)) |\det D\phi(x)| dx \leq \int_{\phi(Q_k)} f(y) dy.$$

Da die  $Q_k$  geschachtelt sind und ihr Durchmesser gegen Null geht, gibt es genau ein  $x_0$ , das in allen  $Q_k$  liegt. Wir verwenden nun folgende asymptotischen Aussagen:

- $f(\phi(x)) |\det D\phi(x)| \approx f(\phi(x_0)) |\det D\phi(x_0)|$  (Stetigkeit von  $f \circ \phi$  und  $\det D\phi$ ).
- $f(y) \approx f(\phi(x_0))$  auf  $\phi(Q_k)$  (Stetigkeit von  $f$ ).
- $|\phi(Q_k)| \approx |\det D\phi(x_0)| |Q_k|$  (da  $\text{diam } Q_k \rightarrow 0$ , ist  $\phi(x) \approx \phi(x_0) + D\phi(x_0)(x - x_0)$ , und die Näherung folgt aus der linearen Transformationsformel).

Setzen wir diese asymptotischen Informationen ein, so ergibt sich asymptotisch für  $k \rightarrow \infty$

$$\theta f(\phi(x_0)) |\det D\phi(x_0)| |Q_k| \lesssim f(\phi(x_0)) |\det D\phi(x_0)| |Q_k|.$$

Da  $f(\phi(x_0)) |\det D\phi(x_0)| > 0$ , ergibt sich nach Division  $\theta \lesssim 1$ . Das ist ein Widerspruch zu unserer Annahme  $\theta > 1$ , womit (2.1) gezeigt ist. Wenden wir nun (2.1) auf die Gitterwürfel in  $B_k$  an, so bekommen wir durch Addition

$$\int_{\phi(B_k)} f(y) dy \leq \int_{B_k} f(\phi(x)) |\det D\phi(x)| dx.$$

Die  $B_k$  schöpfen  $B$  mit  $k \rightarrow \infty$  aus, also folgt

$$\int_{\phi(B)} f(y) dy \leq \int_B f(\phi(x)) |\det D\phi(x)| dx.$$

Schließlich sei  $\psi : V \rightarrow U$  die inverse Transformation. Wir wenden das Gezeigte an auf die Funktion  $g(x) = f(\phi(x)) |\det D\phi(x)|$  und bekommen

$$\begin{aligned} \int_B f(\phi(x)) |\det D\phi(x)| dx &= \int_{\psi(\phi(B))} g(x) dx \\ &\leq \int_{\phi(B)} g(\psi(y)) |\det D\psi(y)| dy \\ &= \int_{\phi(B)} f(y) \underbrace{|\det D\phi(\psi(y))| |\det D\psi(y)|}_{=1} dy \\ &= \int_{\phi(B)} f(y) dy. \end{aligned}$$

Damit ist die Transformationsformel gezeigt. □

**Beispiel 2.2 (Polarkoordinaten im  $\mathbb{R}^2$ )** Die Polarkoordinatentransformation ist

$$\phi : U = (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\} = V, \quad \phi(r, \theta) = (r \cos \theta, r \sin \theta).$$

Die Jacobimatrix und - Determinante sind

$$D\phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}, \det D\phi(r, \theta) = r > 0.$$

Also besagt der Transformationssatz für  $B \subset U$  und  $f = f(x, y)$

$$\int_{\phi(B)} f(x, y) dx dy = \int_B f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Ein interessanter Fall ist  $f(x, y) = e^{-(x^2+y^2)/2}$  und  $\phi(B) = V$ . Mit Fubini erhalten wir

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy &= \int_0^\infty \int_0^{2\pi} e^{-r^2/2} r dr d\theta \\ &= -2\pi \int_0^\infty \frac{d}{dr} e^{-r^2} dr \\ &= (-2\pi) \left[ e^{-r^2} \right]_{r=0}^{r=\infty} = 2\pi. \end{aligned}$$

Andererseits gilt, wieder mit Fubini,

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy = \int_{-\infty}^\infty e^{-x^2/2} \left( \int_{-\infty}^\infty e^{-y^2/2} dy \right) dx = \left( \int_{-\infty}^\infty e^{-x^2/2} dx \right)^2.$$

Also sehen wir

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

Hieraus ergibt sich die Normierung der Gaußschen Normalverteilung: wählt man die Dichtefunktion  $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ , so hat die Verteilungsfunktion  $F(x) = \int_{-\infty}^x f(y) dy$  Integral Eins. In diesem Beispiel haben wir die Transformationsformel für ein unbeschränktes Gebiet benutzt. Um das zu rechtfertigen, könnte man durch beschränkte Gebiete approximieren, zum Beispiel durch Kreisscheiben  $D_R(0)$ .

