

# Vorlesungsskript zur Mathematik 2 für Maschinenbauer im SoSe 2022

Prof. Dr. Helge Glöckner

## Überblick

Die Mathematik 2 gliedert sich in drei Teile:

- Ergänzungen zur Mathematik 1 und Wiederholung  
(Extremstellenbestimmung, Integrationsregeln, uneigentliche Integrale,  
Partialbruchzerlegung, Taylorentwicklung)
- Lineare Gleichungssysteme und lineare Algebra
- Differentialrechnung für Funktionen in mehreren Veränderlichen

## Inhaltsverzeichnis

<b>§1 Wiederholung und Ergänzungen</b> .....	<b>1</b>
Wiederholung zu Extremstellen; Wiederholung zu Integralen; uneigentliche Integrale; Integration rationaler Funktionen mittels Partialbruchzerlegung; Taylorentwicklung von $C^k$ -Funktionen	
<b>§2 Lineare Gleichungssysteme</b> .....	<b>25</b>
Lineare Gleichungssysteme, Matrix-Schreibweise, Gaußsches Lösungsverfahren, Zeilen-Stufenform, freie und abhängige Variablen, Lösbarkeitstest für inhomogene lineare Gleichungssysteme, Struktur der Lösungsmengen	
<b>§3 Vektorräume, lineare Abbildungen, invertierbare Matrizen</b> ..	<b>45</b>
Vektorräume und Untervektorräume, Basen, lineare Unabhängigkeit, zu einer Matrix gehörige lineare Abbildung, invertierbare Matrizen, Gauß-Jordan-Verfahren zur Berechnung der inversen Matrix, Isomorphismen zwischen Vektorräumen, Dimension eines Vektorraums	
<b>§4 Determinanten</b> .....	<b>66</b>
Eigenschaften und Berechnung von Determinanten, Invertierbarkeitstest für quadratische Matrizen, $\det(A) \neq 0$	
<b>§5 Kern und Bild einer linearen Abbildung</b> .....	<b>72</b>
Kern und Bild einer linearen Abbildung, endlich-dimensionale Vektorräume, Dimensionsformel, Endomorphismen endlich-dimensionaler Vektorräume	
<b>§6 Diagonalisierbarkeit</b> .....	<b>78</b>
Eigenvektoren und Eigenwerte, charakteristisches Polynom, Eigenräume, Diagonalisierbarkeit reeller Matrizen und Endomorphismen, komplexe Vektorräume, Diagonalisierbarkeit über $\mathbb{C}$ . Skalarprodukt, spezielle Matrizen. Darstellungsmatrix einer linearen Abbildung	
<b>§7 Abstände in <math>\mathbb{R}^n</math>, konvergente Folgen, stetige Funktionen</b> ....	<b>97</b>
Länge eines Vektors, Cauchy-Schwarzsche Ungleichung, Dreiecksungleichung, Kugeln, Konvergenz von Folgen von Vektoren bzw. Punkten im $\mathbb{R}^n$ , stetige vektorwertige Funktionen, Stetigkeit linearer und affin-linearer Funktionen, Operatornorm, Polynome in mehreren Variablen und Multiindex-Notation. Offene Teilmengen in $\mathbb{R}^n$ , abgeschlossene Teilmengen, beschränkte Teilmengen und kompakte Teilmengen. Rand, Inneres und Abschluss einer Menge. Satz vom Maximum. Gleichmäßige Konvergenz.	

<b>§8 Kurven, Bogenlänge, Kurvenintegrale .....</b>	<b>119</b>
Kurven und Wege, vektorwertige Integrale, Bogenlänge, reguläre Wege, Umparametrisieren auf Bogenlänge, Wegintegrale erster und zweiter Art, stückweise stetig differenzierbare Wege	
<b>§9 Differentialrechnung für reellwertige Funktionen in mehreren Variablen .....</b>	<b>131</b>
Partielle Ableitungen, $C^1$ -Funktionen, Gradient, lokale Extremstellen, kritische Punkte, notwendige Bedingung für lokale Extremstellen, Richtungsableitungen, Richtung des steilsten Anstiegs, affin-lineare Approximation, Fehlerfortpflanzung bei Messfehlern	
<b>§10 Beweise für 9.12, 9.13 und 9.17.....</b>	<b>138</b>
Beweise dreier verwandter technischer Resultate	
<b>§11 Höhere Ableitungen.....</b>	<b>142</b>
$C^2$ -Funktionen, höhere partielle Ableitungen, Satz von Schwarz, Hesse-Matrix; positiv definite, negativ definite und indefinite Matrizen; Hurwitz-Kriterium; hinreichende Bedingung für lokale Extrema, Taylorentwicklung 2. Ordnung, Kettenregel (Spezialfall), stetige Parameterabhängigkeit von Integralen, Differenzieren unter dem Integral, Leibniz-Regel	
<b>§12 Potentialfunktionen .....</b>	<b>163</b>
Gradientenvektorfelder, Potentialfunktionen, Wegunabhängigkeit von Kurvenintegralen, Integrabilitätsbedingung, Methode der sukzessiven Integration zur Berechnung einer Potentialfunktion. Rotation, Divergenz, wirbelfreie Vektorfelder, quellenfreie Vektorfelder	
<b>§13 Differenzierbare vektorwertige Funktionen.....</b>	<b>170</b>
Definition, Jacobi-Matrix, Kettenregel. Mittelwertsatz in Integralform	
<b>§14 Lösen nicht-linearer Gleichungen .....</b>	<b>174</b>
Satz über implizite Funktionen, Satz über die Umkehrfunktion	
<b>§15 Extrema unter Nebenbedingungen .....</b>	<b>179</b>
Problemstellung. Parametrisierungsmethode, direkte Methode, Methode der Lagrange-Multiplikatoren. Lagrange-Multiplikatoren bei mehreren Nebenbedingungen	

<b>§16 Iterationsverfahren</b> .....	<b>186</b>
Fixpunktiteration, vereinfachtes Newton-Verfahren zur approximativen Berechnung von Nullstellen	
<b>§17 Differentialgleichungen erster Ordnung: Grundlagen</b> .....	<b>200</b>
Gewöhnliche Differentialgleichungen 1. Ordnung und Anfangswertprobleme. Elementare Lösungsmethoden: Homogene lineare Differentialgleichungen 1. Ordnung; inhomogene lineare Differentialgleichungen 1. Ordnung (Variation der Konstanten); Differentialgleichungen in getrennten Variablen	
<b>§18 Lineare Differentialgleichungssysteme erster Ordnung</b> ....	<b>211</b>
Lineare Differentialgleichungssysteme mit konstanten Koeffizienten. Existenz globaler Lösungen und Eindeutigkeit, Lösung inhomogener Systeme mittels Variation der Konstanten. Hilfsmittel: Matrixexponentialfunktion	
<b>§19 Lineare Differentialgleichungen zweiter Ordnung</b> .....	<b>227</b>
Umschreiben als System 1. Ordnung. Existenz globaler Lösungen und Eindeutigkeit, Struktur der Lösungsmenge für homogene und inhomogene Gleichungen; Lösungsfundamentalsysteme und Wronski-Determinante. Beispiele aus der Physik: Schwingungsgleichung/harmonischer Oszillator, gedämpfter Oszillator, getriebener Oszillator, getriebener gedämpfter Oszillator	

# 1 Wiederholung und Ergänzungen

## 1.1 Wiederholung: Extremstellen

(a) Ist  $f: ]a, b[ \rightarrow \mathbb{R}$  eine differenzierbare Funktion auf einem offenen Intervall und  $x_0 \in ]a, b[$  eine Extremstelle (also eine lokale Minimalstelle oder lokale Maximalstelle), so ist notwendig

$$f'(x_0) = 0,$$

also  $x_0$  ein sogenannter *kritischer Punkt*.

(b) Ist  $x_0$  ein kritischer Punkt von  $f: ]a, b[ \rightarrow \mathbb{R}$  und macht die Ableitung  $f'$  einen Vorzeichenwechsel von Minus nach Plus an der Stelle  $x_0$  (also  $f'(x) \leq 0$  für  $x < x_0$ ,  $f'(x) \geq 0$  für  $x > x_0$ ), so ist  $x_0$  eine lokale Minimalstelle von  $f$  (denn  $f|_{]a, x_0]}$  ist dann monoton fallend,  $f|_{[x_0, b[}$  monoton wachsend).

(c) Ist  $f: ]a, b[ \rightarrow \mathbb{R}$  zweimal differenzierbar und  $x_0 \in ]a, b[$  ein kritischer Punkt, so ist die Bedingung  $f''(x_0) > 0$  hinreichend dafür, dass  $x_0$  eine lokale Minimalstelle von  $f$  ist (denn  $f'$  macht dann einen Vorzeichenwechsel von Minus nach Plus an der Stelle  $x_0$ ).

(d) Wir haben hierbei folgende Notationen benutzt: Es ist  $f''(x) := (f')'(x)$  wenn existent,  $f'''(x) := (f'')'(x)$  und so weiter; allgemein  $f^{(1)} := f'$ ,  $f^{(k+1)} := (f^{(k)})'$ . Andere Schreibweisen sind  $\frac{df}{dx} := f'$  und  $\frac{d^k f}{dx^k} := f^{(k)}$ .

(e) Damit ein kritischer Punkt  $x_0$  von  $f: ]a, b[ \rightarrow \mathbb{R}$  eine lokale Maximalstelle ist, genügt es entsprechend, einen Vorzeichenwechsel von  $f'$  von Plus nach Minus zu zeigen an der Stelle  $x_0$ . Ist  $f$  zweimal differenzierbar, ist auch  $f''(x_0) < 0$  am kritischen Punkt  $x_0$  eine hinreichende Bedingung für ein lokales Maximum.

(f) Warnung: Ist das Definitionsintervall nicht ein offenes Intervall, so können sich auch an den Intervallenden lokale (oder globale) Extremstellen befinden (egal, ob dort die Ableitung verschwindet oder nicht). Zum Beispiel muss man bei einer Funktion  $f: [a, b[ \rightarrow \mathbb{R}$  immer auch den Punkt  $a$  darauf untersuchen, ob dort eine lokale Extremalstelle vorliegt (meist per Hand).

(g) Beispiel. Die Funktion  $f: [0, \pi] \rightarrow \mathbb{R}$ ,  $x \mapsto \sin(x)$  ist zweimal differenzierbar. Die Ableitung  $f'(x) = \cos(x)$  hat auf dem offenen Intervall  $]0, \pi[$  nur die Nullstelle  $\pi/2$ , dies ist der einzige kritische Punkt von  $f$ . Da

$f''(\pi/2) = -\sin(\pi/2) = -1 < 0$ , ist  $\pi/2$  eine lokale Maximalstelle von  $f$ . Wir müssen noch die Randpunkte im Definitionsintervall betrachten, also 0 und  $\pi$ . Es ist

$$f(0) = f(\pi) = 0 \leq f(x) \text{ für alle } x \in [0, \pi].$$

Also sind 0 und  $\pi$  globale (also auch lokale) Minimalstellen für  $f$ . Beide sind dennoch keine kritischen Punkte: es ist  $f'(0) = \cos(0) = 1 \neq 0$ ,  $f'(\pi) = \cos(\pi) = -1 \neq 0$ .

## 1.2 Wiederholung: Integrale

Für eine stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  haben wir ein Integral  $\int_a^b f(x) dx \in \mathbb{R}$  definiert.

(a) *Anschauliche Bedeutung des Integrals wenn  $f \geq 0$* : Ist  $f(x) \geq 0$  für alle  $x \in [a, b]$ , so können wir  $\int_a^b f(x) dx$  interpretieren als den Flächeninhalt der Fläche, die vom Graphen von  $f$  und der  $x$ -Achse (sowie den vertikalen Geraden  $x = a$  und  $x = b$ ) eingeschlossen wird. Zum Beispiel ist  $\int_0^\pi \sin(x) dx = [-\cos(x)]_0^\pi = 2$  der Flächeninhalt des "ersten Hügels" der Sinusfunktion.

*Der Fall  $f \leq 0$* . Ist  $f(x) \leq 0$  für alle  $x \in [a, b]$ , so ist  $\int_a^b f(x) dx$  das Negative des Flächeninhalts der Fläche, die vom Graphen von  $f$  und der  $x$ -Achse (sowie den Geraden  $x = a$  und  $x = b$ ) eingeschlossen wird. Zum Beispiel ist  $\int_\pi^{3\pi/2} \sin(x) dx = [-\cos(x)]_\pi^{3\pi/2} = -1$ .

*Allgemeines  $f$* : Das Integral  $\int_a^b f(x) dx$  ist die Summe der Flächeninhalte oberhalb der  $x$ -Achse minus die Summe der Flächeninhalte darunter (siehe Skizze in der Vorlesung). Zum Beispiel ist  $\int_0^{3\pi/2} \sin(x) dx = 2 - 1 = 1$ .

Konvention: Für  $f: [a, b] \rightarrow \mathbb{R}$  mit  $a \leq b$  definiert man

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

(b) Integralberechnung mit dem Hauptsatz der Integral- und Differentialrechnung. Wir nennen eine differenzierbare Funktion  $F: [a, b] \rightarrow \mathbb{R}$  eine Stammfunktion für  $f$ , wenn  $F' = f$ . Dann gilt

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

(c) Integrationsregeln. Zur Erinnerung: Eine Funktion  $f: I \rightarrow \mathbb{R}$  auf einem nicht entarteten Intervall (also mit mehr als einem Punkt) heißt *stetig differenzierbar*, wenn  $f$  differenzierbar ist und die Ableitung  $f': I \rightarrow \mathbb{R}$  eine stetige Funktion.

*Partielle Integration.* Sind  $f: [a, b] \rightarrow \mathbb{R}$  und  $g: [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar, so gilt

$$\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx.$$

*Substitutionsregel.* Ist  $f: I \rightarrow \mathbb{R}$  stetig auf einem Intervall  $I$  und  $\phi: [a, b] \rightarrow \mathbb{R}$  stetig differenzierbar mit  $\phi([a, b]) \subseteq I$ , so gilt

$$\int_a^b f(\phi(x)) \phi'(x) dx = \int_{\phi(a)}^{\phi(b)} f(y) dy.$$

Wenn wir  $y := \phi(x)$  substituieren, müssen wir formal also noch  $dy = \phi'(x) dx$  setzen und auch die Grenzen transformieren.

**Beispiel.** Es ist

$$\int_0^{\pi/2} e^{\sin(x)} \cos(x) dx = \int_0^1 e^y dy = e - 1.$$

Wir können nämlich  $y := \phi(x) := \sin(x)$  substituieren mit  $\phi'(x) = \cos(x)$ ,  $\phi(0) = 0$  und  $\phi(\pi/2) = 1$ . Hierbei haben wir die Substitutionsregel von links nach rechts gelesen. Man kann die Substitutionsregel aber auch von rechts nach links lesen, also zur Berechnung des rechten Integrals benutzen. Hierzu muss  $\phi$  geeignet gewählt werden, was oft nur mit Erfahrung möglich ist.

**Beispiel.** Wir wollen das Integral  $\int_0^{\frac{1}{2}\sqrt{2}} \sqrt{1-y^2} dy$  berechnen. Wir substituieren  $\phi(x) = \sin(x)$  mit  $\phi'(x) = \cos(x)$ ,  $0 = \sin(0)$  und  $\frac{1}{2}\sqrt{2} = \sin(\pi/4)$ . Nach der Substitutionsregel ist dann

$$\begin{aligned} \int_0^{\frac{1}{2}\sqrt{2}} \sqrt{1-y^2} dy &= \int_0^{\pi/4} \sqrt{1-\phi(x)^2} \phi'(x) dx = \int_0^{\pi/4} \underbrace{\sqrt{1-\sin^2(x)}}_{=\sqrt{\cos^2(x)}=\cos(x)} \cos(x) dx \\ &= \int_0^{\pi/4} \underbrace{\cos^2(x)}_{=\frac{1}{2}(1+\cos(2x))} dx = \left[ \frac{1}{2}x + \frac{1}{4}\sin(2x) \right]_0^{\pi/4} = \frac{1}{4} + \frac{\pi}{8}. \end{aligned}$$

*Unbestimmte Integrale:* Ist  $f: I \rightarrow \mathbb{R}$  eine stetige Funktion auf einem nicht entarteten Intervall  $I$ , so schreibt man  $\int f(x) dx$  für die Menge aller Stammfunktionen von  $f$ . Ist  $F$  eine solche, so ist also

$$\int f(x) dx = \{F + C : C \in \mathbb{R}\}.$$

Abkürzend schreibt man stattdessen meist  $\int f(x) dx = F + C$ . Man nennt  $\int f(x) dx$  das *unbestimmte Integral* von  $f$ . (Im Gegensatz dazu wird  $\int_a^b f(x) dx$  auch das *bestimmte Integral* genannt).

Zum Beispiel ist  $\int \cos(x) dx = \sin(x) + C$ .

Die Integrationsregel der partiellen Integration kann man auch für unbestimmte Integrale formulieren; sie lautet dann

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx.$$

### 1.3 Uneigentliche Integrale

Es sei  $f: [a, \infty[ \rightarrow \mathbb{R}$  stetig. Wir schreiben

$$\int_a^\infty f(x) dx := \lim_{r \rightarrow \infty} \int_a^r f(x) dx,$$

wenn der Limes existiert. Dies ist ein Funktionengrenzwert im Sinne der Mathematik I; gefordert ist also, dass der Grenzwert

$$\lim_{n \rightarrow \infty} \int_a^{r_n} f(x) dx \tag{1}$$

existiert für jede Folge  $(r_n)_{n \in \mathbb{N}}$  in  $[a, \infty[$  mit  $r_n \rightarrow \infty$  für  $n \rightarrow \infty$ ; das uneigentliche Integral ist dann der Grenzwert in (1) (dieser ist von der gewählten Folge  $(r_n)_{n \in \mathbb{N}}$  dann unabhängig).

**Beispiel.** Es ist  $\int_0^\infty e^{-x} dx = 1$ . Für jede Folge von Zahlen  $r_n \geq 0$  mit  $r_n \rightarrow \infty$  für  $n \rightarrow \infty$  gilt nämlich

$$\int_0^{r_n} e^{-x} dx = [-e^{-x}]_0^{r_n} = -e^{-r_n} + e^0 \rightarrow e^0 = 1$$

für  $n \rightarrow \infty$ , weil  $e^{-r_n} \rightarrow 0$ .

Man kann auch direkt mit  $r \geq 0$  arbeiten, ohne eine Folge zu benutzen:

$$\int_0^r e^{-x} dx = [-e^{-x}]_0^r = -e^{-r} + e^0 \rightarrow e^0 = 1$$

für  $r \rightarrow \infty$ . Dies macht die Rechnung übersichtlicher.

Statt zu sagen, dass ein uneigentliches Integral existiert, sagt man auch, dass das uneigentliche Integral *konvergiert*.

**Beispiel.** Für  $\alpha \in \mathbb{R}$  existiert das uneigentliche Integral  $\int_1^\infty \frac{1}{x^\alpha} dx$  genau dann, wenn  $\alpha > 1$ . In diesem Fall ist  $\int_1^\infty \frac{1}{x^\alpha} dx = \frac{1}{\alpha-1}$ .

Für  $\alpha = 1$  existiert das uneigentliche Integral nicht, denn

$$\int_1^r \frac{1}{x} dx = [\ln(x)]_1^r = \ln(r) - \ln(1) = \ln(r)$$

divergiert bestimmt gegen  $\infty$  für  $r \rightarrow \infty$ .

Im Fall  $\alpha \neq 1$  hat  $1/x^\alpha = x^{-\alpha}$  die Stammfunktion  $\frac{1}{-\alpha+1}x^{-\alpha+1} = x^{1-\alpha}/(1-\alpha)$ . Für  $r \geq 1$  ist also

$$\int_1^r \frac{1}{x^\alpha} dx = \left[ \frac{1}{1-\alpha} x^{1-\alpha} \right]_1^r = \frac{1}{1-\alpha} (r^{1-\alpha} - 1). \quad (2)$$

Im Falle  $\alpha > 1$  ist  $1-\alpha < 0$ , so dass  $r^{1-\alpha} \rightarrow 0$  für  $r \rightarrow \infty$ . Die rechte Seite von (2) konvergiert dann also gegen  $\frac{1}{1-\alpha}(0-1)$ , so dass das uneigentliche Integral existiert und den behaupteten Wert hat.

Im Fall  $\alpha < 1$  ist  $1-\alpha > 0$ , so dass  $r^{1-\alpha}$  bestimmt gegen  $\infty$  divergiert für  $r \rightarrow \infty$ . Die rechte Seite von (2) konvergiert somit nicht für  $r \rightarrow \infty$ , das uneigentliche Integral existiert also nicht.

Ähnlich wie bei der Konvergenz unendlicher Reihen gibt es auch für uneigentliche Integrale ein Majorantenkriterium (mit ähnlichem Beweis).

**Majorantenkriterium.** Es sei  $f: [a, \infty[ \rightarrow \mathbb{R}$  stetig. Gibt es eine stetige Funktion  $g: [a, \infty[ \rightarrow \mathbb{R}$  derart, dass  $|f(x)| \leq g(x)$  für alle  $x \in [a, \infty[$  und das uneigentliche Integral  $\int_a^\infty g(x) dx$  existiert, so existiert auch  $\int_a^\infty f(x) dx$ .

**Beispiel.** Das uneigentliche Integral  $\int_0^\infty \frac{e^{-x} \sin(x)}{1+x} dx$  existiert. Da  $|\sin(x)| \leq 1$  und  $\frac{1}{1+x} \leq 1$  für alle  $x \geq 0$ , ist nämlich  $\left| \frac{e^{-x} \sin(x)}{1+x} \right| = \frac{e^{-x}}{1+x} |\sin(x)| \leq e^{-x}$ . Da

das uneigentliche Integral  $\int_0^\infty e^{-x} dx$  existiert (wie wir oben gesehen haben), ist  $g(x) := e^{-x}$  eine integrierbare Majorante für  $f$  und wir können das Majorantenkriterium anwenden.

**Bemerkung** (weitere Arten uneigentlicher Integrale). Analog definieren wir für eine stetige Funktion  $f: ]a, b[ \rightarrow \mathbb{R}$  das uneigentliche Integral

$$\int_a^b f(x) dx := \lim_{r \uparrow b} \int_a^r f(x) dx$$

wenn der Grenzwert existiert; dies ist also der Grenzwert der Integrale  $\int_a^{r_n} f(x) dx$  für Folgen  $(r_n)_{n \in \mathbb{N}}$  in  $]a, b[$  mit  $r_n \rightarrow b$  für  $n \rightarrow \infty$ . Ist  $f: ]a, b[ \rightarrow \mathbb{R}$  stetig, definieren wir

$$\int_a^b f(x) dx := \lim_{r \downarrow a} \int_r^b f(x) dx$$

wenn existent. Ist  $f: ]a, b[ \rightarrow \mathbb{R}$  stetig, wählen wir  $c \in ]a, b[$ . Existieren beide der uneigentlichen Integrale  $\int_a^c f(x) dx$  (mit Grenzwert an der unteren Integrationsgrenze) und  $\int_c^b f(x) dx$  (mit Grenzwert an der oberen Integrationsgrenze), so sagen wir, dass das uneigentliche Integral  $\int_a^b f(x) dx$  existiert und definieren es als

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

Man kann zeigen, dass Existenz und Zahlenwert von  $\int_a^b f(x) dx$  unabhängig sind vom gewählten Zerlegungspunkt  $c$ .

Auch für die gerade aufgezählten weiteren Arten uneigentlicher Integrale gilt das Majorantenkriterium in der naheliegenden Form.

## 1.4 Reelle Polynome, komplexe Polynome und ihre Nullstellen

Im folgenden Kapitel wollen wir rationale Funktionen der Form  $\frac{q(x)}{p(x)}$  integrieren mit Polynomen  $p$  und  $q$ . Um dies vorzubereiten, stellen wir etwas Grundwissen über reelle und komplexe Polynome zusammen.