**Beispiel 2.3** Sei  $U = (0, \infty) \times (0, \pi) \times (0, 2\pi)$  und  $V = \mathbb{R}^3 \setminus \{(x, 0, z) : x \geq 0\}$ . Dreidimensionale Polarkoordinaten sind gegeben durch

$$\phi : U \rightarrow V, \phi(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta).$$

Die Jacobimatrix von  $\phi$  lautet

$$D\phi(r, \theta, \varphi) = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}.$$

Für die Jacobideterminante ergibt sich  $\det D\phi(r, \theta, \varphi) = r^2 \sin \theta$ . Für ein Gebiet  $\phi(B)$  mit  $B = [r_1, r_2] \times [\theta_1, \theta_2] \times [\varphi_1, \varphi_2]$  liefert der Transformationssatz und der Satz von Fubini

$$|\phi(B)| = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{\varphi_1}^{\varphi_2} r^2 \sin \theta d\varphi d\theta dr = \frac{r_2^3 - r_1^3}{3} (\cos \theta_1 - \cos \theta_2) (\varphi_2 - \varphi_1).$$

Wir kommen jetzt – endlich – zur Umrechnung von Differentialoperatoren auf krummlinige Koordinaten. Bevor wir das in Angriff nehmen, müssen wir allerdings erst überlegen, was eine solche Transformation bedeutet, insbesondere müssen wir uns über das Transformationsverhalten von skalaren und vektorwertigen Funktionen klar werden. Wir nehmen also an, dass wir eine  $C^1$ -Koordinatentransformation gegeben haben:

$$\phi : U \rightarrow V, y = \phi(x), \quad \text{wobei } U, V \text{ offene Teilmengen des } \mathbb{R}^n.$$

Für Funktionen ist die Sache relativ klar: ist zum Beispiel  $v = v(y)$  eine Temperaturverteilung oder Ladungsdichte, so muss man diese eben am Punkt  $\phi(x)$  auswerten, das heißt

$$\text{Transformation eines Skalars } v = v(y) \text{ ist } u(x) = v(\phi(x)).$$

Hier ist eine Hinweis angesagt: in den Anwendungen werden Funktion und transformierte Funktion oft mit demselben Buchstaben bezeichnet, zum Beispiel  $v(y)$  und  $v(r, \theta, \varphi)$ . Der Grund ist, dass es sich um dieselbe Beobachtungsgröße handelt, zum Beispiel die Temperatur, nur in verschiedenen Koordinaten. Leider kann diese Notation manchmal zur Verwirrung führen, wenn die Kettenregel angewandt wird. Für unsere Rechnung behalten wir die Unterscheidung zwischen  $u$  und  $v$  deshalb bei. Es ist weiter wichtig festzustellen, dass für Vektorgrößen ein anderes Transformationsgesetz gilt als für skalare Funktionen, nämlich

$$\text{Transformation eines Vektorfelds } Y = Y(y) \text{ ist } X(x) \text{ mit } D\phi(x)X(x) = Y(\phi(x)).$$

Betrachten wir als Beispiel die Abbildung, die einem Punkt  $x$  auf einem Stadtplan den realen Punkt  $\phi(x)$  zuordnet. Sei  $x_0$  der Punkt auf der Karte, der Ihrem aktuellen Standort entspricht. Angenommen Sie halten die Karte so, dass der Norden auf der Karte real nach Osten ausgerichtet ist. Dann ist  $\phi(x)$  eine Linksdrehung um  $90^\circ$ , gefolgt von einer Streckung mit dem Maßstab. Wollen wir die Windrichtung  $Y(\phi(x))$  in die Karte eintragen, so müssen wir den Vektor um  $90^\circ$  zurück drehen. Allgemein entspricht bei einer Koordinatentransformation dem Vektor  $X(x)$  der mit der Jacobimatrix „gedrehte“ Vektor  $D\phi(x)X(x)$  an der Stelle  $\phi(x)$ .