**Koeffizienten eines Polynoms und Ableitungen.** Aus der Schule (und

der Mathematik 1) kennen wir die Ableitung eines Monoms: Für  $j \in \mathbb{N}_0$  ist

$$\frac{d}{dx}(x^j) = \begin{cases} j x^{j-1} & \text{wenn } j \geq 1; \\ 0 & \text{wenn } j = 0. \end{cases}$$

Induktion nach  $k \in \mathbb{N}$  zeigt nun, dass für jedes  $j \in \mathbb{N}_0$

$$\frac{d^k}{dx^k}(x^j) = \begin{cases} j(j-1) \cdots (j-k+1) x^{j-k} & \text{wenn } j \geq k; \\ 0 & \text{wenn } j < k. \end{cases} \quad (3)$$

Da  $0^{j-k}$  nur für  $j = k$  von 0 verschieden ist (mit  $0^0 = 1$ ), erhalten wir

$$\left. \frac{d^k}{dx^k} \right|_{x=0} (x^j) = k! \delta_{jk} \quad (4)$$

für alle  $j, k \in \mathbb{N}_0$  mit dem Kronecker-Delta  $\delta_{jk}$ ,

$$\delta_{jk} := \begin{cases} 1 & \text{wenn } j = k; \\ 0 & \text{wenn } j \neq k \end{cases}$$

(der Fall  $k = 0$  ist klar, da  $\frac{d^0}{dx^0}(x^j) = x^j$  per Definition). Wir können somit die Koeffizienten  $a_0, \dots, a_n \in \mathbb{R}$  eines Polynoms

$$p: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{j=0}^n a_j x^j$$

aus den Ableitungen des Polynoms zurückgewinnen: Es ist

$$a_k = \frac{p^{(k)}(0)}{k!} \quad \text{für alle } k \in \{0, 1, \dots, n\}. \quad (5)$$

In der Tat haben wir

$$p^{(k)}(0) = \sum_{j=0}^n a_j \left. \frac{d^k}{dx^k} (x^j) \right|_{x=0} = \sum_{j=0}^n a_j k! \delta_{jk} = a_k k!$$

nach (4); beachten Sie, dass alle Summanden mit  $j \neq k$  gleich 0 sind. Auch im Falle eines komplexen Polynoms

$$p: \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \sum_{j=0}^n a_j z^j$$

mit  $a_0, \dots, a_n \in \mathbb{C}$  sind die Koeffizienten  $a_0, \dots, a_n$  durch  $p$  festgelegt: Für  $x \in \mathbb{R}$  ist nämlich der Realteil

$$\operatorname{Re}(p(x)) = \sum_{j=0}^n \operatorname{Re}(a_j) x^j$$

ein reelles Polynom in  $x \in \mathbb{R}$ , legt also die Koeffizienten  $\operatorname{Re}(a_0), \dots, \operatorname{Re}(a_n)$  fest. Ebenso legt der Imaginärteil die Koeffizienten  $\operatorname{Im}(a_0), \dots, \operatorname{Im}(a_n)$  fest.

**Nullstellen eines komplexen Polynoms.** In diesem Abschnitt erläutern wir, dass sich jedes komplexe Polynom  $p(z) = \sum_{j=0}^n a_j z^j$  vom Grad  $n \in \mathbb{N}$  schreiben lässt als

$$p(z) = a_n(z - w_1)^{\nu_1} \cdots (z - w_m)^{\nu_m}, \quad (6)$$

wobei  $w_1, \dots, w_m \in \mathbb{C}$  mit  $m \in \{1, \dots, n\}$  die paarweise verschiedenen komplexen Nullstellen von  $p$  sind, die Zahlen  $\nu_k \in \mathbb{N}$  für  $k \in \{1, \dots, m\}$  eindeutig festgelegt sind und die *Vielfachheit* der Nullstelle  $w_k$  genannt werden. Insbesondere hat  $p$  höchstens  $n$  verschiedene Nullstellen.

Sind  $p$  und  $q$  komplexe Polynome vom Grad  $\leq n$  und ist  $p|_M = q|_M$  für eine Teilmenge  $M \subseteq \mathbb{C}$  mit mindestens  $n + 1$  Elementen, so ist  $p = q$  (denn  $p - q$  ist ein Polynom vom Grad  $\leq n$  mit mindestens  $n + 1$  Nullstellen. Dies ist nur für das Nullpolynom möglich, also ist  $p - q = 0$ ). Also:

$$p|_M = q|_M \quad \Rightarrow \quad p = q \quad \text{wenn } |M| > \text{Grad von } p \text{ und Grad von } q. \quad (7)$$

Wenn Sie (6) als Black Box akzeptieren wollen, können Sie zum Abschnitt über Nullstellen reeller Polynome springen.

Zum Nachweis von (6) erinnern wir an den Hauptsatz der Algebra:

*Jedes nicht konstante komplexe Polynom hat eine komplexe Nullstelle.*

Daraus folgern wir, dass sich für  $n \in \mathbb{N}$  jedes komplexe Polynom  $p(z) = \sum_{j=0}^n a_j z^j$  mit  $a_n \neq 0$  (also vom Grad  $n$ ) in Linearfaktoren zerlegen lässt:

$$p(z) = a_n(z - z_1) \cdots (z - z_n) \quad (8)$$

mit gewissen  $z_1, \dots, z_n \in \mathbb{C}$ . Ist  $n = 1$ , so liefert der Hauptsatz eine Nullstelle  $z_1 \in \mathbb{C}$  für  $p$ . Dann ist

$$0 = p(z_1) = a_1 z_1 + a_0,$$

also  $a_0 = -a_1 z_1$ , also

$$p(z) = a_1 z_1 - a_1 z_1 = a_1(z - z_1).$$

Ist  $n \geq 2$ , so liefert der Hauptsatz eine Nullstelle  $z_n \in \mathbb{C}$  für  $p$ . Eine Polynomdivision liefert

$$p(z) = (z - z_n)q(z) + r(z)$$

mit einem konstanten Polynom  $r(z) = c$  und einem Polynom  $q(z) = a_n z^{n-1} + \dots$  vom Grad  $n - 1$  mit Leitkoeffizient  $a_n$ . Dann ist

$$0 = p(z_n) = (z_n - z_n)q(z_n) + c = c,$$

also  $c = 0$  und somit  $p(z) = (z - z_n)q(z)$ . Per Induktion ist

$$q(z) = a_n(z - z_1) \cdots (z - z_{n-1})$$

und somit hat  $p$  die Form (8).

Einsetzen von  $z := z_j$  in (8) zeigt, dass jedes  $z_j$  eine Nullstelle von  $p$  ist. Weitere Nullstellen gibt es nicht: Ist die komplexe Zahl  $z$  nicht in  $\{z_1, \dots, z_n\}$ , so ist  $z - z_j \neq 0$  für alle  $j$  und somit  $p(z) = a_n(z - z_1) \cdots (z - z_n) \neq 0$ . Wir können (8) nun umschreiben als (6),

$$p(z) = a_n(z - w_1)^{\nu_1} \cdots (z - w_m)^{\nu_m},$$

wobei  $w_1, \dots, w_m$  mit  $m \in \{1, \dots, n\}$  die paarweise verschiedenen komplexen Nullstellen von  $p$  sind und  $\nu_k$  die *Vielfachheit* der Nullstelle  $w_k$  ist, also die Anzahl der  $j \in \{1, \dots, n\}$  mit  $z_j = w_k$ . Die Vielfachheiten sind durch  $p$  eindeutig festgelegt. Setzen wir  $q(z) = a_n(z - w_2)^{\nu_2} \cdots (z - w_m)^{\nu_m}$ , so gilt für  $\ell = \nu_1$  nämlich

$$\frac{p(z)}{(z - w_1)^\ell} = q(z) \rightarrow q(z_1) \neq 0 \tag{9}$$

für  $z \neq z_1$  wenn  $z \rightarrow z_1$ . Für  $\ell < \mu_1$  gilt hingegen

$$\frac{p(z)}{(z - w_1)^\ell} = (z - z_1)^{\mu_1 - \ell} q(z) \rightarrow 0 q(z_1) = 0$$

und für  $\ell > \mu_1$  gilt

$$\left| \frac{p(z)}{(z - w_1)^\ell} \right| = \frac{1}{|z - w_1|^{\ell - \mu_1}} |q(z)| \rightarrow \infty.$$

Also ist  $\nu_1$  über die Eigenschaften von  $p$  festgelegt (das eindeutige  $\ell \in \mathbb{N}$ , für das (9) gilt). Entsprechend sind  $\nu_2, \dots, \nu_m$  eindeutig.

**Nullstellen reeller Polynome.** Ist  $p: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto \sum_{j=0}^n a_j x^j$  ein reelles Polynom vom Grad  $n \in \mathbb{N}$ , wobei  $a_0, \dots, a_n \in \mathbb{R}$  und  $a_n \neq 0$ , so benutzen wir den gleichen Buchstaben,  $p$ , für das zugehörige komplexe Polynom

$$p: \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \sum_{j=0}^n a_j z^j.$$

Ist  $w \in \mathbb{C} \setminus \mathbb{R}$  eine nicht reelle, komplexe Nullstelle von  $p$ , so ist auch das komplex Konjugierte  $\bar{w}$  eine Nullstelle von  $p$ . Da jedes  $a_j$  reell ist und somit  $\overline{a_j} = a_j$ , ist nämlich

$$0 = \bar{0} = \overline{p(w)} = \sum_{j=0}^n a_j (\bar{w})^j = p(\bar{w}).$$

In diesem Abschnitt wollen wir zeigen:

*Ist  $w \in \mathbb{C} \setminus \mathbb{R}$  eine nicht reelle, komplexe Nullstelle von  $p$ , so stimmen die Vielfachheiten der Nullstellen  $w$  und  $\bar{w}$  von  $p$  überein.*

Wenn Sie dies als Black-Box akzeptieren wollen, können Sie gern zum Abschnitt über Partialbruchzerlegungen springen. Wenn gewünscht, finden Sie Details zum Nachlesen jetzt (oder in Zukunft) hier.

Sind  $w_1, \dots, w_m$  die paarweise verschiedenen komplexen Nullstellen von  $p$  und ist  $\nu_k$  die Vielfachheit von  $w_k$ , so ist nach dem Obigen

$$p(z) = a_n (z - w_1)^{\nu_1} \cdots (z - w_m)^{\nu_m} \tag{10}$$

für alle  $z \in \mathbb{C}$ . Da  $p(x) \in \mathbb{R}$  für  $x \in \mathbb{R}$ , mit komplex Konjugiertem  $\overline{p(x)} = p(x)$ , ist für alle  $x \in \mathbb{R}$

$$p(x) = \overline{p(x)} = a_n (x - \bar{w}_1)^{\nu_1} \cdots (x - \bar{w}_m)^{\nu_m}.$$

Nach (7) gilt dann auch

$$p(z) = a_n (z - \bar{w}_1)^{\nu_1} \cdots (z - \bar{w}_m)^{\nu_m} \tag{11}$$

für alle  $z \in \mathbb{C}$ . Sei  $w \in \mathbb{C} \setminus \mathbb{R}$  eine nicht reelle, komplexe Nullstelle von  $p$ . Dann ist  $w = w_j$  für ein  $j \in \{1, \dots, m\}$ . Nach dem Vorigen ist auch das komplex Konjugierte  $\bar{w}_j$  eine Nullstelle von  $p$ , also  $\bar{w}_j = w_k$  für ein  $k \in \{1, \dots, m\}$ . Da Vielfachheiten eindeutig sind, liefert Vergleich von (10) und (11), dass  $\nu_j = \nu_k$ . Die Vielfachheiten von  $w$  und  $\bar{w}$  stimmen also überein.

## 1.5 Integration rationaler Funktionen mittels Partialbruchzerlegung

In diesem Abschnitt wollen wir rationale Funktionen der Form  $f(x) = \frac{q(x)}{p(x)}$  integrieren, wobei  $p \neq 0$  und  $q$  reelle Polynome (Polynomfunktionen) sind. Der Definitionsbereich von  $f$  ist also

$$\mathbb{R} \setminus \{x \in \mathbb{R} : p(x) = 0\}.$$

Mittels Polynomdivision (wie in der Schule) finden Sie stets Polynome  $Q$  und  $R$  derart, dass

$$q = Qp + R$$

und der Grad  $\deg(R)$  von  $R$  kleiner dem Grad von  $p$  ist. Dann ist

$$\frac{q}{p} = Q + \frac{R}{p}.$$

Für das Polynom  $Q$  können wir sofort eine Stammfunktion hinschreiben und Integrale berechnen. Es bleibt also nur noch, Stammfunktionen für die rationale Funktion  $R/p$  zu finden (auf Intervallen außerhalb der Nullstellen von  $p$ ). Schreiben wir statt  $R$  wieder  $q$ , brauchen wir also nur noch den Fall  $\deg(q) < \deg(p)$  zu betrachten. Wir beginnen mit Beispielen.

**Beispiel.** Wir betrachten die rationale Funktion

$$f: \mathbb{R} \setminus \{-2, -3\} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{x^2 + 5x + 6}.$$

Es ist also  $f(x) = \frac{q(x)}{p(x)}$  mit der konstanten Polynomfunktion  $q = 1$  und dem Nennerpolynom  $p(x) = x^2 + 5x + 6$  mit den einfachen reellen Nullstellen  $-2$  und  $-3$ . Da  $p$  normiert ist (also den Leitkoeffizienten 1 vor der höchsten  $x$ -Potenz besitzt), ist dann

$$p(x) = (x + 2)(x + 3),$$

was man durch Ausmultiplizieren sofort verifiziert. Wir behaupten, dass es Konstanten  $A, B \in \mathbb{R}$  gibt derart, dass

$$f(x) = \frac{1}{(x + 2)(x + 3)} = \frac{A}{x + 2} + \frac{B}{x + 3} \quad \text{für alle } x \in \mathbb{R} \setminus \{-2, -3\}. \quad (12)$$

Multiplikation beider Seiten mit  $(x+2)(x+3) \neq 0$  zeigt, dass (12) genau dann gilt, wenn

$$1 = A(x+3) + B(x+2) \quad \text{für alle } x \in \mathbb{R} \setminus \{-2, -3\}.$$

Da beide Seiten der vorigen Gleichung stetige Funktionen auf  $\mathbb{R}$  sind und wir an den zwei Definitionslücken zum Grenzwert übergehen können, gilt die vorige Gleichung genau dann, wenn

$$1 = A(x+3) + B(x+2) \quad \text{für alle } x \in \mathbb{R} \tag{13}$$

(alternativ folgt dies mit (7). Die Polynome  $0x+1$  und  $A(x+3)+B(x+2) = (A+B)x + (3A+2B)$  müssen also gleich sein. Koeffizientenvergleich vor den  $x$ -Potenzen rechts und links zeigt, dass dies genau dann der Fall ist, wenn  $A+B=0$  und (\*\*)  $3A+2B=1$  ist. Ersteres verlangt  $B=-A$  und Einsetzen in (\*\*) zeigt, dass noch  $A=3A-2A=1$  zu verlangen ist, also  $A=1$  (und somit  $B=-1$ ). Wir können also  $f$  auf die Form (12) bringen; es ist

$$f(x) = \frac{1}{x+2} - \frac{1}{x+3}. \tag{14}$$

Auf jedem nicht entarteten Intervall  $I \subseteq \mathbb{R} \setminus \{-2, -3\}$  ist also

$$\ln|x+2| - \ln|x+3|$$

eine Stammfunktion für  $f$ .

Im Falle einfacher Nullstellen (wie hier) kann man den Koeffizientenvergleich vermeiden; es ist schneller, in (13) einfach die speziellen Werte  $x=-2$  und  $x=-3$  (also die Nullstellen von  $p$ ) einzusetzen. Dies liefert die notwendigen Bedingungen

$$1 = A, \quad 1 = -B.$$

**Beispiel.** Die rationale Funktion

$$f: \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}, \quad x \mapsto \frac{x+1}{x^2-2x+1}$$

hat das Nennerpolynom

$$p(x) = x^2 - 2x + 1 = (x-1)^2$$

mit der doppelten reellen Nullstelle 1 (wobei die zweite binomische Formel benutzt wurde). Wir machen den Ansatz

$$f(x) = \frac{x+1}{(x-1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} \quad \text{für alle } x \in \mathbb{R} \setminus \{1\}$$

mit Konstanten  $A, B \in \mathbb{R}$ . Multiplikation mit  $(x-1)^2$  zeigt, dass vorige Gleichung genau dann gilt, wenn

$$x+1 = A(x-1) + B$$

für alle  $x \in \mathbb{R} \setminus \{1\}$ , bzw. (äquivalent zum Vorigen) für alle  $x \in \mathbb{R}$ . Verlangt ist also

$$1x+1 = Ax + (B-A)$$

für alle  $x \in \mathbb{R}$ . Koeffizientenvergleich zeigt, dass die Polynome rechts und links genau dann gleich sind, wenn

$$1 = A \quad \text{und} \quad 1 = B - A.$$

Einsetzen von  $A = 1$  macht aus der zweiten Bedingung  $1 = B - 1$ , was genau für  $B = 2$  erfüllt ist. Wir können in der Tat  $f$  also auf die Form

$$f(x) = \frac{1}{x-1} + \frac{2}{(x-1)^2} \tag{15}$$

bringen. Auf jedem nicht-entarteten Intervall  $I \subseteq \mathbb{R} \setminus \{1\}$  hat  $f$  folglich die Stammfunktion

$$\ln|x-1| - \frac{2}{x-1}.$$

Wir zeigen, dass zu (14) und (15) analoge ‘‘Partialbruchzerlegungen’’ für jede rationale Funktion möglich sind. Eine Komplikation entsteht allerdings dadurch, dass das Nennerpolynom  $p$  nicht immer über  $\mathbb{R}$  in Linearfaktoren zerfällt, sondern nicht-reelle komplexe Nullstellen besitzen kann.

**Notationen.** Wir betrachten ein normiertes Polynom<sup>1</sup>

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

---

<sup>1</sup>Der Leitkoeffizient ist also 1.

mit  $a_0, \dots, a_{n-1} \in \mathbb{R}$ . Hat  $p$  eine komplexe, nicht reelle Nullstelle  $w$  in dem Sinne, dass

$$w^n + a_{n-1}w^{n-1} + \dots + a_1w + a_0 = 0 \text{ in } \mathbb{C},$$

so ist auch die komplex konjugierte Zahl  $\bar{w}$  eine Nullstelle von  $p$ ; zudem stimmt die Vielfachheit  $\nu$  der Nullstelle  $w$  mit derjenigen von  $\bar{w}$  überein (siehe Abschnitt 1.4). Das Produkt der entsprechenden Linearfaktoren ist

$$(x - w)^\nu (x - \bar{w})^\nu = (p_w(x))^\nu$$

mit

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

Dies ist ein normiertes reelles Polynom zweiten Grades ohne reelle Nullstellen. Seien  $\alpha_1, \dots, \alpha_k$  die paarweise verschiedenen reellen Nullstellen von  $p$  mit den Vielfachheiten  $m_1, \dots, m_k$  und

$$\{w_1, \bar{w}_1\}, \{w_2, \bar{w}_2\}, \dots, \{w_\ell, \bar{w}_\ell\}$$

die verschiedenen Paare von komplex konjugierten, nicht reellen Nullstellen von  $p$ . Seien  $n_1, \dots, n_\ell$  die Vielfachheiten von  $w_1, \dots, w_\ell$ . Dann ist

$$p(x) = \prod_{j=1}^k (x - \alpha_j)^{m_j} \prod_{j=1}^{\ell} (x - w_j)^{n_j} (x - \bar{w}_j)^{n_j}$$

und somit

$$p(x) = \prod_{j=1}^k (x - \alpha_j)^{m_j} \prod_{j=1}^{\ell} (p_{w_j}(x))^{n_j}. \quad (16)$$

Wir benutzen folgende Tatsache.<sup>2</sup>

**Satz über die Partialbruchzerlegung.** *Seien  $p, q: \mathbb{R} \rightarrow \mathbb{R}$  Polynomfunktionen mit  $\deg(p) \geq 1$  und  $\deg(q) < \deg(p)$ . Mit Notationen wie in (16) existieren dann reelle Zahlen  $A_{j,\nu}$  für  $j \in \{1, \dots, k\}$  und  $\nu \in \{1, \dots, m_j\}$  sowie reelle Zahlen  $B_{j,\nu}$  und  $C_{j,\nu}$  für  $j \in \{1, \dots, \ell\}$  und  $\nu \in \{1, \dots, n_j\}$  derart, dass für alle  $x \in \mathbb{R}$  mit  $p(x) \neq 0$*

$$\frac{q(x)}{p(x)} = \sum_{j=1}^k \sum_{\nu=1}^{m_j} \frac{A_{j,\nu}}{(x - \alpha_j)^\nu} + \sum_{j=1}^{\ell} \sum_{\nu=1}^{n_j} \frac{B_{j,\nu} + C_{j,\nu}x}{(p_{w_j}(x))^\nu}. \quad (17)$$

---

<sup>2</sup>Einen Beweis finden Sie ggf. im Skript zur Analysis 2 auf der Homepage von Prof. Glöckner.

**Bemerkung.** Die Partialbruchzerlegung für  $\frac{q(x)}{p(x)}$  ist also eine Summe von Beiträgen:

- Für jede einfache reelle Nullstelle  $\alpha_j$  von  $p$  gibt es einen Summanden  $\frac{A_{j,1}}{x-\alpha_j}$ .

- Für jede  $m_j$ -fache reelle Nullstelle  $\alpha_j$  von  $p$  gibt es einen Summanden der Form

$$\frac{A_{j,1}}{x-\alpha_j} + \dots + \frac{A_{j,m_j}}{(x-\alpha_j)^{m_j}}.$$

- Für jedes Paar  $\{w_j, \bar{w}_j\}$  aus einer einfachen komplexen, nicht reellen Nullstelle von  $p$  und ihrem komplex Konjugierten gibt es einen Summanden der Form

$$\frac{B_{j,1} + C_{j,1}x}{p_{w_j}(x)}$$

mit  $p_{w_j}(x) := x^2 - 2 \operatorname{Re}(w_j)x + |w_j|^2$ .

- Für jedes Paar  $\{w_j, \bar{w}_j\}$  aus einer  $n_j$ -fachen komplexen, nicht reellen Nullstelle von  $p$  und ihrem komplex Konjugierten gibt es einen Summanden der Form

$$\frac{B_{j,1} + C_{j,1}x}{p_{w_j}(x)} + \dots + \frac{B_{j,n_j} + C_{j,n_j}x}{p_{w_j}(x)^{n_j}}.$$

**Beispiel.** Wir suchen die Partialbruchzerlegung für die durch

$$f(x) = \frac{1}{(x-4)^2(x-2)}$$

gegebene rationale Funktion. Das Nennerpolynom  $q(x) = (x-4)^2(x-2)$  hat zwei Nullstellen, die beide reell sind, nämlich die Nullstelle  $\alpha_1 = 4$  mit Vielfachheit  $n_1 = 2$  und die Nullstelle  $\alpha_2 = 2$  mit Vielfachheit  $n_2 = 1$ . Es ist also  $k = 2$  und  $\ell = 0$ . Nach dem Satz über Partialbruchzerlegung gibt es reelle Zahlen  $A := A_{1,1}$ ,  $B := A_{1,2}$  und  $C := A_{2,1}$  derart, dass

$$\frac{1}{(x-4)^2(x-2)} = \frac{A}{x-4} + \frac{B}{(x-4)^2} + \frac{C}{x-2}$$

für alle  $x \in \mathbb{R} \setminus \{2, 4\}$ . Äquivalent hierzu ist (nach Multiplizieren beider Seiten mit dem linken Nennerpolynom  $(x-4)^2(x-2)$ )

$$1 = A(x-4)(x-2) + B(x-2) + C(x-4)^2 \tag{18}$$

für alle  $x \in \mathbb{R} \setminus \{2, 4\}$  (und somit für alle  $x \in \mathbb{R}$ , da beide Seiten an Stellen  $x_n \in \mathbb{R} \setminus \{2, 4\}$  gleich sind mit  $x_n \rightarrow 2$  und wir nun bei beiden stetigen Funktionen zum Grenzwert übergehen können). Einsetzen von  $x = 2$  in (18) liefert die Bedingung  $1 = 4C$ , somit

$$C = \frac{1}{4}.$$

Einsetzen von  $x = 4$  in (18) liefert die Bedingung  $1 = 2B$ , also

$$B = \frac{1}{2}.$$

Wir setzen diese Werte von  $B$  und  $C$  in (18) ein und erhalten die Bedingung

$$1 = A(x-4)(x-2) + \frac{1}{2}(x-2) + \frac{1}{4}(x-4)^2 = Ax^2 + \frac{1}{4}x^2 + \text{ein Polynom vom Grad } < 2.$$

Koeffizientengleich rechts und links vor dem Monom  $x^2$  liefert die Bedingung  $0 = A + \frac{1}{4}$ , so dass also

$$A = -\frac{1}{4}$$

(stattdessen könnten Sie natürlich auch in (18) einen Koeffizientenvergleich vor  $x^2$ ,  $x^1$  und  $x^0$  durchführen und ein lineares Gleichungssystem mit drei Gleichungen für  $A$ ,  $B$ ,  $C$  erhalten und nach diesen auflösen). Also ist

$$f(x) := \frac{1}{(x-4)^2(x-2)} = -\frac{1}{4} \frac{1}{x-4} + \frac{1}{2} \frac{1}{(x-4)^2} + \frac{1}{4} \frac{1}{x-2}$$

mit Stammfunktionen der Form

$$\int f(x) dx = -\frac{1}{4} \ln|x-4| - \frac{1}{2} \frac{1}{x-4} + \frac{1}{4} \ln|x-2| + \text{Konst.}$$

auf jedem der Teilintervalle  $]-\infty, 2[$ ,  $]2, 4[$  bzw.  $]4, \infty[$ .

**Beispiel.** Wir suchen die Partialbruchzerlegung für die durch

$$f(x) = \frac{1}{x(x^2 + 1)}$$

gegebene rationale Funktion. Das Nennerpolynom  $q(x) = x(x^2 + 1)$  hat drei Nullstellen, nämlich die einfache reelle Nullstelle  $\alpha_1 = 0$  und die einfachen

komplexen, nicht reellen Nullstellen  $\lambda_1 = i$  und  $\overline{\lambda_1} = -i$ , die zueinander komplex konjugiert sind. Also ist  $k = 1$ ,  $n_1 = 1$ ,  $\ell = 1$  und  $m_1 = 1$ . Nach dem Satz über Partialbruchzerlegung gibt es reelle Zahlen  $A := A_{1,1}$ ,  $B := B_{1,1}$  und  $C := C_{1,1}$  derart, dass

$$\frac{1}{x(x^2 + 1)} = \frac{A}{x} + \frac{B + Cx}{x^2 + 1}$$

für alle  $x \in \mathbb{R} \setminus \{0\}$ . Äquivalent hierzu ist (nach Multiplizieren beider Seiten mit  $q(x)$ )

$$1 = A(x^2 + 1) + (B + Cx)x \quad (19)$$

für alle  $x \in \mathbb{R} \setminus \{0\}$  (und somit für alle  $x \in \mathbb{C}$ , siehe (7)). Einsetzen von  $x = 0$  in (19) liefert  $A = 1$ . Einsetzen von  $x = i$  liefert die Bedingung

$$1 = (B + Ci)i = Bi - C;$$

Vergleich der Real- bzw. Imaginärteile beider Seiten liefert  $C = -1$  und  $B = 0$ . Folglich ist

$$f(x) := \frac{1}{x(x^2 + 1)} = \frac{1}{x} - \frac{x}{x^2 + 1}$$

mit Stammfunktionen der Form

$$\int f(x) dx = \ln|x| - \frac{1}{2} \ln(1 + x^2) + \text{Konst.}$$

auf jedem der Teilintervalle  $]-\infty, 0[$  bzw.  $]0, \infty[$ .

Für die einzelnen Summanden einer Partialbruchzerlegung lässt sich stets eine Stammfunktion angeben:

(a) Für  $\alpha \in \mathbb{R}$  ist

$$\int \frac{1}{x - \alpha} dx = \ln|x - \alpha| + \text{Konst.}$$

auf  $]-\infty, \alpha[$  bzw.  $]\alpha, \infty[$ .

(b) Für  $\alpha \in \mathbb{R}$  und natürliche Zahlen  $n \geq 2$  ist

$$\int \frac{1}{(x - \alpha)^n} dx = \frac{1}{1 - n} \frac{1}{(x - \alpha)^{n-1}} + \text{Konst.}$$

auf  $]-\infty, \alpha[$  bzw.  $]\alpha, \infty[$ .

(c) Für alle reellen Zahlen  $a, b, B, C$  mit  $b \neq 0$  ist

$$\begin{aligned} \int \frac{B + Cx}{(x - a)^2 + b^2} dx \\ = \frac{C}{2} \ln((x - a)^2 + b^2) + \frac{B + aC}{b} \arctan\left(\frac{x - a}{b}\right) + \text{Konst.} \end{aligned} \quad (20)$$

auf  $\mathbb{R}$ .

Hierbei verifiziert man (a) und (b) direkt durch Ableiten.

Nachweis von (c): Ersetzen von  $Cx$  durch  $C(x - a) + Ca$  im Zähler von  $\frac{B+Cx}{(x-a)^2+b^2}$  liefert

$$\begin{aligned} \frac{B + Cx}{(x - a)^2 + b^2} &= \frac{C(x - a)}{(x - a)^2 + b^2} + \frac{B + aC}{(x - a)^2 + b^2} \\ &= \frac{C}{2} \frac{2(x - a)}{(x - a)^2 + b^2} + \frac{B + aC}{b} \frac{1/b}{\left(\frac{x-a}{b}\right)^2 + 1}. \end{aligned}$$

Um den ersten Summanden zu integrieren, substituieren wir  $u = (x - a)^2 + b^2$ ,  $du = 2(x - a) dx$ ; um den zweiten zu integrieren, substituieren wir  $v = \frac{x - a}{b}$ ,  $dv = \frac{1}{b} dx$ . Dies liefert die in (20) angegebene Stammfunktion.

Sei  $n \in \mathbb{N}$ . Auf dem 3. Übungsblatt zeigen Sie:

- (a) Eine Stammfunktion für  $\frac{B+Cx}{((x-a)^2+b^2)^{n+1}}$  lässt sich hinschreiben, sobald wir eine für  $\frac{1}{((x-a)^2+b^2)^{n+1}}$  kennen.
- (b) Für die Stammfunktion von  $\frac{1}{((x-a)^2+b^2)^{n+1}}$  gibt es eine Rekursionsformel, welche diese auf die Stammfunktion von  $\frac{1}{((x-a)^2+b^2)^n}$  zurückführt.

## 1.6 Taylorentwicklung von $C^k$ -Funktionen

In diesem Kapitel untersuchen wir zunächst, wie gut eine differenzierbare Funktion durch ihre Tangente um eine Stelle  $x_0$  approximiert wird. Die Tangente ist der Graph einer affin-linearen Funktion (also eines Polynoms vom Grad  $\leq 1$ ). Anschließend untersuchen wir die Approximation nahe  $x_0$  durch Polynome höheren Grades.

Sei  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall. Einer an einer Stelle  $x_0 \in I$  differenzierbaren Funktion  $f: I \rightarrow \mathbb{R}$  haben wir ihre Tangente zugeordnet, also die Gerade mit der Gleichung

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

Wie gut wird die Funktion  $f$  nahe  $x_0$  durch ihre Tangente approximiert? Die Differenz der beiden ist

$$R(x) := f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Wir zeigen nun, dass  $R(x) \rightarrow 0$  für  $x \rightarrow x_0$  und sogar<sup>3</sup>

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

Letztere Eigenschaft legt die Tangente zudem eindeutig fest:

**Satz.** *Es sei  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $x_0 \in I$ . Eine Funktion  $f: I \rightarrow \mathbb{R}$  ist genau dann an der Stelle  $x_0$  differenzierbar, wenn für ein  $a \in \mathbb{R}$  der Fehler*

$$R(x) := f(x) - f(x_0) - a(x - x_0)$$

der affin-linearen Approximation  $f(x) = f(x_0) + a(x - x_0) + R(x)$  die Bedingung

$$\lim_{x \rightarrow x_0} \frac{R(x)}{x - x_0} = 0 \tag{21}$$

erfüllt. In diesem Fall gilt  $a = f'(x_0)$ .

Gilt (21), so folgt nämlich

$$\frac{f(x) - f(x_0)}{x - x_0} = a + \frac{R(x)}{x - x_0} \rightarrow a$$

für  $x \rightarrow x_0$ , d.h.  $f$  ist an der Stelle  $x_0$  differenzierbar mit Ableitung  $f'(x_0) = a$ . Ist umgekehrt  $f$  an der Stelle  $x_0$  differenzierbar und setzen wir  $a := f'(x_0)$ , so gilt

$$\frac{R(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \rightarrow f'(x_0) - f'(x_0) = 0$$

---

<sup>3</sup>Daraus folgt  $R(x) = (x - x_0) \frac{R(x)}{x - x_0} \rightarrow 0 \cdot 0 = 0$  für  $x \rightarrow x_0$ .

für  $x \rightarrow x_0$ .

Sei  $x_0 \in \mathbb{R}$ . Wie im Abschnitt 1.4 (wo  $x_0 = 0$  war) zeigt man: Für alle  $k \in \mathbb{N}$  und  $j \in \mathbb{N}_0$  gilt

$$\frac{d^k}{dx^k}((x - x_0)^j) = \begin{cases} j(j-1) \cdots (j-k+1)(x - x_0)^{j-k} & \text{wenn } j \geq k; \\ 0 & \text{wenn } j < k \end{cases}$$

(vergleiche (3)). Da  $0^{j-k}$  nur für  $j = k$  von 0 verschieden ist (mit  $0^0 = 1$ ), erhalten wir analog zu (4)

$$\left. \frac{d^k}{dx^k} \right|_{x=x_0} (x - x_0)^j = k! \delta_{jk}$$

für alle  $j, k \in \mathbb{N}_0$  mit dem Kronecker-Delta  $\delta_{jk}$ .

Seien nun  $a_0, \dots, a_n \in \mathbb{R}$  mit  $n \in \mathbb{N}$ . Wir betrachten das Polynom

$$p: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{j=0}^n a_j (x - x_0)^j$$

Wie in (5) sehen wir, dass

$$a_k = \frac{p^{(k)}(x_0)}{k!} \quad \text{für alle } k \in \{0, 1, \dots, n\}.$$

Wir folgern:

**Definition.** Sei  $I \subseteq \mathbb{R}$  eine nicht entartetes Intervall,  $x_0 \in I$  und  $f: I \rightarrow \mathbb{R}$  eine  $C^n$ -Funktion mit  $n \in \mathbb{N}_0$ . Dann ist

$$p: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$$

das einzige Polynom vom Grad  $\leq n$  derart, dass  $p^{(k)}(x_0) = f^{(k)}(x_0)$  für alle  $k \in \{0, 1, \dots, n\}$ , das also die gleichen Funktionswerte und Ableitungen an der Stelle  $x_0$  hat wie  $f$  bis einschließlich der  $n$ ten Ableitung. Wir nennen  $p$  das *n*te *Taylorpolynom* von  $f$  um den Entwicklungspunkt  $x_0$  und schreiben auch  $T_{x_0}^n(f)(x) := p(x)$ . Wir nennen

$$R_n(x) := f(x) - T_{x_0}^n(f)(x)$$

das zugehörige *Restglied*. Es gilt dann also

$$f(x) = P_{x_0}^n(f)(x) + R_n(x) = \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + R_n(x) \quad \text{für alle } x \in I.$$

**Beispiele.** Es ist  $\sin(0) = 0$ ,  $\sin'(0) = \cos(0) = 1$ ,  $\sin''(0) = -\sin(0) = 0$  und  $\sin'''(0) = -\cos(0) = -1$ . Die  $n$ ten Taylorpolynome der Sinusfunktion um den Entwicklungspunkt  $x_0 = 0$  sind für  $n \in \{0, 1, 2, 3\}$  also

$$P_0^0(\sin)(x) = 0, \quad P_0^1(\sin)(x) = P_0^2(\sin)(x) = x, \quad P_0^3(\sin)(x) = x - \frac{1}{6}x^3.$$

Da  $\cos(0) = 1$ ,  $\cos'(0) = -\sin(0) = 0$ ,  $\cos''(0) = -\cos(0) = -1$  und  $\cos'''(0) = \sin(0) = 0$ , lauten die  $n$ ten Taylorpolynome der Cosinusfunktion am Entwicklungspunkt  $x_0 = 0$  für  $n \in \{0, 1, 2, 3\}$

$$P_0^0(\cos)(x) = P_0^1(\cos)(x) = 1, \quad P_0^2(\cos)(x) = P_0^3(\cos)(x) = 1 - \frac{1}{2}x^2.$$

Da  $\exp' = \exp$  und somit  $\exp^{(k)} = \exp$  für alle  $k \in \mathbb{N}_0$ , ist  $\exp^{(k)}(0) = \exp(0) = 1$  und folglich

$$P_0^0(\exp)(x) = 1, \quad P_0^1(\exp)(x) = 1 + x, \quad P_0^2(\exp)(x) = 1 + x + \frac{1}{2}x^2,$$

und  $P_0^3(\exp)(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$ .

Wir fragen uns nun, wie gut  $f(x)$  durch  $P_{x_0}^n f(x)$  approximiert wird, untersuchen also das Restglied  $R_n(x)$  der Taylorapproximation

$$f(x) = P_{x_0}^n f(x) + R_n(x).$$

Um simple, quantitative Formeln für das Restglied  $R_n(x)$  zu haben, setzen wir im folgenden Satz  $f$  als  $C^{n+1}$  voraus.

**Satz von Taylor.** *Es sei  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall,  $x_0 \in I$  und  $f: I \rightarrow \mathbb{R}$  eine  $C^{n+1}$ -Funktion mit  $n \in \mathbb{N}_0$ . Für  $x \in I$  sei*

$$R_n(x) := f(x) - f(x_0) - f'(x_0)(x - x_0) - \dots - \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n,$$

so dass also  $f(x) = f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n(x)$ .  
Dann gilt:

- (a) (Restglied in Integralform). Es ist  $R_n(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt$ .
- (b) (Restglied von Lagrange). Es ist  $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)^{n+1}$  für eine reelle Zahl  $\xi$  zwischen  $x_0$  und  $x$ .
- (c) Es gilt

$$\lim_{x \rightarrow x_0} \frac{R_n(x)}{(x-x_0)^n} = 0.$$

Bevor wir den Satz beweisen, schauen wir ein Beispiel an:

**Beispiel.** Wir wollen sehen, dass für alle  $x \in [-\frac{1}{5}, \frac{1}{5}]$

$$|\cos(x) - P_0^3(\cos)(x)| \leq \frac{1}{10.000},$$

wobei  $P_0^3(\cos)(x) = 1 - \frac{1}{2}x^2$  (siehe oben).

Ersetzen wir  $\cos(x)$  durch  $1 - \frac{1}{2}x^2$ , erhalten wir für  $x \in [-\frac{1}{5}, \frac{1}{5}]$  anschaulich also höchstens einen Fehler in der vierten Stelle hinter dem Komma.

Vorgehen: Es ist  $|\cos(x) - P_0^3(\cos)(x)| = |R_3(x)|$ . Wir benutzen die Lagrange-Form des Restglieds mit  $x_0 = 0$ ; da  $\cos^{(4)} = \cos$ , ist für ein  $\xi$  zwischen 0 und  $x$

$$R_3(x) = \frac{\cos^{(4)}(\xi)}{4!} x^4,$$

folglich

$$|R_3(x)| \leq |\cos(\xi)| \frac{|x|^4}{4!} \leq 1 \cdot \frac{(1/5)^4}{4!} = \frac{1}{25} \frac{1}{25} \frac{1}{24} \leq \frac{1}{25} \frac{1}{20} \frac{1}{20} = \frac{1}{10.000}.$$

**Beweis des Satzes von Taylor.** Wir zeigen zunächst (a) per Induktion nach  $n \in \mathbb{N}_0$ . Im Fall  $n = 0$  gilt nach dem Hauptsatz der Integral- und Differentialrechnung

$$R_0(x) = f(x) - f(x_0) = \int_{x_0}^x f'(t) dt = \frac{1}{0!} \int_{x_0}^x (x-t)^0 f'(t) dt$$

und somit (a). Gilt (a) bereits für  $n$  und ist  $f$  eine  $C^{n+2}$ -Funktion, so wenden wir auf das Restglied

$$R_n(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt$$

aus (a) eine Partielle Integration an mit

$$u'(t) = (x-t)^n, \quad u(t) = -\frac{1}{n+1}(x-t)^{n+1},$$

$v(t) = f^{(n+1)}(t)$  und  $v'(t) = f^{(n+2)}(t)$ . Wir erhalten

$$\begin{aligned} R_n(x) &= \left[-\frac{1}{n+1}(x-t)^{n+1}f^{(n+2)}(t)\right]_{t=x_0}^{t=x} + \frac{1}{n+1} \int_{x_0}^x (x-t)^{n+1}f^{(n+2)}(t) dt \\ &= \frac{1}{n+1}(x-x_0)^{n+1}f^{(n+2)}(x) + \frac{1}{n+1} \int_{x_0}^x (x-t)^{n+1}f^{(n+2)}(t) dt \end{aligned}$$

und somit

$$\begin{aligned} R_{n+1}(x) &= f(x) - \sum_{j=0}^{n+1} \frac{f^{(j)}(x_0)}{j!} (x-x_0)^j \\ &= f(x) - \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x-x_0)^j - \frac{f^{(n+1)}(x_0)}{(n+1)!} (x-x_0)^{n+1} \\ &= R_n(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!} (x-x_0)^{n+1} \\ &= \frac{1}{n+1} \int_{x_0}^x (x-t)^{n+1}f^{(n+2)}(t) dt, \end{aligned}$$

wie benötigt.

(b) Die Formel ist für  $x = x_0$  trivial (man nehmen  $\xi = x_0$ ). Sei nun  $x \neq x_0$ . Wir geben den Beweis im Falle  $x > x_0$  (der Fall  $x < x_0$  lässt sich analog behandeln). Die stetige Funktion  $[x_0, x] \rightarrow \mathbb{R}$ ,  $t \mapsto f^{(n)}(t)$  nimmt ein Maximum (an einer Stelle  $x^* \in [x_0, x]$ ) und ein Minimum (an einer Stelle  $x_* \in [x_0, x]$ ). Da  $(x-t)^{n+1} \geq 0$  für alle  $t \in [x_0, x]$ , folgt mit der Monotonie des Integrals aus der Integralform des Restglieds, dass

$$\frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(x_*) dt \leq R_n(x) \leq \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(x^*) dt.$$

Teilen durch  $K := \frac{1}{n!} \int_{x_0}^x (x-t)^n dt = \frac{(x-x_0)^{n+1}}{(n+1)!}$  liefert

$$f^{(n+1)}(x_*) \leq \frac{R_n(x)}{K} \leq f^{(n+1)}(x^*).$$

Nach dem Zwischenwertsatz für stetige Funktionen existiert ein  $\xi$  zwischen  $x_*$  und  $x^*$  (und somit in  $[x_0, x]$ ) derart, dass

$$\frac{R_n(x)}{K} = f^{(n+1)}(\xi)$$

und somit

$$R_n(x) = K f^{(n+1)}(\xi) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

(c) Ist  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $I \setminus \{x\}$  mit  $x_k \rightarrow x_0$  für  $k \rightarrow \infty$ , so gibt es für jedes  $k \in \mathbb{N}$  ein  $\xi_k$  zwischen  $x_0$  und  $x_k$  derart, dass

$$R_n(x_k) = \frac{f^{(n+1)}(\xi_k)}{(n+1)!} (x_k - x_0)^{n+1}.$$

Da  $|\xi_k - x_0| \leq |x_k - x_0| \rightarrow 0$ , folgt  $\xi_k \rightarrow x_0$ . Wir folgern, dass

$$\frac{R_n(x_k)}{(x_k - x_0)^n} = f^{(n+1)}(\xi_k)(x_k - x_0) \rightarrow f^{(n+1)}(x_0)(x_0 - x_0) = 0$$

für  $k \rightarrow \infty$ .

**Bemerkung.** Im Falle einer  $C^2$ -Funktion  $f: I \rightarrow \mathbb{R}$  haben wir für das Restglied  $R_2(x)$  auch die Formel

$$R_2(x) = \int_{x_0}^x (x-t)(f^{(2)}(t) - f^{(2)}(x_0)) dt \quad (22)$$

zur Verfügung (die uns später nutzen wird). Es ist nämlich

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R_1(x)$$

mit

$$\begin{aligned} R_1(x) &= \int_{x_0}^x (x-t)f^{(2)}(t) dt \\ &= \int_{x_0}^x (x-t)f^{(2)}(x_0) dt + \int_{x_0}^x (x-t)(f^{(2)}(t) - f^{(2)}(x_0)) dt \\ &= \frac{f^{(2)}(x_0)}{2}(x - x_0)^2 + \int_{x_0}^x (x-t)(f^{(2)}(t) - f^{(2)}(x_0)) dt, \end{aligned}$$

also

$$\begin{aligned} R_2(x) &= f(x) - f(x_0) - f'(x_0)(x - x_0) - \frac{f^{(2)}(x_0)}{2}(x - x_0)^2 \\ &= R_1(x) - \frac{f^{(2)}(x_0)}{2}(x - x_0)^2 = \int_{x_0}^x (x-t)(f^{(2)}(t) - f^{(2)}(x_0)) dt. \end{aligned}$$

## 2 Lineare Gleichungssysteme

In diesem Kapitel wollen wir systematisch lineare Gleichungssysteme lösen wie zum Beispiel

$$\begin{cases} x - y = 2 \\ 2x + y = 1. \end{cases}$$

**2.1** Ein *lineares Gleichungssystem* (LGS) mit  $m$  Gleichungen für  $n$  Unbekannte ist von der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Hierbei sind  $b_1, \dots, b_m$  und  $a_{ij}$  gegebene reelle Zahlen für  $i \in \{1, \dots, m\}$ ,  $j \in \{1, \dots, n\}$ . Gesucht sind  $x_1, \dots, x_n \in \mathbb{R}$ , für welche die Gleichungen alle erfüllt sind. Ist  $b_1 = b_2 = \cdots = b_m = 0$ , so heißt das lineare Gleichungssystem *homogen* (andernfalls *inhomogen*). Die Menge aller Lösungen  $(x_1, \dots, x_n)$  des LGS nennt man seine *Lösungsmenge*.

**Beispiel 2.2**

$$\begin{cases} x + y = 1 \\ x^2 - y = 2 \end{cases}$$

ist kein lineares Gleichungssystem, da  $x^2$  vorkommt.

**Beispiel 2.3** Wir wollen die Lösungen des inhomogenen LGS

$$\begin{cases} x - y = 2 \\ 2x + y = 1 \end{cases}$$

finden. Hierzu ziehen wir das Zweifache der ersten Gleichung von der zweiten Gleichung ab (wir deuten das durch eine Notiz rechts vom LGS an). Wir erhalten ein neues LGS, das genau dann erfüllt ist, wenn das vorige es ist. (da wir durch Addieren des Zweifachen der ersten Zeile wieder zum ersten LGS zurück können). Also:

$$\left[ \begin{array}{l} x - y = 2 \\ 2x + y = 1 \end{array} \right] (-2)$$

führt auf

$$\begin{cases} x - y = 2 \\ 3y = -3. \end{cases}$$

Die zweite Gleichung ist genau dann erfüllt, wenn  $y = -1$ . Setzen wir  $y = -1$  in die erste Gleichung ein, wird aus dieser

$$x + 1 = 2,$$

was genau für  $x = 1$  erfüllt ist. Insbesondere gibt es genau eine Lösung.

**Beispiel 2.4** Das inhomogene LGS

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}$$

hat keine Lösung.

**Beispiel 2.5** Es ist

$$3x_1 + 2x_2 = 0 \tag{23}$$

ein homogenes LGS mit einer Gleichung für zwei Unbekannte. Wähle  $x_2 = \lambda \in \mathbb{R}$  beliebig. Dann wird aus (23)

$$3x_1 + 2\lambda = 0 \quad \Leftrightarrow \quad x_1 = -\frac{2}{3}\lambda.$$

Insbesondere gibt es unendlich viele Lösungen!

**2.6** Zwei lineare Gleichungssysteme heißen *äquivalent*, wenn sie die gleiche Lösungsmenge besitzen.

**2.7 Gaußsches Lösungsverfahren** (Gaußsches Eliminationsverfahren). Zwei lineare Gleichungssysteme sind äquivalent, wenn sie durch eine (oder mehrere) der folgenden elementaren Operationen auseinander hervorgehen:

- (i) Vertausche zwei Gleichungen;
- (ii) Multipliziere eine Gleichung mit einer Zahl  $\neq 0$ ;
- (iii) Addiere ein reelles Vielfaches einer Gleichung zu einer anderen (oder ziehe es ab).

Jede der Umformungen lässt sich auch rückgängig machen.

**Bemerkung 2.8** Die obigen elementaren Operationen nennt man auch elementare *Zeilenoperationen*.

**2.9** Wir werden sehen, dass *homogene* lineare Gleichungssysteme immer Lösungen haben und sich alle Lösungen nach dem folgenden Verfahren finden lassen:

- (a) Bringe das LGS mit elementaren Zeilenoperationen auf möglichst einfache Gestalt, die sogenannte “Zeilen-Stufenform”;
- (b) Löse nacheinander nach  $x_n, \dots, x_1$  auf (“Rückwärtssubstitution”).

Bevor wir das Verfahren detaillierter erklären, schauen wir uns Beispiele an.