Für die Umrechnung von Gradient und Divergenz brauchen wir die Gramsche Matrix

$$g(x) = (g_{ij}(x)) \in \mathbb{R}^{n \times n}, \quad g_{ij}(x) = \left\langle \frac{\partial \phi}{\partial x_i}(x), \frac{\partial \phi}{\partial x_j}(x) \right\rangle.$$

Die  $g_{ij}$  sind die Koeffizienten des Euklidischen Skalarprodukts bezüglich der Basis  $\frac{\partial \phi}{\partial x_i} = D\phi \cdot e_i$ . Für Vektoren  $X, Y \in \mathbb{R}^n$  gilt

$$\langle D\phi \cdot X, D\phi \cdot Y \rangle = \sum_{i,j=1}^n \langle D\phi \cdot e_i, D\phi \cdot e_j \rangle X_i Y_j = \sum_{i,j=1}^n g_{ij} X_i Y_j. \quad (2.2)$$

Weiter haben wir

$$g_{ij}(x) = \langle D\phi(x)e_i, D\phi(x)e_j \rangle = \langle e_i, D\phi(x)^T D\phi(x)e_j \rangle = (D\phi(x)^T D\phi(x))_{ij}.$$

Insbesondere gilt

$$\det g(x) = \det (D\phi(x)^T D\phi(x)) = |\det D\phi(x)|^2 > 0.$$

Die Matrix  $g(x)$  ist somit invertierbar, und wir setzen  $g^{-1}(x) = (g^{ij}(x))$ .

**Lemma 2.1 (Transformation von Vektorfeldern)** Sei  $\phi \in C^1(U, V)$  eine Koordinatentransformation mit Gramscher Matrix  $g(x) = (g_{ij}(x))$ , und sei  $Y \circ \phi = D\phi \cdot X$ . Dann gilt

$$X = \sum_{i,j=1}^n g^{ij} \left\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \right\rangle e_j \quad (e_j \text{ Standardbasis}).$$

BEWEIS: Wir berechnen mit (2.2)

$$\left\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \right\rangle = \langle D\phi \cdot X, D\phi \cdot e_i \rangle = \sum_{k=1}^n g_{ki} X_k.$$

Multiplizieren wir mit  $g^{ij}$  und summieren über  $i$ , so folgt

$$\sum_{i=1}^n g^{ij} \left\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \right\rangle = \sum_{k=1}^n X_k \underbrace{\sum_{i=1}^n g_{ki} g^{ij}}_{=\delta_{jk}} = X_j.$$

□

**Satz 2.2 (Transformation von Differentialoperatoren)** Sei  $\phi \in C^1(U, V)$  eine Koordinatentransformation mit Gramscher Matrix  $(g_{ij}(x))$ . Dann gelten folgende Formeln:

(1) Ist  $v = u \circ \phi$ , so folgt  $(\text{grad } v) \circ \phi = D\phi \cdot \text{grad}_g u$  mit

$$\text{grad}_g u = \sum_{i,j=1}^n g^{ij} \frac{\partial u}{\partial x_i} e_j.$$

(2) Ist  $Y \circ \phi = D\phi \cdot X$ , so folgt  $(\text{div } Y) \circ \phi = \text{div}_g X$  mit

$$\text{div}_g X = \frac{1}{\sqrt{\det g}} \sum_{i=1}^n \frac{\partial}{\partial x_i} (\sqrt{\det g} X_i).$$

(3) Ist  $v = u \circ \phi$ , so folgt  $(\Delta v) \circ \phi = \Delta_g u$  mit

$$\Delta_g u = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} (\sqrt{\det g} g^{ij} \frac{\partial u}{\partial x_j}).$$

BEWEIS: Für die Formel (1) wähle  $Y = \text{grad } v$  in Lemma 2.1. Für das zugehörige Vektorfeld  $X$ , also  $(\text{grad } v) \circ \phi = D\phi \cdot X$ , folgt

$$X_j = \sum_{i=1}^n g^{ij} \left\langle (\text{grad } v) \circ \phi, D\phi \cdot e_i \right\rangle = \sum_{i=1}^n g^{ij} Dv \cdot (D\phi \cdot e_i) = \sum_{i=1}^n g^{ij} \frac{\partial u}{\partial x_i}.$$