**Beispiel 2.10**

$$\begin{aligned} & \begin{cases} x + 0 + z = 0 \\ x + y + 2z = 0 \\ 0 + y + 2z = 0 \end{cases} \begin{matrix} ] (-1) \\ \downarrow \\ \end{matrix} \\ \Leftrightarrow & \begin{cases} x + 0 + z = 0 \\ 0 + y + z = 0 \\ 0 + y + 2z = 0 \end{cases} \begin{matrix} ] (-1) \\ \downarrow \\ \end{matrix} \\ \Leftrightarrow & \begin{cases} x + 0 + z = 0 & (*) \\ 0 + y + z = 0 & (**) \\ 0 + 0 + z = 0 & (***) \end{cases} \end{aligned}$$

Löse von unten nach oben auf:

$$\begin{aligned} (***) & \Leftrightarrow z = 0 \\ (**) & \Leftrightarrow y = -z \Leftrightarrow y = 0 \\ (*) & \Leftrightarrow x = -z \Leftrightarrow x = 0. \end{aligned}$$

**Beispiel 2.11**

$$\begin{aligned} & \begin{cases} 3x + 0 + z = 0 \\ x + y + 0 = 0 \end{cases} \begin{matrix} ] (-\frac{1}{3}) \\ \downarrow \\ \end{matrix} \\ \Leftrightarrow & \begin{cases} 3x + 0 + z = 0 & (*) \\ 0 + y - \frac{1}{3}z = 0 & (**) \end{cases} \end{aligned}$$

Sei  $z = \lambda \in \mathbb{R}$  beliebig. Dann  $(**) \Leftrightarrow y = \frac{1}{3}z = \frac{1}{3}\lambda$  und  $(*) \Leftrightarrow 3x = -z = -\lambda$ , also  $x = -\frac{1}{3}\lambda$ .

Alle Lösungen:  $(x, y, z) = (-\frac{1}{3}\lambda, \frac{1}{3}\lambda, \lambda)$  mit  $\lambda \in \mathbb{R}$ .

Hilfsmittel, um weniger anschreiben zu müssen:

**Matrizen, Spaltenvektoren und Zeilenvektoren.** Seien  $m, n \in \mathbb{N}$ .

**2.12** Eine  $m \times n$ -Matrix ist ein rechteckiges Zahlenschema der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

mit  $m$  Zeilen und  $n$  Spalten.

Die  $a_{ij} \in \mathbb{R}$  heißen *Einträge* (oder Komponenten) der Matrix. Man schreibt auch

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

oder kurz  $A = (a_{ij})$ .

$m \times 1$ -Matrizen werden *Spaltenmatrizen* oder *Spaltenvektoren* genannt und haben die Form

$$\mathbf{s} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

$1 \times n$ -Matrizen werden *Zeilenmatrizen* oder *Zeilenvektoren* genannt und haben die Form

$$\mathbf{z} = (a_1 \cdots a_n);$$

wir schreiben auch  $(a_1, \dots, a_n)$  mit Kommas. Eine  $m \times n$ -Matrix  $A = (a_{ij})$  wie oben besteht aus den  $m$  Zeilenvektoren  $\mathbf{z}_i = (a_{i1}, \dots, a_{in})$  für  $i \in \{1, \dots, m\}$ : Es ist

$$A = \begin{pmatrix} \text{--- } \mathbf{z}_1 \text{ ---} \\ \text{--- } \mathbf{z}_2 \text{ ---} \\ \vdots \\ \text{--- } \mathbf{z}_m \text{ ---} \end{pmatrix}.$$

Wir können  $A$  auch aus den  $n$  Spaltenvektoren

$$\mathbf{s}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

aufbauen mit  $j \in \{1, \dots, n\}$ ; es ist

$$A = \left( \begin{array}{c|c|c|c} | & | & \cdots & | \\ \mathbf{s}_1 & \mathbf{s}_2 & & \mathbf{s}_n \\ | & | & & | \end{array} \right).$$

Ist  $m = n$ , so nennen wir  $A$  eine *quadratische Matrix*. Zum Beispiel ist

$$A = \begin{pmatrix} 3 & 4 \\ 1 & -1 \end{pmatrix}$$

eine quadratische Matrix mit  $m = n = 2$ .

**Beispiel 2.13** Es ist

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 0 & 4 \end{pmatrix}$$

eine  $2 \times 3$ -Matrix (also nicht quadratisch). Die Zeilenvektoren sind  $\mathbf{z}_1 = (1, 3, 5)$  und  $\mathbf{z}_2 = (2, 0, 4)$ , die Spaltenvektoren

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{s}_3 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

Zum Vergleich:

$$\begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 3 & 4 \end{pmatrix}$$

ist eine  $3 \times 2$ -Matrix.

**2.14** Wir schreiben  $\mathbb{R}^{m \times n}$  für die Menge aller  $m \times n$ -Matrizen,  $\mathbb{R}^m := \mathbb{R}^{m \times 1}$  für die Menge der Spaltenvektoren und  $\mathbb{R}_n := \mathbb{R}^{1 \times n}$  für die Menge der Zeilenvektoren. Wir schreiben

$$\mathbf{0} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^m$$

und (ohne Verwechslungsgefahr) auch  $\mathbf{0} := (0, \dots, 0) \in \mathbb{R}_n$ .

**2.15** Die Summe von zwei Zeilenvektoren  $\mathbf{v} = (v_1, \dots, v_n)$  und  $\mathbf{w} = (w_1, \dots, w_n)$  definieren wir durch Summieren der Komponenten:

$$\mathbf{v} + \mathbf{w} := (v_1 + w_1, \dots, v_n + w_n).$$

Ebenso wird komponentenweise mit reellen Zahlen  $\lambda$  multipliziert:

$$\lambda \mathbf{v} = (\lambda v_1, \dots, \lambda v_n).$$

Mit Spaltenvektoren verfährt man genauso:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}, \quad \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

**2.16** Für das LGS aus (2.1) schreiben wir kurz

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

oder ganz kurz

$$A\mathbf{x} = \mathbf{b},$$

wobei  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  und

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

gegeben sind und die Lösungen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

gesucht werden.

**Beispiel 2.17** Das inhomogene LGS

$$\begin{cases} 2x_1 + 3x_2 = 1 \\ 5x_1 + 2x_2 = -1 \end{cases}$$

lässt sich schreiben als

$$\begin{pmatrix} 2 & 3 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

**2.18** Wendet man die elementaren Zeilenoperationen (i), (ii) und (iii) aus 2.7 auf ein homogenes LGS  $A\mathbf{x} = \mathbf{0}$  an, wird daraus ein LGS  $A'\mathbf{x} = \mathbf{0}$ , wobei  $A'$  wie folgt aus  $A$  hervorgeht:

- (i) Beim Vertauschen zweier Gleichungen werden die entsprechenden Zeilen der Matrix vertauscht.
- (ii) Multiplizieren wir die  $j$ te Gleichung mit  $\lambda \neq 0$ , wird die  $j$ te Zeile der Matrix mit  $\lambda$  multipliziert.
- (iii) Addieren wir ein Vielfaches der  $j$ ten Gleichung zur  $k$ ten Gleichung mit  $j \neq k$  (oder subtrahieren es), so ist das entsprechende Vielfache der  $j$ ten Zeile zur  $k$ ten Zeile der Matrix zu addieren (bzw. abzuziehen).

**Beispiel 2.19** Das homogene LGS

$$\begin{cases} 2x_1 + 3x_2 = 0 \\ 5x_1 + 2x_2 = 0 \end{cases}$$

lässt sich schreiben als

$$\begin{pmatrix} 2 & 3 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Um es zu lösen, wenden wir elementare Zeilenoperationen (wie in 2.18) auf die Koeffizientenmatrix an:

$$\begin{pmatrix} 2 & 3 \\ 5 & 2 \end{pmatrix} \begin{array}{l} \downarrow \\ \downarrow \end{array} -\frac{5}{2} \rightsquigarrow \begin{pmatrix} 2 & 3 \\ 0 & -\frac{11}{2} \end{pmatrix}.$$

Nun schreiben wir das entsprechende homogene LGS hin und Lösen es von unten nach oben:

$$\begin{cases} 2x_1 + 3x_2 = 0 \\ -\frac{11}{2}x_2 = 0; \end{cases}$$

es ist also  $x_2 = 0$  und  $x_1 = 0$ .

**Bemerkung 2.20** Die Spaltenvektoren  $\mathbf{x}$  und  $\mathbf{0}$  zwischendurch nicht mehr hinschreiben zu müssen, wissen wir natürlich erst bei größeren Gleichungssystemen zu schätzen, dann spart es Schreibarbeit.

**2.21** (Details des Gaußschen Lösungsverfahrens). Wir wollen nun systematisch ein homogenes LGS

$$A\mathbf{x} = \mathbf{0}$$

lösen mit einer  $m \times n$ -Matrix  $A = (a_{ij})$ . Hierzu formen wir  $A$  durch elementare Zeilenoperationen um zu einer Matrix  $A'$  in "Zeilenstufenform," wie folgt.

Ist  $A$  die Nullmatrix, also  $a_{ij} = 0$  für alle  $i$  und  $j$ , so setzen wir  $A' = A$  und springen nach 2.22.

Andernfalls ist eine Spalte  $\mathbf{s}_j$  von  $A$  von  $\mathbf{0}$  verschieden. Wir wählen  $j$  minimal. Dann ist also  $a_{ij} \neq 0$  für ein  $i \in \{1, \dots, m\}$ . Nachdem wir notfalls die erste und die  $j$ te Zeile von  $A$  vertauschen, dürfen wir annehmen, dass  $a_{1j} \neq 0$ . Die Matrix  $A$  ist dann von der Form

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1j} & * & \cdots & * \\ & & & a_{2j} & & & \\ & & & \vdots & & & \\ 0 & \cdots & 0 & a_{mj} & * & \cdots & * \end{pmatrix}$$

mit  $a_{1j} \neq 0$ , wobei  $*$  jeweils ein beliebiger Eintrag ist. Für alle  $i \in \{2, \dots, m\}$  ziehen wir nun nacheinander das  $\frac{a_{ij}}{a_{1j}}$ -fache der ersten Zeile von der  $i$ ten Zeile ab. Wir dürfen danach annehmen, dass  $A$  von der Form

$$\begin{pmatrix} 0 & \cdots & 0 & \blacksquare & * & \cdots & * \\ & & & 0 & & & \\ & & & \vdots & & & B \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}$$

ist mit einer  $(m-1) \times (n-j)$ -Matrix  $B$ ; hierbei bedeutet  $\blacksquare$  einen von Null verschiedenen Matrixeintrag. Wir fahren so fort mit  $B$  an Stelle von  $A$ . Nach endlich vielen Schritten gelangen wir zu einer Matrix  $A'$ , welche im folgenden Sinne Zeilenstufenform besitzt:

**2.22** Man sagt, eine Matrix  $A' \in \mathbb{R}^{m \times n}$  hat *Zeilenstufenform*, wenn sie von

folgender Gestalt ist:

$$A' = \begin{pmatrix} 0 & \cdots & 0 & \blacksquare & * & \cdots & & & & * \\ & & & & & & \blacksquare & * & \cdots & * \\ & & & & & & & & \vdots & \vdots \\ & & & & & & & & \blacksquare & * & \cdots & * \\ 0 & & & & & & & & & & & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & & & & & & & & & & & 0 \end{pmatrix}.$$

Für ein  $r \in \{0, 1, \dots, m\}$  (das man den *Rang* der Zeilenstufenform nennt) sind also alle der ersten  $r$  Zeilen von  $A'$  von Null verschieden. Die letzten  $m-r$  Zeilen sind alle  $\mathbf{0}$ . Für  $i \in \{1, \dots, r\}$  ist das Symbol  $\blacksquare$  der erste Eintrag der  $i$ ten Zeile, der nicht 0 ist. Links von  $\blacksquare$  stehen nur Nullen sowie unterhalb dieser Nullen. Auch unter  $\blacksquare$  stehen nur Nullen. Von Zeile zu Zeile rückt  $\blacksquare$  also mindestens um eine Position nach rechts.

**2.23** Habe  $A'$  Zeilenstufenform wie in 2.22. Steht in der  $j$ ten Spalte ein  $\blacksquare$ , so nennen wir  $x_j$  eine *abhängige Variable*; andernfalls nennen wir  $x_j$  eine *unabhängige Variable*. Letztere dürfen beliebige Werte annehmen. Von links nach rechts seien  $\lambda_1, \dots, \lambda_{n-r} \in \mathbb{R}$  die Werte der unabhängigen Variablen. Wir setzen diese in das homogene LGS

$$A' \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^m$$

ein und lösen von unten nach den abhängigen Variablen auf, die dann jeweils eindeutig festgelegt sind.

**Bemerkung 2.24** Die Zahl der abhängigen Variablen ist gleich der Zahl  $r$  der Quadrate  $\blacksquare$ . Die Zahl der unabhängigen Variablen (Parameter) ist daher  $n - r$ .

**Beispiel 2.25** Wir betrachten das homogene LGS

$$Ax = \mathbf{0} \tag{24}$$

mit der  $3 \times 4$ -Matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

in Zeilenstufenform und

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Dem Schema

$$\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \left( \begin{array}{cccc} \blacksquare & 1 & 2 & 1 \\ 0 & 0 & \blacksquare & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{array}$$

entnehmen wir, dass  $x_1$  und  $x_3$  abhängige Variablen sind,  $x_2$  und  $x_4$  unabhängige Variablen. Dem obigen Vorgehen folgend setzen wir  $x_2 := \lambda_1$ ,  $x_4 := \lambda_2$  mit beliebigen Parametern  $\lambda_1, \lambda_2 \in \mathbb{R}$ . Setzen wir  $x_2$  und  $x_4$  in (24) ein, erhalten wir das LGS

$$\begin{array}{rclcl} x_1 & + & \lambda_1 & + & 2x_3 & + & \lambda_2 & = & 0 \\ & & & & 3x_3 & + & \lambda_2 & = & 0 \\ & & & & & & 0 & = & 0. \end{array}$$

Wir können dieses von unten nach oben lösen: Es ist  $x_3 = -\frac{1}{3}\lambda_2$  und

$$x_1 = -\lambda_1 - 2x_3 - \lambda_2 = -\lambda_1 + \frac{2}{3}\lambda_2 - \lambda_2 = -\lambda_1 - \frac{1}{3}\lambda_2.$$

Die Lösungsmenge des homogenen LGS (24) ist also

$$L_h = \left\{ \left( \begin{pmatrix} \lambda_1 - \frac{1}{3}\lambda_2 \\ \lambda_1 \\ -\frac{1}{3}\lambda_2 \\ \lambda_2 \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right) \right\}.$$

**2.26** Wir erläutern nun genau Vorgehen beim Lösen eines *inhomogenen* linearen Gleichungssystems

$$A\mathbf{x} = \mathbf{b}$$

bei gegebener Matrix  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  und gegebener rechter Seite

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m;$$

gesucht werden alle Lösungen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

(a) Wir bilden die sogenannte *erweiterte Koeffizientenmatrix*

$$B := (A|\mathbf{b}) \in \mathbb{R}^{m \times (n+1)},$$

welche die Spalten von  $A$  enthält, gefolgt von dem Spaltenvektor  $\mathbf{b}$  als letzter Spalte.

(b) Wir führen  $B$  mit elementaren Zeilenoperationen über in eine Matrix  $B'$  in Zeilenstufenform. Die ersten  $n$  Spalten von  $B'$  bilden dann eine Matrix  $A'$  in Zeilenstufenform, die aus  $A$  durch elementare Zeilenoperationen hervorgegangen ist. Die letzte Spalte von  $B'$  nennen wir  $\mathbf{d}$  schreiben

$$\mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}.$$

Sei  $r$  der Rang von  $A'$ . Stimmt dieser mit dem Rang von  $B'$  überein, so ist

$$B' = \left( \begin{array}{cccccccc|cc} 0 & \cdots & 0 & \blacksquare & * & \cdots & * & & * & * \\ & & & & & & \blacksquare & & & * \\ & & & & & & & \ddots & & \vdots \\ & & & & & & & & \blacksquare & * \\ 0 & & & & & & & & 0 & 0 \\ \vdots & & & & & & & & \vdots & \vdots \\ 0 & & & & & & & & 0 & 0 \end{array} \right) \quad (25)$$

mit  $r$  von  $\mathbf{0}$  verschiedenen Zeilen. Es ist dann  $r = m$  oder  $r < m$  und  $d_{r+1} = 0$ . Ist  $r$  kleiner als der Rang von  $B'$ , so ist

$$B' = \left( \begin{array}{cccccccc|c} 0 & \cdots & 0 & \blacksquare & * & \cdots & * & * & d_1 \\ & & & & & & \blacksquare & & d_2 \\ & & & & & & & \ddots & \vdots \\ & & & & & & & & d_r \\ 0 & & & & & & & \blacksquare & 0 \\ 0 & & & & & & & & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & & & & & & & & 0 \end{array} \right)$$

mit  $d_{r+1} \neq 0$ , also

$$B' = \left( \begin{array}{cccccccc|c} 0 & \cdots & 0 & \blacksquare & * & \cdots & * & * & d_1 \\ & & & & & & \blacksquare & & d_2 \\ & & & & & & & \ddots & \vdots \\ & & & & & & & & d_r \\ 0 & & & & & & & \blacksquare & 0 \\ 0 & & & & & & & & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & & & & & & & & 0 \end{array} \right).$$

(c) Lösbarkeitstest: *Das inhomogene LGS  $A\mathbf{x} = \mathbf{b}$  ist genau dann lösbar, wenn der Rang  $r$  der Zeilenstufenform  $A'$  mit dem Rang von  $B'$  übereinstimmt, also  $d_j = 0$  ist für alle  $j \in \{r+1, \dots, m\}$ .*

(In anderen Worten,  $r = m$  oder  $r < m$  und  $d_{r+1} = 0$ ).

Hat  $B'$  größeren Rang als  $A'$ , so hat das zu  $A\mathbf{x} = \mathbf{b}$  äquivalente LGS  $A'\mathbf{x} = \mathbf{d}$  die  $(r+1)$ te Gleichung

$$0x_1 + \cdots + 0x_n = d_{r+1}$$

mit  $d_{r+1} \neq 0$ . Diese Gleichung ist für kein  $\mathbf{x}$  erfüllbar.

(d) Ist der Lösbarkeitstest bestanden, also  $r$  gleich dem Rang von  $B'$ , so ist  $B'$  von der Form (25). Ähnlich wie bei homogenen Systemen nennen wir wieder die  $x_j$  mit  $j \in \{1, \dots, n\}$ , für welche die  $j$ -te Spalte von  $B'$  ein  $\blacksquare$  enthält, die

abhängigen Variablen und die anderen  $n - r$  der  $x_j$  die unabhängigen Variablen. Letzteren können wir wieder beliebige Werte  $\lambda_1, \dots, \lambda_{n-r}$  zuweisen. Wir substituieren diese Parameter für die unabhängigen Variablen in das LGS

$$A'\mathbf{x} = \mathbf{d}$$

und lösen von unten nach oben auf nach den abhängigen Variablen. Insbesondere gibt es eine Lösung.

**Beispiel 2.27** Ist das inhomogene LGS

$$\begin{cases} x_1 & & + x_3 & = & 0 \\ & x_2 & + x_3 & = & 1 \\ x_1 & + x_2 & + 2x_3 & = & 0 \end{cases} \quad (26)$$

lösbar?

Wir bilden die erweiterte Koeffizientenmatrix und bringen diese auf Zeilenstufenform:

$$\begin{aligned} & \left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \end{array} \right) \begin{array}{l} \updownarrow \\ \updownarrow \end{array} (-1) \\ \rightsquigarrow & \left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right) \begin{array}{l} \updownarrow \\ \updownarrow \end{array} (-1) \\ \rightsquigarrow & \left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{array} \right) =: B'. \end{aligned}$$

Die Matrix

$$A' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

in Zeilenstufenform hat Rang 2, die Matrix  $B'$  in Zeilenstufenform jedoch echt größeren Rang 3 (es ist  $d_3 = -1 \neq 0$ ). Das inhomogene LGS (26) besteht also nicht den Lösbarkeitstest, es hat keine Lösung.

**Beispiel 2.28** Finden Sie die Lösungen des inhomogenen LGS

$$\begin{cases} x_1 & & + x_3 & = & 0 \\ & x_2 & + x_3 & = & 1 \\ x_1 & + x_2 & + 2x_3 & = & 1. \end{cases} \quad (27)$$

Die erweiterte Koeffizientenmatrix ist diesmal

$$\left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \end{array} \right)$$

und mit den gleichen Zeilenoperationen wie im vorigen Beispiel führen wir diese über in die Zeilenstufenform

$$\left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) =: B'.$$

Diesmal haben  $B'$  und die Untermatrix

$$A' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

in Zeilenstufenform den gleichen Rang  $r = 2$ . Der Lösbarkeitstest ist also bestanden. Da  $A'$  von der Form

$$A' = \begin{pmatrix} \blacksquare & 0 & 1 \\ 0 & \blacksquare & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

ist, sind  $x_1$  und  $x_2$  abhängige Variablen,  $x_3$  eine unabhängige Variable. Wir setzen  $x_3 := \lambda \in \mathbb{R}$  in das LGS

$$A' \mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

ein und erhalten das LGS

$$\begin{cases} x_1 + \lambda = 0 \\ x_2 + \lambda = 1 \\ 0 = 0. \end{cases}$$

Auflösen von unten nach oben liefert  $x_2 = 1 - \lambda$  und  $x_1 = -\lambda$ . Die Lösungsmenge unseres inhomogenen LGS ist also

$$L_i = \left\{ \left( \begin{pmatrix} -\lambda \\ 1 - \lambda \\ \lambda \end{pmatrix} : \lambda \in \mathbb{R} \right) \right\}.$$

Wir wollen uns die Lösungsmengen homogener und inhomogener linearer Gleichungssysteme noch einmal systematischer anschauen. Es ist hilfreich, zunächst einige Rechenregeln festzuhalten.

**2.29** Für alle  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^m$  mit

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

und alle  $\lambda, \mu \in \mathbb{R}$  gilt:

(a)  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$  (Assoziativgesetz);

(b)  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$  (Kommutativgesetz);

(c) Es ist  $\mathbf{0} + \mathbf{v} = \mathbf{v}$ ;

(d)  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$  mit

$$-\mathbf{v} := \begin{pmatrix} -v_1 \\ \vdots \\ -v_m \end{pmatrix};$$

(e)  $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ ;

(f)  $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$ ;

(g)  $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$ ;

(h)  $1\mathbf{v} = \mathbf{v}$ .

Dies rechnet man sofort komponentenweise nach. Exemplarisch begründen wir (a), (b) und (e).

(a): Es ist

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \begin{pmatrix} (u_1 + v_1) + w_1 \\ \vdots \\ (u_m + v_m) + w_m \end{pmatrix} = \begin{pmatrix} u_1 + (v_1 + w_1) \\ \vdots \\ u_m + (v_m + w_m) \end{pmatrix} = \mathbf{u} + (\mathbf{v} + \mathbf{w}).$$

(b): Es ist

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_m + v_m \end{pmatrix} = \begin{pmatrix} v_1 + u_1 \\ \vdots \\ v_m + u_m \end{pmatrix} = (\mathbf{v} + \mathbf{u}).$$

(e): Es ist

$$(\lambda + \mu)\mathbf{v} = \begin{pmatrix} (\lambda + \mu)v_1 \\ \vdots \\ (\lambda + \mu)v_m \end{pmatrix} = \begin{pmatrix} \lambda v_1 + \mu v_1 \\ \vdots \\ \lambda v_m + \mu v_m \end{pmatrix} = \lambda\mathbf{v} + \mu\mathbf{v}.$$

**2.30** Zu (a)–(h) in 2.29 analoge Rechenregeln gelten für Zeilenvektoren  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w} \in \mathbb{R}_m$  mit  $\mathbf{v} = (v_1, \dots, v_m)$ , wobei wir in (d) stattdessen

$$-\mathbf{v} := (-v_1, \dots, -v_m)$$

nehmen.

**2.31** (Matrixmultiplikation). Für Matrizen  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  und  $B = (b_{jk}) \in \mathbb{R}^{n \times \ell}$  definieren wir  $AB \in \mathbb{R}^{m \times \ell}$  also die  $m \times \ell$ -Matrix  $AB = (c_{ik})$  mit den Einträgen

$$c_{ik} = \sum_{\nu=1}^n a_{i\nu} b_{\nu k}$$

für  $i \in \{1, \dots, m\}$  und  $k \in \{1, \dots, \ell\}$ .

**Beispiel 2.32** Es sei  $A = \mathbf{a} = (a_1, \dots, a_m) \in \mathbb{R}_m$  ein Zeilenvektor und

$$B = \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

ein Spaltenvektor. Dann ist  $AB = \mathbf{a}\mathbf{b}$  eine  $1 \times 1$ -Matrix  $C = (c)$  mit einem reellen Eintrag  $c$ , d.h.  $AB$  ist eine reelle Zahl  $c$ . Die Formel aus 2.31 zeigt, dass  $\mathbf{a}\mathbf{b} = \sum_{\nu=1}^m a_\nu b_\nu$ , also

$$(a_1, \dots, a_m) \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \sum_{\nu=1}^m a_\nu b_\nu$$

(dem Skalarprodukt zweier Spaltenvektoren ähnelnd; vgl. 6.28).

**Bemerkung 2.33** Die Berechnen von Produkten von Matrizen wie in 2.31 lässt sich auf den Spezialfall aus Beispiel 2.32 zurückführen. Schreiben wir in 2.31

$$A = \begin{pmatrix} - & \mathbf{a}_1 & - \\ - & \mathbf{a}_2 & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{pmatrix}$$

mit den Zeilenvektoren  $\mathbf{a}_j := (a_{j1}, \dots, a_{jn}) \in \mathbb{R}_n$  und schreiben wir

$$B = \begin{pmatrix} | & | & \cdots & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_\ell \\ | & | & & | \end{pmatrix}$$

mit den Spaltenvektoren

$$\mathbf{b}_k = \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} \in \mathbb{R}^n,$$

so ist einfach

$$AB = \begin{pmatrix} \mathbf{a}_1 \mathbf{b}_1 & \cdots & \mathbf{a}_1 \mathbf{b}_\ell \\ \vdots & & \vdots \\ \mathbf{a}_m \mathbf{b}_1 & \cdots & \mathbf{a}_m \mathbf{b}_\ell \end{pmatrix},$$

also  $AB = (c_{ik})$  mit  $c_{ik} = \mathbf{a}_i \mathbf{b}_k$ . Mit dieser Idee berechnet man Produkte von Matrizen an der Tafel, auf dem Papier und im Kopf.

**Beispiel 2.34** Es ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{pmatrix} = \begin{pmatrix} (1,2) \begin{pmatrix} 5 \\ 8 \end{pmatrix} & (1,2) \begin{pmatrix} 6 \\ 9 \end{pmatrix} & (1,2) \begin{pmatrix} 7 \\ 10 \end{pmatrix} \\ (3,4) \begin{pmatrix} 5 \\ 8 \end{pmatrix} & (3,4) \begin{pmatrix} 6 \\ 9 \end{pmatrix} & (3,4) \begin{pmatrix} 7 \\ 10 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 21 & 24 & 27 \\ 47 & 54 & 71 \end{pmatrix}$$

**Bemerkung 2.35** Insbesondere können wir das Matrixprodukt bilden aus einer  $m \times n$ -Matrix  $A = (a_{ij})$  und einer Spaltenmatrix  $\mathbf{x} \in \mathbb{R}^n$ , also einem Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Wie erhalten

$$A\mathbf{x} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix},$$

wie bereits in 2.16 (und anschließend) benutzt. Im Kopf würde man dies meist wie folgt berechnen: Schreiben wir  $\mathbf{a}_1, \dots, \mathbf{a}_m$  für die Zeilen von  $A$ , so dass also  $\mathbf{a}_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}_n$ , so ist

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a}_1\mathbf{x} \\ \vdots \\ \mathbf{a}_m\mathbf{x} \end{pmatrix}.$$

**2.36** Rechenregeln für das Multiplizieren von Matrizen und Vektoren. Es sei  $A \in \mathbb{R}^{m \times n}$ . Für alle  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  und alle  $\lambda \in \mathbb{R}$  gilt dann

- (a)  $A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w}$ .
- (b)  $A(\lambda\mathbf{v}) = \lambda(A\mathbf{v})$ .

Dies rechnet man direkt nach durch Hinschreiben in Komponenten. Zum Beispiel ist die  $i$ te Komponente von  $A(\mathbf{v} + \mathbf{w})$  gleich

$$\sum_{j=1}^n a_{ij}(v_j + w_j) = \sum_{j=1}^n a_{ij}v_j + \sum_{j=1}^n a_{ij}w_j$$

und dies stimmt mit der  $i$ ten Komponente von  $A\mathbf{v} + A\mathbf{w}$  überein.

**2.37** (Die Lösungsmenge eines homogenen linearen Gleichungssystems). Gegeben  $A \in \mathbb{R}^{m \times n}$  betrachten wir die Lösungsmenge

$$L_h := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

des homogenen LGS  $A\mathbf{x} = \mathbf{0}$ . Dann gilt:

- (a) Der Nullvektor  $\mathbf{0} \in \mathbb{R}^n$  ist in  $L_h$ .
- (b) Sind  $\mathbf{x}, \mathbf{y} \in L_h$ , so ist auch  $\mathbf{x} + \mathbf{y} \in L_h$ .
- (c) Ist  $\mathbf{x} \in L_h$  und  $\lambda \in \mathbb{R}$ , so ist auch  $\lambda\mathbf{x} \in L_h$ .

Es ist also immer  $\mathbf{0}$  eine Lösung. Summen von Lösungen sind Lösungen und reelle Vielfache von Lösungen sind Lösungen.

Es ist klar, dass  $A\mathbf{0} = \mathbf{0}$ . Sind  $\mathbf{x}, \mathbf{y} \in L_h$ , so ist nach 2.36

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

also  $\mathbf{x} + \mathbf{y} \in L_h$ . Analog sieht man mit 2.36 (b), dass  $\lambda\mathbf{x} \in L_h$ .

**2.38** (Struktur der Lösungsmenge eines inhomogenen linearen Gleichungssystems). Wir betrachten ein inhomogenes lineares Gleichungssystem

$$A\mathbf{x} = \mathbf{b}$$

mit  $A \in \mathbb{R}^{m \times n}$  und  $\mathbf{b} \in \mathbb{R}^m$ , welches eine Lösung besitzt, so dass also die Lösungsmenge

$$L_i := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}\}$$

nicht leer ist. Es existiert also eine Lösung  $\mathbf{x}_p \in L_i$ . Dann gilt

$$L_i = \{\mathbf{x}_p + \mathbf{y} : \mathbf{y} \in L_h\}, \quad (28)$$

wobei

$$L_h := \{\mathbf{y} \in \mathbb{R}^n : A\mathbf{y} = \mathbf{0}\}$$

die Lösungsmenge des zugehörigen homogenen LGS ist.

Kennen wir also irgendeine Lösung  $\mathbf{x}_p$  (die man traditionell eine “partikuläre Lösung” des inhomogenen LGS nennt), so erhalten wir:

*Allgemeine Lösung des inhomogenen LGS*

= *partikuläre Lösung des inhomogenen LGS plus allgemeine Lösung des homogenen LGS.*

Zum Nachweis von (28) sei  $\mathbf{x}$  in der rechten Seite, also  $\mathbf{x} = \mathbf{x}_p + \mathbf{y}$  für ein  $\mathbf{y} \in L_h$ . Dann ist

$$A\mathbf{x} = A(\mathbf{x}_p + \mathbf{y}) = A\mathbf{x}_p + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

also  $\mathbf{x} \in L_i$ . Ist umgekehrt  $\mathbf{x}$  ein Element von  $L_i$ , so ist

$$A(\mathbf{x} - \mathbf{x}_p) = A\mathbf{x} - A\mathbf{x}_p = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

also  $\mathbf{y} := \mathbf{x} - \mathbf{x}_p \in L_h$  und somit  $\mathbf{x} = \mathbf{x}_p + (\mathbf{x} - \mathbf{x}_p) = \mathbf{x}_p + \mathbf{y}$  in der rechten Seite von (28).

**2.39** Es sei  $A \in \mathbb{R}^{m \times n}$  und  $A' \in \mathbb{R}^{m \times n}$  eine zugehörige Zeilenstufenform, vom Rang  $r$ . Dann ist  $0 \leq r \leq \min\{n, m\}$  und es gilt:

- (a) Genau dann sind für jedes  $\mathbf{b} \in \mathbb{R}^m$  die Lösungen des LGS  $A\mathbf{x} = \mathbf{b}$  eindeutig, wenn  $r = n$ .
- (b) Genau dann existiert für jedes  $\mathbf{b} \in \mathbb{R}^m$  eine Lösung für  $A\mathbf{x} = \mathbf{b}$ , wenn  $r = m$ .
- (c) Genau dann ist das LGS  $A\mathbf{x} = \mathbf{b}$  für alle  $\mathbf{b} \in \mathbb{R}^m$  eindeutig lösbar, wenn  $r = n = m$ .

Der Rang ist die Zahl nicht verschwindender Zeilen der Zeilenstufenform  $A'$ , also kleiner gleich der Gesamtzahl der Zeilen. Somit ist  $r \leq m$ . Da jede der genannten Zeilen an einer anderen Stelle ein  $\blacksquare$  enthält, ist  $r \leq n$ .

Begründung für (a): Die Lösungen sind genau dann eindeutig, wenn es keine freien Variablen gibt, also  $r = n$  ist.

(b) Ist  $r = m$ , so ist  $(A', \mathbf{d})$  von Zeilen-Stufenform und hat den Rang  $m = r$ , für alle  $\mathbf{d} \in \mathbb{R}^m$ . Für gegebenes  $\mathbf{b} \in \mathbb{R}^m$  entsteht solche eine Matrix durch elementare Zeilenumformungen aus der erweiterten Koeffizientenmatrix  $(A|\mathbf{b})$ . Nach dem Lösbarkeitstest aus 2.26 ist  $A\mathbf{x} = \mathbf{b}$  also lösbar. Ist  $r < m$ , so ist  $A'\mathbf{x} = \mathbf{d}$  nicht lösbar für  $\mathbf{d} := \mathbf{e}_{r+1} \in \mathbb{R}^m$ , den Vektor, der eine 1 in der  $(r + 1)$ ten Komponente hat und sonst Nullen. Machen wir rückwärts die elementaren Zeilenoperationen rückgängig, die von  $A$  zu  $A'$  geführt haben, kommen wir von  $(A'|\mathbf{d})$  zu  $(A, \mathbf{b})$  mit einem  $\mathbf{b} \in \mathbb{R}^m$ . Da die linearen Gleichungssysteme  $A\mathbf{x} = \mathbf{b}$  und  $A'\mathbf{x} = \mathbf{d}$  äquivalent sind, hat auch  $A\mathbf{x} = \mathbf{b}$  keine Lösung.

(c) kombiniert (a) und (b).

### 3 Vektorräume und lineare Abbildungen, invertierbare Matrizen

Wir kennen bereits

- Vektoren in der Ebene
- Vektoren im Raum
- Spaltenvektoren in  $\mathbb{R}^m$
- Zeilenvektoren in  $\mathbb{R}_n$ .

Als ein gemeinsames Dach für diese (und neue) Beispiele führen wir nun *Vektorräume* ein. Die Elemente eines Vektorraums (die Vektoren) können wir addieren und mit reellen Zahlen (Skalaren) multiplizieren. Dabei sollen die Rechenregeln gelten, die wir aus den obigen Beispielen kennen.

#### Vektorräume und Untervektorräume

**Definition 3.1** Ein Vektorraum ist eine Menge  $V$  (z.B.  $V = \mathbb{R}^m$ ), auf der eine Summe

$$\mathbf{v} + \mathbf{w} \in V$$

definiert ist für alle  $\mathbf{v}, \mathbf{w} \in V$  und das Produkt

$$\lambda \mathbf{v} \in V$$

definiert ist für  $\mathbf{v} \in V$  und  $\lambda \in \mathbb{R}$ , so dass die übliche Rechenregeln (wie in  $\mathbb{R}^m$  in 2.29) gelten:

- Für alle  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  ist  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$  (Assoziativgesetz);
- Für alle  $\mathbf{u}, \mathbf{v} \in V$  ist  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$  (Kommutativgesetz);
- Es gibt ein  $\mathbf{0} \in V$  derart, dass  $\mathbf{0} + \mathbf{v} = \mathbf{v}$  für alle  $\mathbf{v} \in V$ ;
- Für jedes  $\mathbf{v} \in V$  gibt es ein  $-\mathbf{v} \in V$  mit  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ ;
- $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$  für alle  $\lambda, \mu \in \mathbb{R}$  und  $\mathbf{v} \in V$ ;
- $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$  für alle  $\lambda \in \mathbb{R}$  und  $\mathbf{v}, \mathbf{w} \in V$ ;

(g)  $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$  für alle  $\lambda, \mu \in \mathbb{R}$  und  $\mathbf{v} \in V$ ; und

(h)  $1\mathbf{v} = \mathbf{v}$  für alle  $\mathbf{v} \in V$ .

Die Elemente  $\mathbf{v} \in V$  nennen wir *Vektoren*, die Zahlen  $\lambda \in \mathbb{R}$  *Skalare*.

Der Nullvektor  $\mathbf{0} \in V$  in (c) ist hierbei eindeutig festgelegt, denn ist auch  $\mathbf{0}'$  ein solcher, so ist

$$\mathbf{0} + \mathbf{0}' = \mathbf{0}'$$

nach (c) und

$$\mathbf{0} + \mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0},$$

unter Benutzung von (c) für  $\mathbf{0}'$  an Stelle von  $\mathbf{0}$ . Also ist  $\mathbf{0}' = \mathbf{0}$ .

Ähnliche kleine Rechnungen zeigen, dass  $-\mathbf{v}$  durch (d) eindeutig festgelegt ist;<sup>4</sup> zudem gilt<sup>5</sup>

$$0\mathbf{v} = \mathbf{0}$$

und

$$-\mathbf{v} = (-1)\mathbf{v},$$

da  $(-1)\mathbf{v} + \mathbf{v} = ((-1) + 1)\mathbf{v} = 0\mathbf{v} = \mathbf{0}$ .

**Beispiel 3.2** (a)  $\mathbb{R}^m$  ist ein Vektorraum für alle  $m \in \mathbb{N}_0$  (wobei  $\mathbb{R}^0 := \{0\}$ ), denn (a)–(h) wurden in 2.29 nachgeprüft. Entsprechend gilt:

(b)  $\mathbb{R}_n$  ist ein Vektorraum für alle  $n \in \mathbb{N}_0$ .

**Definition 3.3** Ist  $V$  ein Vektorraum, so nennen wir eine Teilmenge  $W \subseteq V$  einen *Untervektorraum*, wenn gilt:

(a) Für alle  $\mathbf{v}, \mathbf{w} \in W$  ist  $\mathbf{v} + \mathbf{w} \in W$  (Abgeschlossenheit unter Addition);

(b) Für alle  $\mathbf{w} \in W$  und  $\lambda \in \mathbb{R}$  ist  $\lambda\mathbf{w} \in W$  (Abgeschlossenheit unter Multiplikation mit Skalaren);

(c) Es ist  $\mathbf{0} \in W$ .

**Beispiel 3.4** Wir betrachten ein homogenes lineares Gleichungssystem  $A\mathbf{x} = \mathbf{0}$  mit  $A \in \mathbb{R}^{m \times n}$ . Die Rechnung in 2.37 zeigt, dass die Lösungsmenge  $L_h$  des LGS ein Untervektorraum von  $\mathbb{R}^n$  ist.

<sup>4</sup>Ist  $\mathbf{v} + \mathbf{w} = \mathbf{0}$ , so zeigt Addition von  $-\mathbf{v}$  auf beiden Seiten, dass  $\mathbf{w} = -\mathbf{v}$ .

<sup>5</sup>Hierfür addieren wir  $-\mathbf{v}$  auf beiden Seiten von  $0\mathbf{v} + 0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v}$ .

**3.5** Ist  $V$  ein Vektorraum und  $W \subseteq V$  ein Untervektorraum, so ist  $W$  ein Vektorraum unter Benutzung der Addition

$$\mathbf{v} + \mathbf{w}$$

für  $\mathbf{v}, \mathbf{w} \in W$  wie in  $V$  und der Multiplikation

$$\lambda \mathbf{w}$$

für  $\lambda \in \mathbb{R}$  und  $\mathbf{w} \in W$  wie in  $V$ .

**Bemerkung 3.6** In der Praxis wollen wir nie von Hand nachrechnen müssen, dass eine gegebene Menge  $W$  ein Vektorraum ist. So gut wie immer ist  $W$  einfach ein Untervektorraum eines Vektorraums, den wir schon kennen, und somit  $W$  nach 3.5 selbst ein Vektorraum.

Rechnungen sind nur noch einmal nötig, für das folgende Beispiel. Mit Summen  $f + g$  von Funktionen und Vielfache  $\lambda f$  gehen wir seit Langem um. Wir erinnern daran:

**Definition 3.7** Ist  $X$  eine Menge, so schreiben wir  $\mathbb{R}^X$  für die Menge aller Funktionen  $f: X \rightarrow \mathbb{R}$ . Für  $f, g \in \mathbb{R}^X$  definieren wir die Funktion  $f + g: X \rightarrow \mathbb{R}$  via

$$(f + g)(x) := f(x) + g(x) \quad \text{für } x \in X.$$

Für  $\lambda \in \mathbb{R}$  und  $f \in \mathbb{R}^X$  definieren wir die Funktion  $\lambda f: X \rightarrow \mathbb{R}$  via

$$(\lambda f)(x) := \lambda f(x)$$

Weiter sei  $0: X \rightarrow \mathbb{R}, x \mapsto 0$  die konstante Funktion mit Wert 0 und  $-f$  die Funktion  $X \rightarrow \mathbb{R}, x \mapsto -f(x)$ .

**Beispiel 3.8** Für jede Menge  $X$  ist  $\mathbb{R}^X$  ein Vektorraum. Die konstante Nullfunktion  $0: X \rightarrow \mathbb{R}$  ist der Nullvektor. Für  $f \in \mathbb{R}^X$  ist  $-f: x \mapsto -f(x)$  die Funktion mit  $f + (-f) = 0$  (wie im Vektorraumaxiom (d) verlangt).

Die Vektorraumaxiome (a)–(h) rechnet man punktweise nach (ähnlich, wie diejenigen für  $\mathbb{R}^m$  komponentenweise nachgerechnet wurden). Zum Beispiel gilt für alle  $f, g, h \in \mathbb{R}^X$

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) = (f + (g + h))(x) \end{aligned}$$

für alle  $x \in X$  und somit das Assoziativitätsgesetz  $(f + g) + h = f + (g + h)$ .  
Für alle  $f, g \in \mathbb{R}^X$  gilt

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$$

für alle  $x \in X$ , somit das Kommutativgesetz  $f + g = g + f$  (usw.)

Viele interessante Vektorräume bestehen aus Funktionen und sind Untervektorräume von  $\mathbb{R}^X$ .

**Beispiel 3.9** (a) Für reelle Zahlen  $a < b$  ist die Menge  $C[a, b]$  aller stetigen Funktionen  $f: [a, b] \rightarrow \mathbb{R}$  ein Untervektorraum des Vektorraums  $\mathbb{R}^{[a, b]}$  aller Funktionen  $f: [a, b] \rightarrow \mathbb{R}$ . Denn sind  $f: [a, b] \rightarrow \mathbb{R}$  und  $g: [a, b] \rightarrow \mathbb{R}$  stetig, so auch  $f + g$  und  $\lambda f$  für alle  $\lambda \in \mathbb{R}$ . Weiter ist die Nullfunktion  $0: [a, b] \rightarrow \mathbb{R}$ ,  $x \mapsto 0$  stetig, also in  $C[a, b]$ .

(b) Ebenso sieht man, dass die Menge  $C^1[a, b]$  aller stetig differenzierbaren<sup>6</sup> Funktionen  $f: [a, b] \rightarrow \mathbb{R}$  einen Untervektorraum von  $\mathbb{R}^{[a, b]}$  (und von  $C[a, b]$ ) bildet.

(c) Die Menge  $W$  aller Polynome  $p: \mathbb{R} \rightarrow \mathbb{R}$  der Form  $p(x) = \sum_{k=0}^n a_k x^k$  (mit einem  $n \in \mathbb{N}_0$  und Koeffizienten  $a_0, \dots, a_n \in \mathbb{R}$ ) ist ein Untervektorraum des Vektorraums  $\mathbb{R}^{\mathbb{R}}$  aller Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$ , denn Summen  $p + q$  von Polynomen  $p$  und  $q$  sind wieder Polynome und auch Vielfache  $\lambda p$  mit  $\lambda \in \mathbb{R}$ .

(d) Analog sieht man: Für festes  $n \in \mathbb{N}_0$  ist die Menge  $W$  aller Polynome  $p: \mathbb{R} \rightarrow \mathbb{R}$  vom Grad  $\leq n$  ein Untervektorraum von  $\mathbb{R}^{\mathbb{R}}$ .

## Basen und lineare Unabhängigkeit

**Definition 3.10** Es sei  $V$  ein Vektorraum. Eine *Linearkombination* von Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  ist eine Summe der Form

$$\sum_{k=1}^n \lambda_k \mathbf{v}_k = \lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n \in V$$

mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ .

---

<sup>6</sup>Eine Funktion auf  $[a, b]$  wurde *stetig differenzierbar* genannt, wenn sie überall differenzierbar ist und  $f': [a, b] \rightarrow \mathbb{R}$ ,  $x \mapsto f'(x)$  eine stetige Funktion.

**Beispiel 3.11** In  $\mathbb{R}^2$  ist jeder Vektor

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

eine Linearkombination der Standard-Basisvektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Für  $\lambda_1, \lambda_2 \in \mathbb{R}$  ist nämlich

$$\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

genau dann gleich  $\mathbf{v}$ , wenn  $\lambda_1 = v_1$  und  $\lambda_2 = v_2$ .

**Definition 3.12** Ist  $V$  ein Vektorraum, so nennt man Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine *Basis* von  $V$ , wenn sich jeder Vektor  $\mathbf{v} \in V$  als Linearkombination

$$\mathbf{v} = \sum_{k=1}^n \lambda_k \mathbf{b}_k$$

der Basisvektoren schreiben lässt mit eindeutig festgelegten Zahlen  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ .

Zum Beispiel ist  $\mathbf{e}_1, \mathbf{e}_2$  eine Basis für  $\mathbb{R}^2$ , wie die Rechnung in Beispiel 3.11 zeigt. Analog gilt:

**Beispiel 3.13** Die Vektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

bilden eine Basis für  $\mathbb{R}^n$ , denn jeder Vektor

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

ist von der Form

$$\mathbf{v} = v_1 \mathbf{e}_1 + \cdots + v_n \mathbf{e}_n$$

und legt die Koeffizienten  $v_1, \dots, v_n$  fest. Wir nennen  $\mathbf{e}_1, \dots, \mathbf{e}_n$  die *Standard-Basis* für  $\mathbb{R}^n$ .

Wir wollen sehen:

**3.14** *Jede andere Basis von  $\mathbb{R}^n$  hat ebenfalls  $n$  Elemente.*

Beim Nachweis hilft uns der folgende Begriff.

**Definition 3.15** Es sei  $V$  ein Vektorraum. Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  werden *linear unabhängig* genannt, wenn aus

$$\sum_{k=1}^n \lambda_k \mathbf{v}_k = \mathbf{0}$$

mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  stets folgt, dass  $\lambda_1 = \cdots = \lambda_n = 0$ .

Der Nullvektor darf also nur auf triviale Art als Linearkombination der  $\mathbf{v}_1, \dots, \mathbf{v}_n$  darstellbar sein.

Sind  $\mathbf{v}_1, \dots, \mathbf{v}_n$  nicht linear unabhängig, so werden  $\mathbf{v}_1, \dots, \mathbf{v}_n$  *linear abhängig* genannt.

**Beispiel 3.16** Für  $n \in \mathbb{N}_0$  sind die Monome  $f_0, \dots, f_n: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f_k(x) := x^k$  im Vektorraum  $V$  aller Polynome linear unabhängig.

Seien nämlich  $\lambda_0, \dots, \lambda_n \in \mathbb{R}$  mit

$$p := \sum_{k=0}^n \lambda_k f_k = 0,$$

also  $p(x) = \sum_{k=0}^n \lambda_k x^k$  für alle  $x \in \mathbb{R}$ . Für  $j \in \{0, \dots, n\}$  ergibt sich für die  $j$ te Ableitung von  $p$

$$0 = p^{(j)}(0) = \sum_{k=0}^n \lambda_k \underbrace{\frac{d^j}{dx^j} \Big|_{x=0} x^k}_{=j! \delta_{j,k}} = j! \lambda_j$$

nach (4) in Kapitel 1; also ist  $\lambda_j = 0$  für alle  $J \in \{0, 1, \dots, k\}$ .

**3.17** Es seien  $V$  ein Vektorraum und  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ . Dann sind äquivalent:

- (a)  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sind linear unabhängig.
- (b) Koeffizienten in Linearkombinationen sind eindeutig: Sind  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  und  $\mu_1, \dots, \mu_n \in \mathbb{R}$  mit

$$\sum_{k=1}^n \lambda_k \mathbf{v}_k = \sum_{k=1}^n \mu_k \mathbf{v}_k \quad (29)$$

so folgt  $\lambda_j = \mu_j$  für alle  $j \in \{1, \dots, n\}$ .