Das zeigt (1). Zum Beweis von (2) verwenden wir partielle Integration. Die Divergenz geht dabei über in den Gradienten, dessen Umrechnung wir bereits kennen. Sei  $\varphi \in C^1(U)$  eine

beliebige Funktion mit  $\varphi = 0$  außerhalb einer abgeschlossenen und beschränkten Teilmenge von  $U$ . Mit  $\psi = \varphi \circ \phi^{-1}$  berechnen wir

$$\begin{aligned}
\int_U \varphi(\operatorname{div} Y) \circ \phi \sqrt{\det g} \, dx &= \int_V \psi \operatorname{div} Y \, dy \quad (\text{Transformationsatz}) \\
&= - \int_V \langle \operatorname{grad} \psi, Y \rangle \, dy \quad (\text{partielle Integration}) \\
&= - \int_U \langle (\operatorname{grad} \psi) \circ \phi, Y \circ \phi \rangle \sqrt{\det g} \, dx \quad (\text{Transformationsatz}) \\
&= - \int_U \langle D\phi \cdot \operatorname{grad}_g \varphi, D\phi \cdot X \rangle \sqrt{\det g} \, dx \quad (\text{Umrechnung Gradient, } Y) \\
&= - \int_U \sum_{j,k=1}^n g_{jk} (\operatorname{grad}_g \varphi)_j X_k \sqrt{\det g} \, dx \quad \text{Gleichung (2.2)} \\
&= - \int_U \sum_{i,k=1}^n \sum_{j=1}^n g_{jk} g^{ij} \frac{\partial \varphi}{\partial x_i} X_k \sqrt{\det g} \, dx \quad (\text{Formel (1)}) \\
&= - \int_U \left( \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i} X_i \right) \sqrt{\det g} \, dx \quad (\sum_{j=1}^n g_{jk} g^{ij} = \delta_{ik}) \\
&= \int_U \varphi \sum_{i=1}^n \frac{\partial}{\partial x_i} (\sqrt{\det g} X_i) \, dx \quad (\text{partielle Integration}).
\end{aligned}$$

Da diese Gleichung für alle  $\varphi$  gilt, folgt mit dem üblichen Schluss

$$(\operatorname{div} Y) \circ \phi \sqrt{\det g} = \sum_{i=1}^n \frac{\partial}{\partial x_i} (\sqrt{\det g} X_i).$$

Damit ist (2) bewiesen. Formel (3) ergibt sich direkt aus (1) und (2), es gilt

$$(\Delta v) \circ \phi = (\operatorname{div} \operatorname{grad} v) \circ \phi = \operatorname{div}_g \operatorname{grad}_g u.$$

□

**Beispiel 2.4** Für Polarkoordinaten im  $\mathbb{R}^3$  berechnet man

$$g(r, \theta, \varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}.$$

Damit ergeben sich die Operatoren wie folgt:

$$\begin{aligned}
\operatorname{grad}_g u &= \frac{\partial u}{\partial r} e_1 + \frac{1}{r^2} \frac{\partial u}{\partial \theta} e_2 + \frac{1}{r^2 \sin^2 \theta} \frac{\partial u}{\partial \varphi} e_3 \\
\operatorname{div}_g X &= \frac{1}{r^2 \sin \theta} \left( \frac{\partial}{\partial r} (r^2 \sin \theta X_1) + \frac{\partial}{\partial \theta} (r^2 \sin \theta X_2) + \frac{\partial}{\partial \varphi} (r^2 \sin \theta X_3) \right) \\
&= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 X_1) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta X_2) + \frac{\partial}{\partial \varphi} X_3 \\
\Delta_g u &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \varphi^2}.
\end{aligned}$$

Wir können  $\Delta_g$  in der Form  $\Delta_g u = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \frac{\partial u}{\partial r}) + \frac{1}{r^2} \Delta_{\mathbb{S}^2} u$  schreiben, wobei

$$\Delta_{\mathbb{S}^2} u = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial u}{\partial \theta}) + \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \varphi^2}.$$

Die Formel (1) liefert die Entwicklung des Gradienten  $(\text{grad } v) \circ \phi$  bezüglich der Basis  $\frac{\partial \phi}{\partial x_i}$ :

$$(\text{grad } v) \circ \phi = \sum_{i,j=1}^n g^{ij} \frac{\partial u}{\partial x_i} \frac{\partial \phi}{\partial x_j} \quad \text{mit } u = v \circ \phi.$$