Aus (a) folgt (b): Seien  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig. Gilt (29), so ist

$$\mathbf{0} = \left( \sum_{k=1}^n \lambda_k \mathbf{v}_k \right) - \left( \sum_{k=1}^n \mu_k \mathbf{v}_k \right) = \sum_{k=1}^n (\lambda_k - \mu_k) \mathbf{v}_k.$$

Da  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig angenommen sind, folgt  $\lambda_k - \mu_k = 0$  für alle  $k \in \{1, \dots, n\}$  und somit  $\lambda_k = \mu_k$ .

Ist (a) falsch, so auch (b): Ist (a) falsch, so gibt es  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ , die nicht alle 0 sind, so dass

$$\sum_{k=1}^n \lambda_k \mathbf{v}_k = \mathbf{0}.$$

Es ist aber auch  $\sum_{k=1}^n \mu_k \mathbf{v}_k = \mathbf{0}$ , wenn wir  $\mu_k := 0$  setzen für alle  $k \in \{1, \dots, n\}$ .

**Beispiel 3.18** Ist  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine Basis für einen Vektorraum  $V$ , so sind die Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_n$  linear unabhängig.

Per Definition einer Basis sind die Koeffizienten in Linearkombinationen der  $\mathbf{b}_1, \dots, \mathbf{b}_n$  nämlich eindeutig. Also sind  $\mathbf{b}_1, \dots, \mathbf{b}_n$  linear unabhängig nach “(b) $\Rightarrow$ (a)” in 3.17.

Wir zeigen nun insbesondere 3.14.

**3.19** Für linear unabhängige Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  in  $\mathbb{R}^m$  gilt:

- (a) Es ist  $n \leq m$ .
- (b) Genau dann ist  $n = m$ , wenn  $\mathbf{v}_1, \dots, \mathbf{v}_n$  eine Basis für  $\mathbb{R}^m$  bilden.

Nachweis für (a): Es sei  $A \in \mathbb{R}^{m \times n}$  die Matrix mit  $\mathbf{v}_1, \dots, \mathbf{v}_n$  als Spaltenvektoren. Für alle  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  ist dann also

$$\sum_{k=1}^n \lambda_k \mathbf{v}_k = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \quad (30)$$

Da  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^m$  linear unabhängig angenommen sind, sind nach 3.17 (b) die Lösungen von

$$A\mathbf{x} = \mathbf{b}$$

für jedes  $\mathbf{b} \in \mathbb{R}^m$  eindeutig (falls existent). Ist  $r$  der Rang einer zu  $A$  gehörigen Zeilenstufenform  $A'$ , so ist nach 2.39 (a) also  $n = r$  und somit  $n \leq m$ , da stets  $r \leq m$ .

Nachweis für (b): Es sei  $A'$  eine zu  $A$  gehörige Zeilenstufenform und  $r$  ihre Rang. da wir  $\mathbf{v}_1, \dots, \mathbf{v}_n$  als linear unabhängig annehmen, wissen wir aus dem Beweis von (a) bereits, dass  $r = n$ . Wegen (30) bilden die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  genau dann eine Basis für  $\mathbb{R}^m$ , wenn das LGS

$$A\mathbf{x} = \mathbf{b}$$

für jedes  $\mathbf{b} \in \mathbb{R}^m$  eine eindeutige Lösung besitzt. Nach 2.39 (c) ist dies genau dann der Fall, wenn  $n = r = m$ . Da wir bereits wissen, dass  $n = r$  ist, ist letztere Bedingung zu  $n = m$  äquivalent.

Die folgende Charakterisierung ist manchmal nützlich.

**3.20** *Es seien  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängige Vektoren in einem Vektorraum  $V$ . Dann sind äquivalent:*

- (a)  $\mathbf{v}_1, \dots, \mathbf{v}_n$  ist eine Basis für  $V$ ;
- (b)  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sind in  $V$  **maximal linear unabhängig**, d.h. für jedes  $\mathbf{v}_{n+1} \in V$  sind  $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$  linear abhängig.

(a) $\Rightarrow$ (b): Da  $\mathbf{v}_1, \dots, \mathbf{v}_n$  eine Basis ist, gibt es  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  derart, dass

$$\mathbf{v}_{n+1} = \sum_{j=1}^n \lambda_j \mathbf{v}_j.$$

Dann ist also

$$\mathbf{0} = \left( \sum_{j=1}^n \lambda_j \mathbf{v}_j \right) - \mathbf{v}_{n+1} = \sum_{j=1}^{n+1} \lambda_j \mathbf{v}_j$$

mit  $\lambda_{n+1} := -1$ . Da nicht alle der  $\lambda_1, \dots, \lambda_{n+1}$  gleich 0 sind, sind  $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$  linear abhängig.

(b) $\Rightarrow$ (a): ist  $\mathbf{v}_{n+1} \in V$  beliebig, so gibt es nach (b) Zahlen  $\lambda_1, \dots, \lambda_{n+1} \in \mathbb{R}$ , die nicht alle 0 sind, so dass

$$\mathbf{0} = \sum_{j=1}^{n+1} \lambda_j \mathbf{v}_j.$$

Wäre  $\lambda_{n+1} = 0$ , so würde aus  $\mathbf{0} = \sum_{j=1}^{n+1} \lambda_j \mathbf{v}_j = \sum_{j=1}^n \lambda_j \mathbf{v}_j$  folgen, dass  $\lambda_1 = \dots = \lambda_n = 0$ , da  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig angenommen sind. Andernfalls wären doch alle  $\lambda_j = 0$ , Widerspruch. Also muss doch  $\lambda_{n+1} \neq 0$  sein. Folglich ist

$$\mathbf{v}_{n+1} = \sum_{j=1}^n \left( -\frac{\lambda_j}{\lambda_{n+1}} \right) \mathbf{v}_j$$

eine Linearkombination der  $\mathbf{v}_1, \dots, \mathbf{v}_n$  und da letztere Vektoren linear unabhängig sind, sind die Koeffizienten in der Linearkombination eindeutig (nach 3.17). Also ist  $\mathbf{v}_1, \dots, \mathbf{v}_n$  eine Basis für  $V$ .

## Nachprüfen von linearer Unabhängigkeit

Seien  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^m$  und  $A \in \mathbb{R}^{m \times n}$  die Matrix mit den Spalten  $\mathbf{v}_1, \dots, \mathbf{v}_n$ . Es sind  $\mathbf{v}_1, \dots, \mathbf{v}_n$  genau dann linear unabhängig, wenn

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

nur die Lösung  $\lambda_1 = \dots = \lambda_n = 0$  hat, die Zahl  $n - r$  der freien Variablen einer Zeilenstufenform  $A'$  also 0 ist, wobei  $r$  der Rang von  $A'$  ist. Lineare Unabhängigkeit ist also äquivalent zu  $r = n$  (vergleiche auch 2.39 (a)). Dies liefert ein praktikables Verfahren zum Nachweis von linearer Unabhängigkeit, indem man mit elementaren Zeilenoperationen von  $A$  zu  $A'$  übergeht.

**Beispiel 3.21** Sind die  $n = 3$  Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

in  $\mathbb{R}^3$  linear unabhängig?

Wir bringen die  $3 \times 3$ -Matrix mit den genannten Vektoren als Spalten durch elementare Zeilenoperationen auf Zeilenstufenform:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{array}{l} \downarrow \\ \downarrow \\ \uparrow \end{array} \begin{array}{l} (-1) \\ \\ (-1) \end{array} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{array}{l} \downarrow \\ \downarrow \\ \uparrow \end{array} \begin{array}{l} (-1) \\ \\ (-1) \end{array} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Letztere Zeilenstufenform hat den Rang  $r = 3$ . Da  $r = n$  ist, sind die Vektoren linear unabhängig.

## Lineare Abbildungen und Matrizen; inverse Matrix

**Definition 3.22** Es seien  $V$  und  $W$  Vektorräume. Eine Abbildung  $\phi: V \rightarrow W$  wird *linear* genannt, wenn gilt:

- (a)  $\phi(\mathbf{v} + \mathbf{w}) = \phi(\mathbf{v}) + \phi(\mathbf{w})$  für alle  $\mathbf{v}, \mathbf{w} \in V$ .
- (b)  $\phi(\lambda \mathbf{v}) = \lambda \phi(\mathbf{v})$  für alle  $\mathbf{v} \in V$  und  $\lambda \in \mathbb{R}$ .

**Beispiel 3.23** Für jede Matrix  $A \in \mathbb{R}^{m \times n}$  ist die zugehörige Abbildung

$$\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{x} \mapsto A\mathbf{x}$$

linear (siehe 2.36).

**Beispiel 3.24** Eine  $2 \times 2$ -Matrix der Form

$$A = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

mit  $\phi \in \mathbb{R}$  wird *Drehmatrix* genannt, denn die zugehörige lineare Abbildung

$$\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cos \phi - y \sin \phi \\ x \sin \phi + y \cos \phi \end{pmatrix}$$

beschreibt eine Drehung der Ebene um den Ursprung im Gegenuhrzeigersinn um den Winkel  $\phi$ .

**3.25** Für jede lineare Abbildung  $\phi: V \rightarrow W$  zwischen Vektorräumen gilt

$$\phi(\mathbf{0}) = \mathbf{0}$$

und

$$\phi(-\mathbf{v}) = -\phi(\mathbf{v}) \quad \text{für alle } \mathbf{v} \in V.$$

Es ist nämlich

$$\phi(\mathbf{0}) = \phi(0\mathbf{0}) = 0\phi(\mathbf{0}) = \mathbf{0}$$

und

$$\phi(-\mathbf{v}) = \phi((-1)\mathbf{v}) = (-1)\phi(\mathbf{v}) = -\phi(\mathbf{v}).$$

Wir wollen erste Einblicke in das Zusammenspiel von Matrizen und linearen Abbildungen bekommen. Halten wir zunächst fest, dass das Multiplizieren von Matrizen assoziativ ist:

**3.26** Es sei  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ ,  $B = (b_{jk}) \in \mathbb{R}^{n \times p}$ ,  $C = (c_{k\ell}) \in \mathbb{R}^{p \times q}$  mit  $n, m, p, q \in \mathbb{N}$ . Dann gilt

$$(AB)C = A(BC).$$

Beweis. Der  $(i, k)$ -Eintrag von  $AB$  ist  $d_{ik} := \sum_{j=1}^n a_{ij}b_{jk}$ . Der  $(i, \ell)$ -Eintrag von  $(AB)C$  ist somit

$$\sum_{k=1}^p d_{ik}c_{k\ell} = \sum_{j=1}^n \sum_{k=1}^p a_{ij}b_{jk}c_{k\ell}. \quad (31)$$

Der  $(j, \ell)$ -Eintrag von  $BC$  ist

$$e_{j,\ell} := \sum_{k=1}^p b_{jk}c_{k\ell}.$$

Der  $(i, \ell)$ -Eintrag von  $A(BC)$  ist also

$$\sum_{j=1}^n a_{ij}e_{j\ell} = \sum_{j=1}^n \sum_{k=1}^p a_{ij}b_{jk}c_{k\ell},$$

was mit (31) übereinstimmt.

Das Produkt von Matrizen entspricht dem Komponieren der entsprechenden linearen Abbildungen.

**3.27** (a) Für alle  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{n \times k}$  gilt

$$\phi_A \circ \phi_B = \phi_{AB}.$$

(b) Für die Einheitsmatrix  $\mathbf{1}_n \in \mathbb{R}^{n \times n}$  gilt  $\phi_{\mathbf{1}_n} = \text{id}_{\mathbb{R}^n}$ .

(c) Es habe  $A \in \mathbb{R}^{m \times n}$  die Spaltenvektoren  $\mathbf{s}_1, \dots, \mathbf{s}_n \in \mathbb{R}^m$ . Für die Standard-Basisvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_n$  von  $\mathbb{R}^n$  gilt dann

$$\phi_A(\mathbf{e}_j) = \mathbf{s}_j$$

für alle  $j \in \{1, \dots, n\}$ . Insbesondere ist  $A$  durch  $\phi := \phi_A$  eindeutig festgelegt: Es ist

$$A = (\phi(\mathbf{e}_1) \cdots \phi(\mathbf{e}_n)).$$

Nachweis: (a) Für alle  $\mathbf{v} \in \mathbb{R}^k$  gilt wegen des Assoziativgesetzes für die Matrixmultiplikation

$$(\phi_A \circ \phi_B)(\mathbf{v}) = A(B\mathbf{v}) = (AB)\mathbf{v} = \phi_{AB}(\mathbf{v}).$$

(b) Für alle  $\mathbf{v} \in \mathbb{R}^n$  gilt  $\phi_{\mathbf{1}_n}(\mathbf{v}) = \mathbf{1}_n \mathbf{v} = \mathbf{v} = \text{id}_V(\mathbf{v})$ .

(c) Es ist

$$A\mathbf{x} = \sum_{i=1}^n \mathbf{s}_i x_i$$

für  $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ . Im Falle  $\mathbf{x} = \mathbf{e}_j$  ist  $x_i = \delta_{ij}$ , also

$$A\mathbf{e}_j = \sum_{i=1}^n \mathbf{s}_i \delta_{ij} = \mathbf{s}_j.$$

Wir zeigen nun, dass jede lineare Abbildung  $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  von der Form

$$\psi = \phi_A \tag{32}$$

ist für eine Matrix  $A \in \mathbb{R}^{m \times n}$ . Die Matrix  $A$  ist durch die Bedingung (32) dann eindeutig festgelegt, nach 3.27 (c).

**3.28** Es sei  $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung. Sind  $\mathbf{e}_1, \dots, \mathbf{e}_n$  die Standard-Basisvektoren für  $\mathbb{R}^n$ , so setzen wir

$$\mathbf{a}_j := \psi(\mathbf{e}_j) \in \mathbb{R}^m$$

für  $j \in \{1, \dots, n\}$  und  $A := (\mathbf{a}_1 \cdots \mathbf{a}_n) \in \mathbb{R}^{m \times n}$ . Dann gilt  $\psi = \phi_A$ , wobei  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \mapsto A\mathbf{x}$  die zu  $A$  gehörige lineare Abbildung ist.

Begründung: Für alle  $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  gilt wegen  $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j$

$$\phi_A(\mathbf{x}) = \phi_A\left(\sum_{j=1}^n x_j \mathbf{e}_j\right) = \sum_{j=1}^n x_j A \mathbf{e}_j = \sum_{j=1}^n x_j \mathbf{a}_j = \sum_{j=1}^n x_j \psi(\mathbf{e}_j) = \psi\left(\sum_{j=1}^n x_j \mathbf{e}_j\right) = \psi(\mathbf{x}).$$

Also ist  $\psi = \phi_A$ .

**3.29** Ist eine lineare Abbildung  $\psi: V \rightarrow W$  zwischen Vektorräumen  $V$  und  $W$  invertierbar, so ist auch  $\psi^{-1}: W \rightarrow V$  linear.

Beweis. Für  $\mathbf{w}_1, \mathbf{w}_2 \in W$  sei  $\mathbf{v}_1 := \psi^{-1}(\mathbf{w}_1)$  und  $\mathbf{v}_2 := \psi^{-1}(\mathbf{w}_2)$ . Dann gilt

$$\psi(\mathbf{v}_1 + \mathbf{v}_2) = \psi(\mathbf{v}_1) + \psi(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2.$$

Anwendung von  $\psi^{-1}$  liefert

$$\mathbf{v}_1 + \mathbf{v}_2 = \psi^{-1}(\mathbf{w}_1 + \mathbf{w}_2),$$

wobei die linke Seite gleich  $\psi^{-1}(\mathbf{w}_1) + \psi^{-1}(\mathbf{w}_2)$  ist. Gegeben  $\mathbf{w} \in W$  und  $\lambda \in \mathbb{R}$  gilt

$$\psi(\lambda \psi^{-1}(\mathbf{w})) = \lambda \psi(\psi^{-1}(\mathbf{w})) = \lambda \mathbf{w}.$$

Anwendung von  $\psi^{-1}$  liefert  $\lambda \psi^{-1}(\mathbf{w}) = \psi^{-1}(\lambda \mathbf{w})$ .

**Definition 3.30** Eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *invertierbar*, wenn es eine Matrix  $B \in \mathbb{R}^{n \times n}$  derart gibt, dass

$$AB = BA = \mathbf{1}_n$$

die  $n \times n$ -Einheitsmatrix ist. Dann ist  $B$  eindeutig festgelegt;<sup>7</sup> wir schreiben  $A^{-1} := B$ . Es ist also  $AA^{-1} = A^{-1}A = \mathbf{1}_n$ .

Da  $\mathbf{1}_n \mathbf{1}_n = \mathbf{1}_n$ , ist die Einheitsmatrix  $\mathbf{1}_n$  invertierbar.

**Beispiel 3.31** Ist eine  $n \times n$ -Matrix  $A$  invertierbar, so können Sie das lineare Gleichungssystem

$$A\mathbf{x} = \mathbf{b} \tag{33}$$

für jedes  $\mathbf{b} \in \mathbb{R}^m$  ganz einfach lösen: Multiplikation mit  $A^{-1}$  von links führt auf die notwendige Bedingung

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}.$$

Diese ist auch hinreichend für (33), denn Multiplikation mit  $A$  von links führt zurück zu (33).

<sup>7</sup>Ist auch  $AC = CA = \mathbf{1}_n$ , so ist  $C = \mathbf{1}_n C = BAC = B\mathbf{1}_n = B$ .

**3.32** (a) Ist eine Matrix  $A \in \mathbb{R}^{n \times n}$  invertierbar, so ist auch  $A^{-1}$  invertierbar mit  $(A^{-1})^{-1} = A$ .

[Denn für  $B := A$  gilt  $BA^{-1} = A^{-1}B = \mathbf{1}_n$ .]

(b) Sind die Matrizen  $A \in \mathbb{R}^{n \times n}$  und  $B \in \mathbb{R}^{n \times n}$  beide invertierbar, so ist auch die Matrix  $AB$  invertierbar und es ist  $(AB)^{-1} = B^{-1}A^{-1}$ .

[Es gilt  $B^{-1}A^{-1}AB = B^{-1}\mathbf{1}_nB = B^{-1}B = \mathbf{1}_n$  und analog  $ABB^{-1}A^{-1} = \mathbf{1}_n$ .]

(c) Eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann invertierbar, wenn die zugehörige lineare Abbildung  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$  invertierbar (also bijektiv) ist.

[Ist  $A$  invertierbar, so ist  $\phi_{A^{-1}} \circ \phi_A = \phi_{A^{-1}A} = \phi_{\mathbf{1}_n} = \text{id}_{\mathbb{R}^n}$  und analog  $\phi_A \circ \phi_{A^{-1}} = \text{id}_{\mathbb{R}^n}$ , so dass  $\phi_A$  invertierbar ist. Sei umgekehrt  $\phi_A$  invertierbar. Nach 3.29 ist dann auch  $\phi_A^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear. Nach 3.28 gibt es also eine  $n \times n$ -Matrix  $B$  mit  $(\phi_A)^{-1} = \phi_B$ . Aus

$$\phi_{\mathbf{1}_n} = \text{id}_{\mathbb{R}^n} = (\phi_A)^{-1} \circ \phi_A = \phi_B \circ \phi_A = \phi_{BA}$$

folgt  $\mathbf{1}_n = BA$ . Analog sieht man, dass  $\mathbf{1}_n = AB$ , so dass also  $A$  invertierbar ist mit  $A^{-1} = B$ .]

**Beispiel 3.33** Eine Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$  mit Diagonaleinträgen  $\lambda_1, \dots, \lambda_n$  ist invertierbar, wenn alle Diagonaleinträge  $\lambda_1, \dots, \lambda_n$  von 0 verschieden sind.

Dann ist nämlich

$$B := \text{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)$$

eine  $n \times n$ -Matrix derart, dass  $DB = \text{diag}(\lambda_1 \frac{1}{\lambda_1}, \dots, \lambda_n \frac{1}{\lambda_n}) = \text{diag}(1, \dots, 1) = \mathbf{1}_n$  und analog  $BD = \mathbf{1}_n$ .

Insbesondere ist für  $i \in \{1, \dots, n\}$  und  $\lambda \in \mathbb{R} \setminus \{0\}$  die Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  mit  $\lambda_i := \lambda$ ,  $\lambda_j := 1$  für  $j \in \{1, \dots, n\} \setminus \{i\}$  invertierbar.

**Beispiel 3.34** Gegeben  $n \in \mathbb{N}$  und  $i, j \in \{1, \dots, n\}$  mit  $i < j$  können wir die lineare Abbildung  $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,

$$(x_1, \dots, x_n)^T \mapsto (x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n)^T$$

betrachten, welche die  $i$ te und  $j$ te Komponente von  $x = (x_1, \dots, x_n)^T$  vertauscht und alle anderen festhält. Dann ist offenbar  $\psi \circ \psi = \text{id}_{\mathbb{R}^n}$ , also

$\psi$  invertierbar. Nach 3.28 ist  $\psi = \phi_P$  für eine  $n \times n$ -Matrix  $P$ . Aus  $\phi_{P^2} = \phi_P \circ \phi_P = \psi \circ \psi = \text{id}_{\mathbb{R}^n} = \phi_{\mathbf{1}_n}$  folgt  $P \circ P = \mathbf{1}_n$ . Also ist  $P$  invertierbar.

**Beispiel 3.35** Für  $a \in \mathbb{R}$  betrachten wir die Matrix

$$S(a) := \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix};$$

die zugehörige lineare Abbildung ist dann eine Scherung parallel zur  $x$ -Achse. Für alle  $a, b \in \mathbb{R}$  ist

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+b \\ 0 & 1 \end{pmatrix},$$

also  $S(a)S(b) = S(a+b)$ . Insbesondere ist

$$S(a)S(-a) = S(-a)S(a) = S(0) = \mathbf{1}_2$$

für alle  $a \in \mathbb{R}$ , also  $S(a)$  invertierbar mit  $S(a)^{-1} = S(-a)$ .

**Beispiel 3.36** Analog zum vorigen Beispiel sieht man, dass für  $m \geq 2$  die Matrix

$$S(a) = \begin{pmatrix} 1 & a & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{1}_{m-2} & & \end{pmatrix}$$

invertierbar ist für jedes  $a \in \mathbb{R}$ , mit  $S(a)^{-1} = S(-a)$ . Gegeben  $i \neq j$  in  $\{1, \dots, m\}$  gilt Analoges gilt für die Matrix  $S(a) \in \mathbb{R}^{m \times m}$ , deren  $k$ te Spalte  $\mathbf{e}_k$  ist für alle  $k \in \{1, \dots, m\} \setminus \{i\}$  und deren  $i$ te Spalte  $\mathbf{e}_i + a \mathbf{e}_j$  ist (mit den Standard-Basisvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_m$  für  $\mathbb{R}^m$ ).

Die zugehörige lineare Abbildung  $\phi_{S(a)}: \mathbb{R}^m \rightarrow \mathbb{R}^m$  hält dann also jeden der Basisvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_m$  fest und bildet  $\mathbf{e}_i$  auf  $\mathbf{e}_i + a \mathbf{e}_j$  ab.

Für jede Matrix  $A \in \mathbb{R}^{m \times n}$  entsteht dann

$$S(a)A$$

aus  $A$ , indem man das  $a$ -fache der  $i$ ten Zeile zur  $j$ ten Zeile von  $A$  addiert.

## Gauß-Jordan-Verfahren zur Berechnung inverser Matrizen

In diesem Abschnitt lernen wir ein praktikables Verfahren zum Invertieren von Matrizen kennen. Es basiert auf der folgenden Tatsache, die wir zuerst begründen:

**3.37** *Es sei  $A \in \mathbb{R}^{m \times n}$  eine  $m \times n$ -Matrix. Entsteht eine Matrix  $B \in \mathbb{R}^{m \times n}$  aus  $A$  durch Anwendung einer oder mehrerer elementarer Zeilenoperationen, so ist*

$$B = TA$$

*mit einer invertierbaren Matrix  $T \in \mathbb{R}^{m \times m}$ .*

Seien  $A_0, A_1, A_2, \dots, A_k$  Matrizen mit  $A_0 = A$  und  $A_k = B$ , so dass für jedes  $j \in \{1, \dots, k\}$  die Matrix  $A_j$  aus  $A_{j-1}$  durch eine elementare Zeilenoperation hervorgeht. Können wir zeigen, dass

$$A_j = T_j A_{j-1}$$

für eine invertierbare Matrix  $T_j \in \mathbb{R}^{m \times m}$ , so ist  $B = TA$  mit

$$T = T_k \cdots T_1.$$

Als Produkt invertierbarer Matrizen ist  $T$  nach 3.32 (b) invertierbar. Es genügt also, den Fall zu diskutieren, in welchem  $B$  aus  $A$  durch eine einzelne elementare Zeilenoperation hervorgeht. Alle drei Arten von elementaren Zeilenoperationen lassen sich nach den Beispielen 3.33, 3.34 und 3.36 aber durch Multiplikation von  $A$  mit einer invertierbaren Matrix  $T$  der Form  $D(\lambda)$ ,  $P$  oder  $S(a)$  von links realisieren.

**3.38** (Gauß-Jordan-Verfahren). Um die inverse Matrix  $A^{-1}$  einer invertierbaren  $n \times n$ -Matrix  $A$  zu finden, schreiben wir  $A$  und die Einheitsmatrix  $\mathbf{1}_n$  nebeneinander, betrachten also die  $n \times (2n)$ -Matrix

$$(A|\mathbf{1}_n).$$

Man wendet auf diese Matrix nun elementare Zeilenoperationen an, wobei wir  $A$  zuerst auf Zeilenstufenform bringen (also obere Dreiecksform) und anschließend auf die Form  $\mathbf{1}_n$ . Wir enden dann also bei einer  $n \times (2n)$ -Matrix der Form

$$(\mathbf{1}_n|B) \tag{34}$$

mit einer  $n \times n$ -Matrix  $B$ . Wir werden sehen, dass dieses Vorgehen immer möglich ist, und dass

$$B = A^{-1}.$$

Wissen wir vorher noch nicht, dass  $A$  invertierbar ist, kommen aber nach elementaren Zeilenumformungen an bei einer  $n \times (2n)$ -Matrix der Form (34), so ist  $A$  invertierbar und es ist wieder  $B = A^{-1}$ .

Bevor wir dies theoretisch begründen, schauen wir uns ein Beispiel an.

**Beispiel 3.39** Führen Sie das Gauß-Jordan-Verfahren durch für die Matrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}.$$

Zeigen Sie auf diesem Wege, dass  $A$  invertierbar ist und finden Sie  $A^{-1}$ .

Lösung: Wir führen wie folgt elementare Zeilenumformungen durch:

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \downarrow \\ \downarrow \end{array} \begin{array}{l} (-1) \\ (-1) \end{array} \\ \rightsquigarrow & \left( \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} \downarrow \\ \downarrow \end{array} \begin{array}{l} (-1) \\ (-1) \end{array} \\ \rightsquigarrow & \left( \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right) \cdot (-1) \\ \rightsquigarrow & \left( \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right) \begin{array}{l} \uparrow \\ \uparrow \end{array} \begin{array}{l} (-2) \\ (-2) \end{array} \\ \rightsquigarrow & \left( \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & -1 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right) \begin{array}{l} \uparrow \\ \uparrow \end{array} \begin{array}{l} (+1) \\ (+1) \end{array} \\ \rightsquigarrow & \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 0 & -2 & -1 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right) \end{aligned}$$

Also ist  $A$  invertierbar mit

$$A^{-1} = \begin{pmatrix} -1 & -1 & 2 \\ -2 & -1 & 2 \\ 1 & 1 & -1 \end{pmatrix}.$$

### 3.40 Begründung des Verfahrens.

(i) Angenommen, wir können  $(A|\mathbf{1}_n)$  durch Anwendung von  $k$  Stück elementaren Zeilenoperationen auf die Form  $(\mathbf{1}_n|B)$  bringen. Sei  $(A_j|B_j)$  die nach  $j$  der Zeilenoperationen erhaltene Matrix. Dann ist also

$$(A_0|B_0) = (A|\mathbf{1}_n) \quad \text{und} \quad (A_k|B_k) = (\mathbf{1}_n|B).$$

Weiter gibt es nach 3.37 für jedes  $j \in \{1, \dots, k\}$  eine invertierbare  $n \times n$ -Matrix  $T_j$  derart, dass

$$T_j(A_{j-1}|B_{j-1}) = (A_j|B_j),$$

also  $A_j = T_j A_{j-1}$  und  $B_j = T_j B_{j-1}$ . Dann ist

$$\mathbf{1}_n = A_k = T_k \cdots T_2 T_1 A_0 = T_k \cdots T_2 T_1 A = TA$$

mit der Matrix  $T := T_k \cdots T_2 T_1$ , die nach 3.32 (b) invertierbar ist als Produkt invertierbarer Matrizen. Also ist auch

$$A = T^{-1}TA = T^{-1}\mathbf{1}_n = T^{-1}$$

invertierbar, mit  $A^{-1} = T$ . Hierbei ist  $B = B_k = T_k \cdots T_2 T_1 B_0 = T_k \cdots T_2 T_1 \mathbf{1}_n = T$ , also  $A^{-1} = B$ .

(ii) Nun sei  $A$  als invertierbar angenommen. Bringen wir  $A$  durch elementare Zeilenoperationen auf Zeilen-Stufenform  $A'$ , so ist nach 3.37  $A' = TA$  mit einer invertierbaren  $n \times n$ -Matrix  $T$ . Nach 3.32 (b) ist dann auch  $A'$  invertierbar, also  $\phi_{A'}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  invertierbar, somit  $A'\mathbf{x} = \mathbf{b}$  stets eindeutig lösbar und folglich der Rang  $r$  von  $A'$  gleich  $n$ . Also ist  $A'$  eine obere Dreiecksmatrix und alle Einträge  $\lambda_1, \dots, \lambda_n$  auf der Diagonalen sind von Null verschieden. Indem wir nacheinander für  $j \in \{1, \dots, n\}$  die  $j$ te Zeile von  $A'$  mit  $\lambda_j^{-1}$  multiplizieren, dürfen wir annehmen, dass  $\lambda_j = 1$  für alle  $j \in \{1, \dots, n\}$ . Ist  $A' = (a'_{ij})$ , so dürfen wir nach Subtraktion des  $a'_{in}$ -fachen der  $n$ ten Zeile von  $A'$  von der  $i$ ten Zeile von  $A'$  für  $i \in \{1, \dots, n-1\}$  annehmen, dass

$a_{in} = 0$  für alle  $i \in \{1, \dots, n-1\}$ . Analog dürfen wir annehmen, dass für  $j \in \{n-1, \dots, 2\}$  die  $j$ te Spalte von  $A'$  außer  $a_{jj} = 1$  nur Nullen enthält. Also ist  $A' = \mathbf{1}_n$ , somit

$$TA = A' = \mathbf{1}_n.$$

Multiplikation mit  $T^{-1}$  von links liefert

$$A = T^{-1},$$

sodass also  $T = A^{-1}$ . Nun ist

$$T(A|\mathbf{1}_n) = (TA, T\mathbf{1}_n) = (A'|T) = (\mathbf{1}_n|T)(\mathbf{1}_n|A^{-1}).$$

## Isomorphismen zwischen Vektorräumen

**Definition 3.41** Ein Isomorphismus zwischen Vektorräumen  $V$  und  $W$  ist eine bijektive lineare Abbildung  $\psi: V \rightarrow W$ .

Nach 3.29 ist dann auch die bijektive Abbildung  $\psi^{-1}: W \rightarrow V$  linear, somit auch  $\psi^{-1}$  ein Isomorphismus.

**3.42** *Es sei  $\psi: V \rightarrow W$  ein Isomorphismus zwischen Vektorräumen. Dann gilt:*

- (a) *Sind  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  linear unabhängig, so sind  $\psi(\mathbf{v}_1), \dots, \psi(\mathbf{v}_n)$  in  $W$  linear unabhängig.*
- (b) *Sind  $\mathbf{v}_1, \dots, \mathbf{v}_n$  eine Basis für  $V$ , so ist  $\psi(\mathbf{v}_1), \dots, \psi(\mathbf{v}_n)$  eine Basis für  $W$ .*

Beweis: Ist  $\sum_{j=1}^n \lambda_j \psi(\mathbf{v}_j) = \mathbf{0}$  mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ , so liefert Anwendung von  $\psi^{-1}$ , dass

$$\mathbf{0} = \sum_{j=1}^n \lambda_j \psi^{-1}(\psi(\mathbf{v}_j)) = \sum_{j=1}^n \lambda_j \mathbf{v}_j.$$

Da  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig sind, folgt  $\lambda_1 = \dots = \lambda_n = 0$ . Also sind auch  $\psi(\mathbf{v}_1), \dots, \psi(\mathbf{v}_n)$  linear unabhängig.

(b) Nach (a) sind die Vektoren  $\psi(\mathbf{v}_1), \dots, \psi(\mathbf{v}_n)$  in  $W$  linear unabhängig. Für jedes  $\mathbf{w}$  sind die Vektoren

$$\mathbf{v}_1, \dots, \mathbf{v}_n, \psi^{-1}(\mathbf{w})$$

in  $V$  linear abhängig, da die Basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  maximal linear unabhängig ist. Also gibt es  $\lambda_1, \dots, \lambda_n, \lambda \in \mathbb{R}$ , die nicht alle 0 sind, mit

$$\lambda \psi^{-1}(\mathbf{w}) + \sum_{j=1}^n \lambda_j \mathbf{v}_j = \mathbf{0}.$$

Anwendung von  $\psi$  liefert

$$\lambda \mathbf{w} + \sum_{j=1}^n \lambda_j \psi(\mathbf{v}_j) = \mathbf{0}.$$

Also sind  $\psi(\mathbf{v}_1), \dots, \psi(\mathbf{v}_n)$  maximal linear unabhängig in  $W$  und somit eine Basis für  $W$ .

**3.43** *Es sei  $V$  ein Vektorraum mit einer Basis  $\mathbf{b}_1, \dots, \mathbf{b}_n$ . Weiter sei  $W$  ein Vektorraum und  $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$ . Dann gilt:*

- (a) *Es gibt genau eine lineare Abbildung  $\psi: V \rightarrow W$  derart, dass  $\psi(\mathbf{b}_j) = \mathbf{w}_j$  für alle  $j \in \{1, \dots, n\}$ .*
- (b) *Ist  $\mathbf{w}_1, \dots, \mathbf{w}_n$  eine Basis für  $W$ , so ist  $\psi$  ein Isomorphismus.*

Beweis. (a) Die lineare Abbildung  $\psi$  ist eindeutig wenn existent, denn jeder Vektor  $\mathbf{x} \in V$  ist von der Form  $\mathbf{x} = \sum_{j=1}^n \lambda_j \mathbf{b}_j$  mit eindeutigen  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  und dann ist notwendig

$$\psi(\mathbf{x}) = \psi\left(\sum_{j=1}^n \lambda_j \mathbf{b}_j\right) = \sum_{j=1}^n \lambda_j \psi(\mathbf{b}_j) = \sum_{j=1}^n \lambda_j \mathbf{w}_j.$$

Um die Existenz von  $\psi$  zu zeigen, seien für  $\mathbf{x} \in V$  die eindeutigen  $\lambda_1, \dots, \lambda_n$  wie zuvor; wir setzen

$$\psi(\mathbf{x}) := \sum_{j=1}^n \lambda_j \mathbf{w}_j.$$

Für jedes  $\lambda \in \mathbb{R}$ , ist  $\lambda \mathbf{x} = \sum_{j=1}^n \lambda \lambda_j \mathbf{b}_j = \sum_{j=1}^n \mu_j \mathbf{b}_j$  mit  $\mu_j = \lambda \lambda_j$ , somit

$$\psi(\lambda \mathbf{x}) = \sum_{j=1}^n \mu_j \mathbf{w}_j = \sum_{j=1}^n \lambda \lambda_j \mathbf{w}_j = \lambda \sum_{j=1}^n \lambda_j \mathbf{w}_j = \lambda \psi(\mathbf{x}).$$

Ist  $\mathbf{x}' \in V$  mit  $\mathbf{x}' = \sum_{j=1}^n \lambda'_j \mathbf{b}_j$ , so ist  $\mathbf{x} + \mathbf{x}' = \sum_{j=1}^n (\lambda_j + \lambda'_j) \mathbf{b}_j$ , also

$$\psi(\mathbf{x} + \mathbf{x}') = \sum_{j=1}^n (\lambda_j + \lambda'_j) \mathbf{w}_j = \sum_{j=1}^n \lambda_j \mathbf{w}_j + \sum_{j=1}^n \lambda'_j \mathbf{w}_j = \psi(\mathbf{x}) + \psi(\mathbf{x}').$$

Somit ist  $\psi$  linear. Da  $\mathbf{b}_i = \sum_{j=1}^n \delta_{ij} \mathbf{b}_j$  für  $i \in \{1, \dots, n\}$ , ist weiter  $\psi(\mathbf{b}_i) = \sum_{j=1}^n \delta_{ij} \mathbf{w}_j = \mathbf{w}_i$ .

(b) Sei  $\psi$  wie in (a) und  $\theta: W \rightarrow V$  die eindeutige lineare Abbildung mit  $\theta(\mathbf{w}_j) = \mathbf{b}_j$  für alle  $j \in \{1, \dots, n\}$ . Dann ist  $\theta \circ \psi: V \rightarrow V$  die eindeutige lineare Abbildung, die  $\mathbf{b}_j$  auf  $\mathbf{b}_j$  abbildet für alle  $j \in \{1, \dots, n\}$ , also  $\theta \circ \psi = \text{id}_V$ . Analog ist  $\psi \circ \theta = \text{id}_W$ . Also ist  $\psi$  invertierbar mit  $\psi^{-1} = \theta$ , somit  $\psi$  ein Isomorphismus von Vektorräumen.

**3.44** Hat ein Vektorraum  $W$  eine Basis  $\mathbf{b}_1, \dots, \mathbf{b}_m$  mit  $m \in \mathbb{N}$ , so gilt:

- (a) Für jede Basis  $\mathbf{w}_1, \dots, \mathbf{w}_n$  von  $W$  ist  $n = m$ .
- (b) Sind  $\mathbf{w}_1, \dots, \mathbf{w}_n$  linear unabhängige Vektoren in  $W$ , so gilt  $n \leq m$ . Es gilt genau dann  $n = m$ , wenn  $\mathbf{w}_1, \dots, \mathbf{w}_n$  eine Basis für  $W$  ist.

Beweis. Nach 3.43 gibt es einen Isomorphismus  $\psi: \mathbb{R}^m \rightarrow W$  von Vektorräumen derart, dass  $\psi(\mathbf{e}_j) = \mathbf{b}_j$  für alle  $j \in \{1, \dots, n\}$ ; hierbei ist  $\mathbf{e}_1, \dots, \mathbf{e}_n$  die Standard-Basis für  $\mathbb{R}^n$ . Wegen 3.42 folgen die Aussagen nun aus 3.19.

**Bemerkung 3.45** Hat ein Vektorraum  $W$  eine endliche Basis, so lässt sich ihm nach (a) eine wohldefinierte *Dimension* zuordnen, nämlich die Zahl

$$\dim(W) := n$$

der Elemente einer (und dann auch jeder) endlichen Basis für  $W$ .

Die Dimension des trivialen Vektorraums  $\{0\}$  ist per Definition 0 und  $\emptyset$  eine Basis.

**Bemerkung 3.46** Ein Vektorraum  $W$  braucht keine endliche Basis zu besitzen. Wir nennen ihn dann unendlich-dimensional.

## 4 Determinanten

Aus der Mathematik 1 kennen wir bereits die Determinante einer  $2 \times 2$ -Matrix  $A = (a_{ij}) \in \mathbb{R}^{2 \times 2}$ ,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

und die Determinante einer  $3 \times 3$ -Matrix  $A = (a_{ij}) \in \mathbb{R}^{3 \times 3}$ ,

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33},$$

die nach der Sarrusschen Regel berechnet werden kann (siehe Tafel), also grob gesagt

$$\begin{array}{ccc} \searrow & \searrow & \searrow \\ \swarrow & \swarrow & \swarrow \end{array} - \begin{array}{ccc} \swarrow & \swarrow & \swarrow \\ \searrow & \searrow & \searrow \end{array},$$

wenn man die erste und zweite Spalte noch einmal rechts neben die Matrix schreibt. In beiden Fällen haben wir gesehen, dass  $A$  genau dann invertierbar ist, wenn  $\det(A) \neq 0$ .

In diesem Kapitel wollen wir Determinanten für  $n \times n$ -Matrizen einführen und berechnen, die entsprechende Eigenschaften haben.

**4.1** Wir definieren Determinanten von  $n \times n$ -Matrizen  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  induktiv wie folgt:

(i) Ist  $n = 1$  und  $A = (a_{11})$  eine  $1 \times 1$ -Matrix, so setzen wir

$$\det(a_{11}) := a_{11}$$

als den Eintrag der Matrix. Zum Beispiel ist also  $\det(5) = 5$ .

(ii) Im Falle  $n \geq 2$  definieren wir die Determinante von  $A$  als<sup>8</sup>

$$\begin{aligned} \det(A) &:= \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A_{i1}) \\ &= a_{11} \det(A_{11}) - a_{21} \det(A_{21}) + \dots + (-1)^{n+1} a_{n1} \det(A_{n1}). \end{aligned}$$

---

<sup>8</sup>In späterer Sprechweise definieren wir  $\det(A)$  also durch Entwickeln nach der ersten Spalte. In der Vorlesung hatten wir stattdessen nach der ersten Zeile entwickelt, was wegen 4.5 aber das gleiche Ergebnis liefert.

Für  $i, j \in \{1, \dots, n\}$  ist hierbei  $A_{ij} \in \mathbb{R}^{(n-1) \times (n-1)}$  die Matrix, die aus  $A$  durch streichen der  $i$ ten Zeile und  $j$ ten Spalte entsteht.

**Beispiel 4.2** Betrachten wir

$$A := \begin{pmatrix} 2 & 0 & -7 \\ 3 & 4 & -3 \\ 0 & 2 & 1 \end{pmatrix},$$

so ist z.B.

$$A_{13} = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 2 & -7 \\ 0 & 1 \end{pmatrix}.$$

**Beispiel 4.3** (a) Für

$$A = \begin{pmatrix} 1 & 2 & -2 & -1 \\ 2 & 1 & -1 & 0 \\ 3 & 4 & 0 & -2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

ist

$$\begin{aligned} \det(A) &= 1 \det \begin{pmatrix} 1 & -1 & 0 \\ 4 & 0 & -2 \\ 1 & 1 & 3 \end{pmatrix} - 2 \det \begin{pmatrix} 2 & -2 & -1 \\ 4 & 0 & -2 \\ 1 & 1 & 3 \end{pmatrix} + 3 \det \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix} \\ &= 1 \cdot 16 - 2 \cdot 28 + 3 \cdot (-2) = -46. \end{aligned}$$

(b) Ist

$$A = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

eine Diagonalmatrix, so ist einfach

$$\det(A) = a_{11} a_{22} \cdots a_{nn}$$

das Produkt der Diagonaleinträge.

Dies zeigt man per Induktion nach  $n$ . Für  $n = 1$  ist die Aussage klar. Da

$a_{21}, \dots, a_{n1} = 0$  gilt und  $A_{11} = \text{diag}(a_{22}, \dots, a_{nn})$  diagonal ist mit  $\det(A_{11}) = a_{22} \cdots a_{nn}$ , ist weiter

$$\det(A) = (-1)^{1+1} a_{11} \det(A_{11}) + 0 = a_{11} a_{22} \cdots a_{nn}.$$

Insbesondere gilt für die Einheitsmatrix  $\mathbf{1} \in \mathbb{R}^{n \times n}$ , dass

$$\det(\mathbf{1}) = 1 \cdot \dots \cdot 1 = 1.$$

(c) Ist

$$A = \begin{pmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix, also  $a_{ij} = 0$  für alle  $i, j \in \{1, \dots, n\}$  mit  $i > j$ , so gilt ebenfalls einfach

$$\det(A) = a_{11} a_{22} \cdots a_{nn}.$$

Wiederum zeigt man dies per Induktion nach  $n$ , wobei die Aussage für  $n = 1$  klar ist. Da

$$A_{11} = \begin{pmatrix} a_{22} & & & * \\ & a_{33} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix mit Diagonaleinträgen  $a_{22}, \dots, a_{nn}$  ist, gilt  $\det(A_{11}) = a_{22} \cdots a_{nn}$  per Induktionsvoraussetzung. Da  $a_{21}, \dots, a_{n1} = 0$ , folgt

$$\det(A) = (-1)^{1+1} a_{11} \det(A_{11}) + 0 = a_{11} a_{22} \cdots a_{nn}.$$

**4.4** (Eigenschaften von Determinanten). Wir betrachten eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$ .

(a) Änderungen der Determinante bei elementaren Zeilen- bzw. Spaltenumformungen:

- (i) Addiert man ein Vielfaches einer Zeile (Spalte) zu einer anderen Zeile (bzw. Spalte) von  $A$ , so ändert sich der Wert der Determinante nicht.

Beispiel:

$$\det \begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix} = \det \begin{pmatrix} 2 & 3 \\ 0 & -7 \end{pmatrix};$$

hier wurde das doppelte der ersten Zeile von der zweiten abgezogen.

- (ii) Multiplikation einer Zeile (oder Spalte) mit einer reellen Zahl  $\lambda$  verändert die Determinante um einen Faktor  $\lambda$ . Sind etwa  $\mathbf{s}_1, \dots, \mathbf{s}_n$  die Spaltenvektoren von  $A$ , so ist

$$\det(\mathbf{s}_1 \cdots \mathbf{s}_{j-1} \lambda \mathbf{s}_j \mathbf{s}_{j+1} \cdots \mathbf{s}_n) = \lambda \det(A).$$

Beispiel:

$$\det \begin{pmatrix} 40 & 10 \\ 4 & -1 \end{pmatrix} = 10 \det \begin{pmatrix} 4 & 1 \\ 4 & -1 \end{pmatrix}.$$

- (iii) Das Vertauschen von zwei Zeilen (oder Spalten) ändert das Vorzeichen der Determinanten.

Beispiel:

$$\det \begin{pmatrix} 4 & -1 \\ 3 & 2 \end{pmatrix} = -\det \begin{pmatrix} 3 & 2 \\ 4 & -1 \end{pmatrix}.$$

- (b) Symmetrie in Zeilen und Spalten: Für die transponierte Matrix gilt

$$\det(A^T) = \det(A).$$

- (c) Invertierbarkeit: Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann invertierbar, wenn  $\det(A) \neq 0$ .

Nachweis für (c): Mit elementaren Zeilenoperationen lässt sich  $A$  in eine obere Dreiecksmatrix  $A'$  verwandeln (eine Matrix in Zeilenstufenform). Nach 4.4 (a) ist dann

$$\det(A') = \lambda \det(A) \tag{35}$$

mit einer reellen Zahl  $\lambda \neq 0$ . Somit:

$A$  invertierbar

$\Leftrightarrow \mathbf{Ax} = \mathbf{b}$  ist für alle  $\mathbf{b} \in \mathbb{R}^n$  eindeutig lösbar

$\Leftrightarrow$  alle Diagonalelemente von  $A'$  sind von 0 verschieden

$\Leftrightarrow \det(A') \neq 0$  (siehe Beispiel 4.3 (c));

$\Leftrightarrow \det(A) \neq 0$  (siehe (35)).

Aus den obigen Eigenschaften folgt:

**4.5** (Laplacescher Entwicklungssatz). Für alle  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  kann die Determinante wie folgt berechnet werden:

- Entwicklung nach der  $j$ ten Spalte für  $j \in \{1, \dots, n\}$ :

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

(dies folgt mit (iii) in 4.4 (a));

- Entwicklung nach der  $i$ ten Zeile für  $i \in \{1, \dots, n\}$ :

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

(dies folgt mit 4.4 (b) aus der Entwicklung nach der  $i$ ten Spalte für  $\det(A^T)$ ).

**Beispiel 4.6** Entwickeln nach der 3. Zeile liefert für

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 2 & 4 & 1 \\ -2 & 0 & 0 & 1 \\ -1 & 3 & -2 & 3 \end{pmatrix}$$

die Determinante

$$\begin{aligned} \det(A) &= -2 \det \begin{pmatrix} 2 & 3 & 0 \\ 2 & 4 & 1 \\ 3 & -2 & 3 \end{pmatrix} - 1 \det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ -1 & 3 & -2 \end{pmatrix} \\ &= -2(24 + 9 + 4 - 18) - 1(-4 - 8 + 18 + 6 + 8 - 12) = -46. \end{aligned}$$

**4.7** (Multiplikationssatz) Für alle  $n \times n$ -Matrizen  $A$  und  $B$  ist

$$\det(AB) = \det(A) \det(B).$$

**4.8** (Folgerungen) Seien  $A, B \in \mathbb{R}^{n \times n}$ .

(a)  $\det(AB) = \det(BA)$

[Denn  $\det(AB) = \det(A) \det(B) = \det(B) \det(A) = \det(BA)$ ].

(b)  $\det(A^k) = (\det(A))^k$  für alle  $k \in \mathbb{N}$ .

(c)  $\det(A^{-1}) = (\det(A))^{-1} = \frac{1}{\det(A)}$  wenn  $A$  invertierbar ist.

[Denn  $1 = \det(\mathbf{1}) = \det(AA^{-1}) = \det(A) \det(A^{-1})$ .]