Die  $\frac{\partial \phi}{\partial x_i}$  bilden im allgemeinen keine Orthonormalbasis, ihre Skalarprodukte sind ja genau die Gramsche Matrix  $g_{ij}$ . Im Fall der Polarkoordinaten sind die  $\frac{\partial \phi}{\partial x_i}$  immerhin orthogonal. Es gilt

$$(\text{grad } v) \circ \phi = \frac{\partial u}{\partial r} \frac{\partial \phi}{\partial r} + \frac{1}{r^2} \frac{\partial u}{\partial \theta} \frac{\partial \phi}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial u}{\partial \varphi} \frac{\partial \phi}{\partial \varphi}.$$

Für orthogonale Koordinaten kann man eine Orthonormalbasis im Bild herstellen, indem man die Vektoren  $\frac{\partial \phi}{\partial x_i}$  normiert. Im Fall der Polarkoordinaten wird diese Basis oft mit  $\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi$  bezeichnet, die Darstellung des Gradienten lautet dann

$$(\text{grad } v) \circ \phi = \frac{\partial u}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial u}{\partial \theta} \vec{e}_\theta + \frac{1}{r \sin \theta} \frac{\partial u}{\partial \varphi} \vec{e}_\varphi.$$

Der Ansatz in Satz 2.2 ist aber aus meiner Sicht systematischer, und er ist auch nicht auf orthogonale Koordinaten beschränkt.

### 3 Der Satz von Gauß

In diesem Abschnitt diskutieren wir den Integralsatz von Gauß (Englisch: *divergence theorem*), der eine mehrdimensionale Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung ist. Dem Charakter nach handelt es sich um ein Erhaltungsgesetz, deshalb ist der Satz ein wichtiges mathematisches Hilfsmittel zum Beispiel in der Strömungslehre und der Elektrodynamik.

**Satz 3.1 (Integralsatz von Gauß)** Sei  $\Omega \subset \mathbb{R}^n$  ein beschränktes Gebiet mit glattem Rand. Bezeichne mit  $\nu(x)$ ,  $x \in \partial\Omega$ , die nach außen weisende Einheitsnormale am Rand. Ist  $X : \bar{\Omega} \rightarrow \mathbb{R}^n$  Vektorfeld mit stetigen Ableitungen bis zum Rand, so gilt

$$\int_{\Omega} \text{div } X \, dx = \int_{\partial\Omega} \langle X, \nu \rangle \, dA.$$

BEWEIS: Wir geben einen Beweis im Spezialfall dass das zugrundeliegende Gebiet ein  $n$ -dimensionaler Quader  $Q$  ist, das heißt  $Q = (a_1, b_1) \times \dots \times (a_n, b_n)$ . Die äußere Einheitsnormale ist dann wie folgt:

$$\nu(x) = \begin{cases} -e_i & \text{für } x \in \partial Q \text{ mit } x_i = a_i \\ e_i & \text{für } x \in \partial Q \text{ mit } x_i = b_i. \end{cases}$$

Wir setzen  $Q_i = (a_1, b_1) \times \dots \times \widehat{(a_i, b_i)} \times \dots \times (a_n, b_n) \subset \mathbb{R}^{n-1}$ , wobei das Dach bedeutet, dass der Faktor wegzulassen ist. Mit dem Satz von Fubini und dem Hauptsatz der Differential- und Integralrechnung folgt

$$\begin{aligned}
\int_Q \operatorname{div} X \, dx &= \sum_{i=1}^n \int_Q \frac{\partial X_i}{\partial x_i} \, dx \\
&= \sum_{i=1}^n \int_{Q_i} \int_{a_i}^{b_i} \frac{\partial X_i}{\partial x_i}(x_1, \dots, x_i, \dots, x_n) \, dx_i \, dx_1 \dots \widehat{dx_i} \dots dx_n \\
&= \sum_{i=1}^n \int_{Q_i} X_i(x_1, \dots, b_i, \dots, x_n) \, dx_1 \dots \widehat{dx_i} \dots dx_n \\
&\quad - \sum_{i=1}^n \int_{Q_i} X_i(x_1, \dots, a_i, \dots, x_n) \, dx_1 \dots \widehat{dx_i} \dots dx_n \\
&= \sum_{i=1}^n \int_{\{x_i=b_i\}} \langle X(x), e_i \rangle \, dA + \sum_{i=1}^n \int_{\{x_i=a_i\}} \langle X(x), -e_i \rangle \, dA \\
&= \int_{\partial Q} \langle X, \nu \rangle \, dA.
\end{aligned}$$

Wir wollen einen zweiten Beweis andeuten, der für allgemeine Gebiete funktioniert und ebenfalls aufschlussreich ist. Für  $x \in \bar{\Omega}$  sei  $d(x) \geq 0$  der Abstand zum Rand von  $\Omega$ . Man kann zeigen: auf der Parallelfäche  $\{x \in \Omega : d(x) = r\}$  ist die nach außen weisende Einheitsnormale  $\nu(x)$  gleich  $-\operatorname{grad} d(x)$ . Definiere nun die reelle Funktion

$$\eta_\varepsilon : [0, \infty) \rightarrow \mathbb{R}, \quad \eta_\varepsilon(r) = \begin{cases} \frac{r}{\varepsilon} & \text{falls } r \leq \varepsilon, \\ 1 & \text{sonst.} \end{cases}$$

Die Funktion  $\varphi_\varepsilon(x) = \eta_\varepsilon(d(x))$  hat Null-Randwerte, also folgt (siehe Lemma 1.1)

$$0 = \int_\Omega \operatorname{div}(\varphi_\varepsilon X) \, dx = \int_\Omega \varphi_\varepsilon \operatorname{div} X \, dx + \int_\Omega \langle \operatorname{grad} \varphi_\varepsilon, X \rangle \, dx.$$

Es gilt  $\lim_{\varepsilon \searrow 0} \varphi_\varepsilon(x) = 1$  für alle  $x \in \Omega$ , deshalb folgt

$$\lim_{\varepsilon \searrow 0} \int_\Omega \varphi_\varepsilon \operatorname{div} X \, dx = \int_\Omega \operatorname{div} X \, dx.$$

Im zweiten Integral ist  $\operatorname{grad} \varphi_\varepsilon(x) = 0$  für  $d(x) > \varepsilon$ . Für  $0 < d(x) < \varepsilon$  ist

$$\operatorname{grad} \varphi_\varepsilon(x) = \eta'_\varepsilon(d(x)) \operatorname{grad} d(x) = -\frac{1}{\varepsilon} \nu(x).$$

Dabei ist  $\nu(x)$  die Normale auf den Parallelfächen. Für das rechte Integral sollte also gelten

$$\int_\Omega \langle \operatorname{grad} \varphi_\varepsilon, X \rangle \, dx = -\frac{1}{\varepsilon} \int_{\{0 < d(x) < \varepsilon\}} \langle X(x), \nu(x) \rangle \, dx \rightarrow - \int_{\partial \Omega} \langle X, \nu \rangle \, dA.$$

□

Der Satz von Gauß gilt auch für Gebiete mit gewissen Singularitäten, zum Beispiel Kanten. Das können wir hier nicht behandeln. Stattdess machen wir noch ein paar Folgerungen und Anwendungen.

**Beispiel 3.1** Wählen wir im Satz von Gauß als Vektorfeld  $X(x) = x$ , so folgt

$$|\Omega| = \frac{1}{n} \int_{\Omega} \operatorname{div} X \, dx = \frac{1}{n} \int_{\partial\Omega} \langle x, \nu \rangle \, dA.$$

Insbesondere folgt  $|\mathbb{B}^n| = \frac{1}{n} |\mathbb{S}^{n-1}|$ .

**Folgerung 3.1 (Greensche Formel)** Sei  $\Omega$  beschränktes Gebiet im  $\mathbb{R}^n$  mit glattem Rand. Dann gilt für  $u, v \in C^2(\overline{\Omega})$

$$\int_{\Omega} (u\Delta v - v\Delta u) \, dx = \int_{\partial\Omega} \left( u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) \, dA.$$

BEWEIS: Wegen  $\operatorname{div}(u \operatorname{grad} v) = \langle \operatorname{grad} u, \operatorname{grad} v \rangle + u\Delta v$  liefert der Satz von Gauß

$$\int_{\Omega} (u\Delta v + \langle \operatorname{grad} u, \operatorname{grad} v \rangle) \, dx = \int_{\partial\Omega} u \frac{\partial v}{\partial \nu} \, dA.$$

Die Behauptung folgt nun durch Vertauschen von  $u, v$  und Subtraktion.  $\square$

**Beispiel 3.2 (Strömungen)** Sei  $v : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $v = v(x, t)$ , ein evtl. zeitabhängiges Geschwindigkeitsfeld auf  $\mathbb{R}^n$ . Wir bezeichnen mit  $\phi^x(t)$  die Bahnkurve eines Partikels, dass zur Zeit  $t = 0$  bei  $x$  startet und im Punkt  $\phi^x(t)$  die Geschwindigkeit  $v(\phi^x(t), t)$  hat. Zusammenfassen aller Bahnkurven ergibt eine Abbildung, den Fluss von  $v$ ,

$$\phi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n, \phi(x, t) = \phi^x(t).$$

Nach Definition gilt

$$\frac{\partial \phi}{\partial t}(x, t) = v(\phi(x, t), t) \quad \text{und} \quad \phi(x, 0) = x.$$

Fließen wir erst von  $x$  nach  $\phi(x, t)$ , und dann weiter von  $\phi(x, t)$  nach  $\phi(\phi(x, t), s)$ , so ist das im Ergebnis dasselbe wie ein Fluss direkt nach  $\phi(x, s + t)$ . Also gilt für die Abbildungen  $\phi_s : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\phi_s(x) = \phi(x, s)$ ,

$$\phi_s \circ \phi_t = \phi_{s+t}.$$

Insbesondere ist  $\phi_s \circ \phi_{-s} = \phi_0 = \operatorname{id}_{\mathbb{R}^n}$ , und die Abbildungen  $\phi_s$  sind bijektiv. Zur Zeit  $t = 0$  berechnen wir

$$\frac{\partial}{\partial t}(D_x \phi) = D_x \left( \frac{\partial \phi}{\partial t} \right) = D_x v,$$

insbesondere wegen  $D_x \phi(x, 0) = \operatorname{Id}_{\mathbb{R}^n}$

$$\frac{\partial}{\partial t}(\det D_x \phi)|_{t=0} = \operatorname{tr} D_x v = \operatorname{div} v.$$

Sei nun  $\varrho : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  die evtl. zeitabhängige Dichteverteilung der Strömung. Für  $\Omega \subset \mathbb{R}^n$  beschränkt sei  $m_{\Omega}(t)$  die Masse im Gebiet  $\phi_t(\Omega)$ . Mit dem Transformationssatz und Differentiation unter dem Integral folgt

$$\begin{aligned} m'_{\Omega}(0) &= \frac{d}{dt} \int_{\phi_t(\Omega)} \varrho(y, t) \, dy|_{t=0} \\ &= \frac{d}{dt} \int_{\Omega} \varrho(\phi_t(x), t) \det D_x \phi(x, t) \, dx|_{t=0} \\ &= \int_{\Omega} \left( D_x \varrho \cdot v + \frac{\partial \varrho}{\partial t} + \varrho \operatorname{div} v \right) \, dx \\ &= \int_{\Omega} \left( \frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho v) \right) \, dx. \end{aligned}$$

Für beliebige  $t \in \mathbb{R}$  ergibt sich aus dem Fall  $t = 0$  dieselbe Formel, und zwar

$$m'_{\Omega}(t) = \frac{d}{ds} \int_{\phi_s(\phi_t(\Omega))} \varrho(y, s+t) dy|_{s=0} = \int_{\phi_t(\Omega)} \left( \frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho v) \right) dx.$$

Gilt daher auf  $\mathbb{R}^n \times \mathbb{R}$  die Kontinuitätsgleichung

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho v) = 0,$$

so ist die Funktion  $m_{\Omega}(t)$  für alle  $\Omega$  zeitlich konstant. Wählen wir  $\varrho \equiv 1$ , so ist  $m_{\Omega}(t)$  das Volumen von  $\phi_t(\Omega)$ , und der Satz von Gauß impliziert

$$m'_{\Omega}(0) = \int_{\partial\Omega} \langle v, \nu \rangle dA.$$

Das Flächenelement  $dA$  trägt mit der Normalkomponente  $\langle v, \nu \rangle$  zur Änderung des eingeschlossenen Volumens bei, was eine sehr anschauliche Deutung des Gaußschen Satzes ist.