(d)  $\det(C^{-1}AC) = \det(A)$  für jede invertierbare  $n \times n$ -Matrix  $C$ .

Ergebnisse aus Kapitel 3 und 4 zusammenfassend, können wir nun festhalten:

**4.9** Für eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  mit den Spalten  $\mathbf{s}_1, \dots, \mathbf{s}_n \in \mathbb{R}^n$  sind äquivalent:

- (a)  $A$  ist invertierbar;
- (b) Das LGS  $A\mathbf{x} = \mathbf{b}$  ist für jedes  $\mathbf{b} \in \mathbb{R}^n$  eindeutig lösbar;
- (c) Die Spaltenvektoren  $\mathbf{s}_1, \dots, \mathbf{s}_n$  bilden eine Basis für  $\mathbb{R}^n$ ;
- (d) Die Spaltenvektoren  $\mathbf{s}_1, \dots, \mathbf{s}_n$  sind in  $\mathbb{R}^n$  linear unabhängig;
- (e)  $\det(A) \neq 0$ .

Übrigens lässt sich aus 4.4 (a) ein praktikables Verfahren zur Berechnung von Determinanten herleiten.

**4.10** (Determinanten berechnen mit Gaußverfahren). Sei  $A \in \mathbb{R}^{n \times n}$ . Um  $\det(A)$  zu berechnen, bringen Sie  $A$  mit elementaren Zeilenoperationen auf obere Dreiecksgestalt  $A'$ . (z.B. auf Zeilenstufenform). In jedem Schritt wissen wir aus 4.4 (a), wie sich die Determinante ändert. Die Determinante der oberen Dreiecksmatrix ist das Produkt der Diagonaleinträge (siehe 4.4 (c)). Sie dürfen zusätzlich sogar auch elementare Spaltenoperationen benutzen.

Wir geben zur Erläuterung ein  $3 \times 3$ -Beispiel. Richtig lohnend ist das Verfahren bei größeren Matrizen, z.B.  $10 \times 10$ -Matrizen.

**Beispiel 4.11** Es ist

$$\begin{aligned} \det \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 4 & 1 & 1 \end{pmatrix} &\stackrel{\downarrow(-2)}{\leftarrow} = \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 4 & 1 & 1 \end{pmatrix} \stackrel{\downarrow(-4)}{\leftarrow} = \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & -3 \end{pmatrix} \stackrel{\downarrow(-1)}{\leftarrow} \\ &= \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -4 \end{pmatrix} = -4. \end{aligned}$$

## 5 Kern und Bild einer linearen Abbildung

Wir gehen noch einmal ausführlicher auf Untervektorräume ein und lernen neue Beispiele kennen: Den Kern und das Bild einer linearen Abbildung. Deren Dimensionen sind verknüpft über die sogenannte *Dimensionsformel*.

**Definition 5.1** Ein Vektorraum  $E$  heißt *endlich-dimensional*, wenn er eine endliche Basis  $\mathbf{b}_1, \dots, \mathbf{b}_m$  besitzt.

Hier ist  $m \in \mathbb{N}_0$  (der Fall  $m = 0$  entspricht dem trivialen Vektorraum  $\{0\}$  mit  $\emptyset$  als Basis). Wir beobachten:

**Bemerkung 5.2** Ist  $E$  endlich-dimensional, so auch jeder Untervektorraum  $F \subseteq E$ . Es ist  $\dim(F) \leq \dim(E)$ . Ist  $F \neq E$ , so ist  $\dim(F) < \dim(E)$ .

Ist  $m := \dim(E)$ , so sind alle linear unabhängigen Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_n \in F$  auch in  $E$  linear unabhängig, also  $n \leq m$ . Somit kann  $n$  in  $\{0, 1, \dots, m\}$  maximal gewählt werden und dann ist  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine Basis von  $F$ . Ist  $n = m$ , so bilden die linear unabhängigen Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_m$  auch eine Basis für  $E$  (wie wir gesehen haben), da ihre Anzahl gleich  $\dim(E)$  ist; somit ist  $F = \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_m\} = E$ .

**Definition 5.3** Ist  $\phi: V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen, so definieren wir ihren *Kern* als

$$\ker(\phi) := \{\mathbf{v} \in V : \phi(\mathbf{v}) = \mathbf{0}\}.$$

Dies ist eine Teilmenge von  $V$ . Wie für jede Abbildung ist das *Bild* von  $\phi$  definiert als

$$\phi(V) := \{\phi(\mathbf{v}) : \mathbf{v} \in V\}.$$

Dies ist eine Teilmenge von  $W$ .

**5.4** Es ist  $\ker(\phi)$  ein Untervektorraum von  $V$ .

Sind nämlich  $\mathbf{v}_1, \mathbf{v}_2 \in \ker(\phi)$ , so ist

$$\phi(\mathbf{v}_1 + \mathbf{v}_2) = \phi(\mathbf{v}_1) + \phi(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

also  $\mathbf{v}_1 + \mathbf{v}_2 \in \ker(\phi)$ . Analog sieht man, dass  $t\mathbf{v} \in \ker(\phi)$  für alle  $t \in \mathbb{R}$  und  $\mathbf{v} \in \ker(\phi)$ . Außerdem ist offenbar der Nullvektor  $\mathbf{0} \in V$  in  $\ker(\phi)$ .

**5.5** Es ist  $\phi(V)$  ein Untervektorraum von  $W$ .

Ist nämlich  $\mathbf{w} \in \phi(V)$  und  $t \in \mathbb{R}$ , so gibt es ein  $\mathbf{v} \in V$  mit  $\phi(\mathbf{v}) = \mathbf{w}$ . Dann ist  $t\mathbf{v} \in V$  und

$$\phi(t\mathbf{v}) = t\phi(\mathbf{v}) = t\mathbf{w},$$

also  $t\mathbf{w} \in \phi(V)$ . Analog ist  $\mathbf{w}_1 + \mathbf{w}_2 \in \phi(V)$  für alle  $\mathbf{w}_1, \mathbf{w}_2 \in \phi(V)$ . Weiter ist  $\mathbf{0} = \phi(\mathbf{0}) \in \phi(V)$ .

**Beispiel 5.6** Der Lösungsraum  $L_h$  eines homogenen linearen Gleichungssystems  $A\mathbf{x} = \mathbf{0}$  mit  $A \in \mathbb{R}^{m \times n}$  ist der Kern der linearen Abbildung

$$\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{x} \mapsto A\mathbf{x}$$

und somit ein Untervektorraum von  $\mathbb{R}^n$  (was wir schon wissen, und damals per Hand nachrechnen konnten).

**5.7** Eine lineare Abbildung  $\phi: V \rightarrow W$  ist genau dann injektiv, wenn  $\ker(\phi) = \{\mathbf{0}\}$ .

Ist  $\ker(\phi) \neq \{\mathbf{0}\}$  mit dem Nullvektor  $\mathbf{0} \in V$ , so existiert ein  $\mathbf{v} \in V$  mit  $\mathbf{v} \neq \mathbf{0}$  und

$$\phi(\mathbf{v}) = \mathbf{0},$$

dem Nullvektor in  $W$ . Es ist aber auch  $\phi(\mathbf{0}) = \mathbf{0}$ , also  $\phi$  nicht injektiv.

Ist  $\ker(\phi) = \{\mathbf{0}\}$  und sind  $\mathbf{v}_1, \mathbf{v}_2 \in V$  mit  $\phi(\mathbf{v}_1) = \phi(\mathbf{v}_2)$ , so ist

$$\mathbf{0} = \phi(\mathbf{v}_2) - \phi(\mathbf{v}_1) = \phi(\mathbf{v}_2 - \mathbf{v}_1)$$

und somit  $\mathbf{v}_2 - \mathbf{v}_1 \in \ker(\phi) = \{\mathbf{0}\}$ , also  $\mathbf{v}_2 - \mathbf{v}_1 = \mathbf{0}$  und somit  $\mathbf{v}_1 = \mathbf{v}_2$ . Also ist  $\phi$  injektiv.

**5.8** (Dimensionsformel). Für jede lineare Abbildung  $\phi: V \rightarrow W$  zwischen endlich-dimensionalen Vektorräumen  $V$  und  $W$  gilt

$$\dim \ker(\phi) + \dim \phi(V) = \dim(V).$$

Zum Nachweis sei  $\mathbf{b}_1, \dots, \mathbf{b}_k$  eine Basis von  $\ker(\phi)$ . Wir ergänzen diese Basis zu einer maximalen linear unabhängigen Teilmenge  $\mathbf{b}_1, \dots, \mathbf{b}_n$  von  $V$ . Dies ist möglich, da es in  $V$  nicht mehr als  $\dim(V)$  linear unabhängige Vektoren

gibt. Dann ist  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine Basis von  $V$  und somit  $n = \dim(V)$ . Setzen wir  $S := \text{span}\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ , so ist

$$\mathbf{b}_{k+1}, \dots, \mathbf{b}_n$$

eine Basis für  $S$ , also  $\dim(S) = n - k$ . Wir behaupten, dass  $\phi|_S: S \rightarrow \phi(V)$  bijektiv und somit ein Isomorphismus von Vektorräumen ist. Also ist

$$\dim \phi(V) = \dim S = n - k$$

und somit  $\dim(V) = n = \dim \phi(V) + k = \dim \phi(V) + \dim \ker(\phi)$ , wie benötigt. Zum Beweis der Behauptung sei  $\mathbf{v} \in V$ . Dann  $\mathbf{v} = \sum_{j=1}^n t_j \mathbf{b}_j$  mit geeigneten  $t_1, \dots, t_n \in \mathbb{R}$  und

$$\phi(\mathbf{v}) = \sum_{j=1}^n t_j \phi(\mathbf{b}_j) = \sum_{j=1}^k t_j \underbrace{\phi(\mathbf{b}_j)}_{=\mathbf{0}} + \sum_{j=k+1}^n t_j \phi(\mathbf{b}_j) = \phi\left(\sum_{j=k+1}^n t_j \mathbf{b}_j\right) \in \phi(S),$$

also  $\phi(V) = \phi(S)$ . Weiter ist

$$\ker(\phi|_S) = \ker(\phi) \cap S = \{\mathbf{0}\}, \quad (36)$$

also  $\phi|_S$  injektiv und somit  $\phi|_S: S \rightarrow \phi(S) = \phi(V)$  ein Isomorphismus von Vektorräumen. Zum Nachweis von (36) sei  $\mathbf{v} = \sum_{j=1}^n t_j \mathbf{b}_j \in \ker(\phi) \cap S$ . da  $\mathbf{v} \in \ker(S)$ , ist  $t_j = 0$  für alle  $j \in \{k+1, \dots, n\}$ . Da  $\mathbf{v} \in S$ , ist  $t_j = 0$  für alle  $j \in \{1, \dots, k\}$ . Also ist  $\mathbf{v} = \mathbf{0}$ .

**Definition 5.9** Es sei  $V$  ein Vektorraum. Lineare Abbildungen  $\phi: V \rightarrow V$  nennt man *Endomorphismen* von  $V$ .

Definitionsbereich und Wertebereich stimmen also beide mit  $V$  überein.

**5.10** *Es sei  $V$  ein endlich-dimensionaler Vektorraum und  $\phi: V \rightarrow V$  ein Endomorphismus. Dann sind äquivalent:*

- (a)  $\phi$  ist bijektiv (also ein Isomorphismus);
- (b)  $\phi$  ist injektiv;
- (c)  $\phi$  ist surjektiv.

Die Implikationen “(a) $\implies$ (b)” und “(a) $\implies$ (c)” sind trivial. Gilt (b), so ist  $\ker(\phi) = \{0\}$ , nach der Dimensionsformel also

$$\dim(V) = \dim \phi(V) + \underbrace{\dim \ker(\phi)}_{=0}.$$

Der Untervektorraum  $\phi(V)$  von  $V$  hat also Dimension  $\dim(V)$ . Nach Bemerkung 5.2 ist also  $\phi(V) = V$ , somit  $\phi$  surjektiv und somit  $\phi$  bijektiv.

Gilt (c), so ist  $\phi(V) = V$ , nach der Dimensionsformel also

$$\dim \ker(\phi) = \dim(V) - \dim \phi(V) = \dim(V) - \dim(V) = 0.$$

Somit ist  $\ker(\phi) = \{0\}$ , also  $\phi$  injektiv und somit  $\phi$  bijektiv.

**Definition 5.11** Ist  $A \in \mathbb{R}^{m \times n}$  eine  $(m \times n)$ -Matrix und  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \mapsto A\mathbf{x}$  die zugehörige lineare Abbildung, so nennt man

$$\text{rk}(A) := \dim(\phi_A(\mathbb{R}^n)) = \dim(A\mathbb{R}^n)$$

den *Rang* der Matrix  $A$ .

**5.12** *Es ist  $\text{rk}(A)$  gleich der maximalen Zahl linear unabhängiger Spalten von  $A$ .*

Sind nämlich  $\mathbf{b}_1, \dots, \mathbf{b}_k$  linear unabhängige Spalten und ist für jede weitere Spalte  $\mathbf{s}$

$$\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{s}$$

linear abhängig, so ist  $\mathbf{s}$  eine Linearkombination von  $\mathbf{b}_1, \dots, \mathbf{b}_k$ , also  $\mathbf{s} \in \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_k\} =: S$ . Für jedes  $j \in \{1, \dots, n\}$  ist die  $j$ te Spalte  $\mathbf{s}_j$  also in  $S$  und somit ist

$$\phi_A(\mathbb{R}^n) = \text{span}\{\mathbf{s}_1, \dots, \mathbf{s}_n\} \subseteq S.$$

Aber  $S \subseteq \text{span}\{\mathbf{s}_1, \dots, \mathbf{s}_n\}$ , also herrscht Gleichheit:  $\phi_A(\mathbb{R}^n) = \text{span}\{\mathbf{s}_1, \dots, \mathbf{s}_n\} = S$ . Per Konstruktion hat  $S$  die Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_k$  als Basis. Also  $\text{rk}(A) = \dim \phi(\mathbb{R}^n) = \dim S = k$ .

**5.13** *Können wir  $A \in \mathbb{R}^{m \times n}$  durch elementare Zeilenoperationen auf Zeilenstufenform  $A'$  bringen, so ist  $\text{rk}(A) = \text{rk}(A')$ .*

Es gibt eine invertierbare Matrix  $T \in \mathbb{R}^{m \times m}$  mit  $A' = TA$ . Für die zugehörigen linearen Abbildungen gilt dann

$$\phi_{A'} = \phi_T \circ \phi_A,$$

wobei  $\phi_T: \mathbb{R}^m \rightarrow \mathbb{R}^m$  ein Isomorphismus ist. Es ist dann

$$\phi_{A'}(\mathbb{R}^n) = \phi_T(\phi_A(\mathbb{R}^n)),$$

also

$$\alpha := \phi_T|_{\phi_A(\mathbb{R}^n)}: \phi_A(\mathbb{R}^n) \rightarrow \phi_{A'}(\mathbb{R}^n)$$

surjektiv. Als Einschränkung der injektiven Abbildung  $\phi_T$  ist die lineare Abbildung  $\alpha$  auch injektiv, also bijektiv und somit ein Isomorphismus. Also ist

$$\dim \phi_A(\mathbb{R}^n) = \dim \phi_{A'}(\mathbb{R}^n)$$

und somit  $\text{rk}(A) = \text{rk}(A')$ .

**5.14** Ist  $A \in \mathbb{R}^{m \times n}$  eine Matrix in Zeilenstufenform, so ist  $\text{rk}(A)$  gleich der Zahl  $r$  der von  $\mathbf{0}$  verschiedenen Zeilen.

Für

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

mit  $b_j = 0$  für  $j \in \{r+1, \dots, m\}$  ist die Lösbarkeitsbedingung erfüllt für das inhomogene lineare Gleichungssystem

$$A\mathbf{x} = \mathbf{b},$$

es ist also  $\mathbf{b} = A\mathbf{x}$  für ein  $\mathbf{x} \in \mathbb{R}^n$  und somit  $\mathbf{b} \in \phi_A(\mathbb{R}^n)$ . Sind  $\mathbf{e}_1, \dots, \mathbf{e}_m \in \mathbb{R}^m$  die Standardbasisvektoren in  $\mathbb{R}^m$  und

$$S := \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_r\},$$

so ist also  $S \subseteq \phi_A(\mathbb{R}^n)$ . Umgekehrt ist jede der Spalten  $\mathbf{s}_j$  von  $A$  in  $S$ . Für  $\mathbf{x} \in \mathbb{R}^n$  mit den Komponenten  $x_j$  ist

$$\phi_A(\mathbf{x}) = \sum_{j=1}^n x_j \mathbf{s}_j \in S.$$

Also ist  $\phi_A(\mathbb{R}^n) \subseteq S$  und somit  $\phi_A(\mathbb{R}^n) = S$ . Da die Vektoren  $\mathbf{e}_1, \dots, \mathbf{e}_r$  linear unabhängig sind, bilden sie eine Basis für

$$\text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_r\} = S = \phi_A(\mathbb{R}^n).$$

Also ist  $\text{rk}(A) = \dim \phi_A(\mathbb{R}^n) = r$ .

## 6 Diagonalisierbarkeit

Wir studieren Endomorphismen, die sich bezüglich einer geeigneten Basis durch Diagonalmatrizen beschreiben lassen.

**Definition 6.1** Es sei  $V$  ein endlich-dimensionaler Vektorraum und  $\phi: V \rightarrow V$  ein Endomorphismus.

(a) Ein Vektor

$$\mathbf{v} \neq \mathbf{0}$$

in  $V$  wird ein *Eigenvektor* für  $\phi$  genannt, wenn er von  $\phi$  auf ein Vielfaches von sich selbst abgebildet wird:

$$\phi(\mathbf{v}) = \lambda \mathbf{v} \quad \text{für ein } \lambda \in \mathbb{R}.$$

Man nennt dann  $\lambda$  den zugehörigen *Eigenwert*.

(b) Hat  $V$  eine Basis  $\mathbf{b}_1, \dots, \mathbf{b}_n$  aus Eigenvektoren zu  $\phi$ , so wird  $\phi$  ein *diagonalisierbarer* Endomorphismus genannt.

In (a) wird die von  $\mathbf{v}$  aufgespannte Ursprungsgerade  $\mathbb{R}\mathbf{v}$  von  $\phi$  also in sich selbst abgebildet,

$$\phi(\mathbb{R}\mathbf{v}) \subseteq \mathbb{R}\mathbf{v}.$$

**Beispiel 6.2** (a) Der zur Diagonalmatrix

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix}$$

gehörige Endomorphismus  $\phi_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 3x_1 \\ 7x_2 \end{pmatrix}$$

hat die Standard-Basisvektoren  $\mathbf{e}_1$  und  $\mathbf{e}_2$  in  $\mathbb{R}^2$  als Eigenvektoren, zu den Eigenwerten 3 bzw. 7. Diese bilden eine Basis,  $\phi_A$  ist also diagonalisierbar.

(b) Der zur Matrix

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 7 \end{pmatrix}$$

gehörige Endomorphismus  $\phi_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 3x_1 + x_2 \\ 7x_2 \end{pmatrix}$$

hat  $\mathbf{b}_1 := \mathbf{e}_1$  als Eigenvektor zum Eigenwert 3 und

$$\mathbf{b}_2 := \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

als Eigenvektor zum Eigenwert 7, denn es ist

$$\begin{pmatrix} 3 & 1 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 28 \end{pmatrix} = 7 \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$

Da  $\mathbf{b}_1$  und  $\mathbf{b}_2$  eine Basis bilden, ist  $\phi_A$  diagonalisierbar.

Wie kommt man darauf? Wie kann man Eigenwerte und Eigenvektoren berechnen?

**6.3** Ist  $A \in \mathbb{R}^{n \times n}$  gegeben und  $\mathbf{1}_n \in \mathbb{R}^{n \times n}$  die Einheitsmatrix, so ist

$$p(\lambda) := \det(A - \lambda \mathbf{1}_n)$$

ein Polynom in  $\lambda \in \mathbb{R}$ , genannt das *charakteristische Polynom*. Es hat den Grad  $n$ .

Dann gilt:

**6.4** Die Eigenwerte von  $\phi_A$  sind genau die Nullstellen des charakteristischen Polynoms.

Bevor wir dies nachweisen, schauen wir Beispiele an.

**Beispiel 6.5** (a) In Beispiel 6.2 (a) lautet das charakteristische Polynom

$$p(\lambda) = \det \left( \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \det \begin{pmatrix} 3 - \lambda & 0 \\ 0 & 7 - \lambda \end{pmatrix} = (3 - \lambda)(7 - \lambda).$$

Es hat die Nullstellen 3 und 7, dies sind also die Eigenwerte von  $\phi_A$ .

(b) In Beispiel 6.2 (a) lautet das charakteristische Polynom

$$p(\lambda) = \det \left( \begin{pmatrix} 3 & 1 \\ 0 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \det \begin{pmatrix} 3 - \lambda & 1 \\ 0 & 7 - \lambda \end{pmatrix} = (3 - \lambda)(7 - \lambda).$$

Es hat die Nullstellen 3 und 7, dies sind also die Eigenwerte von  $\phi_A$ .

Um 6.4 zu begründen, stellen wir folgende Überlegung an: Eine reelle Zahl  $\lambda$  ist genau dann ein Eigenwert des Endomorphismus  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\mathbf{x} \mapsto A\mathbf{x}$ , wenn ein Vektor  $\mathbf{v} \neq \mathbf{0}$  in  $\mathbb{R}^n$  existiert mit

$$A\mathbf{v} = \lambda\mathbf{v},$$

also  $A\mathbf{v} = \lambda\mathbf{1}_n\mathbf{v}$ , also  $(A - \lambda\mathbf{1}_n)\mathbf{v} = \mathbf{0}$ , also  $\mathbf{v} \in \ker(\phi_A - \lambda\text{id}_{\mathbb{R}^n})$ . Somit ist  $\lambda$  genau dann ein Eigenwert, wenn der Endomorphismus  $\phi_A - \lambda\text{id}_{\mathbb{R}^n}$  nicht injektiv ist. Dies gilt nach 5.10 genau dann, wenn er nicht bijektiv ist, die zugehörige Matrix  $A - \lambda\mathbf{1}_n$  also nicht invertierbar ist, also

$$\det(A - \lambda\mathbf{1}_n) = 0$$

gilt (d.h.  $\lambda$  ist Nullstelle des charakteristischen Polynoms).

**Definition 6.6** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *diagonalisierbar*, wenn der entsprechende Endomorphismus  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\mathbf{x} \mapsto A\mathbf{x}$  diagonalisierbar ist, es also eine Basis  $\mathbf{b}_1, \dots, \mathbf{b}_n$  von Eigenvektoren für  $\phi_A$  gibt.

Die Eigenvektoren von  $\phi_A$  nennen wir auch einfach *Eigenvektoren für A*. Es ist  $\mathbf{v} \in \mathbb{R}^n$  ein Eigenvektor für  $A$  zum Eigenwert  $\lambda \in \mathbb{R}$  genau dann, wenn  $\mathbf{v} \neq \mathbf{0}$  und

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Wir werden sehen:

**6.7** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann diagonalisierbar, wenn eine invertierbare Matrix  $S \in \mathbb{R}^{n \times n}$  existiert derart, dass  $D := S^{-1}AS$  eine Diagonalmatrix ist.

Dann ist also  $A = SDS^{-1}$ .

**Bemerkung 6.8** Ist  $A = SDS^{-1}$ , so ist  $A^k = SD^kS^{-1}$  für alle  $k \in \mathbb{N}$ , per Induktion. Für  $k = 1$  ist dies die Ausgangsformel. Für  $k \geq 2$  ist unter Benutzung der Induktionsannahme und des Induktionsanfangs

$$A^k = A^{k-1}A = SD^{k-1}S^{-1}SDS^{-1} = SD^kS^{-1}.$$

Für diagonalisierbare Matrizen können wir also z.B. Potenzen gut berechnen.

**Beispiel 6.9** Wir betrachten die Matrix  $A := SDS^{-1}$  mit

$$D = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Dann ist

$$SD = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & 2 \end{pmatrix},$$

also

$$A = SDS^{-1} = \begin{pmatrix} -1 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 0 & 2 \end{pmatrix}.$$

Was ist  $A^{10}$ ? Es ist

$$D^{10} = \begin{pmatrix} (-1)^{10} & 0 \\ 0 & 2^{10} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1024 \end{pmatrix}.$$

Wegen

$$SD^{10} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1024 \end{pmatrix} = \begin{pmatrix} 1 & 1024 \\ 0 & 1024 \end{pmatrix}$$

ist dann

$$A^{10} = SD^{10}S^{-1} = \begin{pmatrix} 1 & 1024 \\ 0 & 1024 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1023 \\ 0 & 1024 \end{pmatrix}.$$

Eine diagonalisierbare Matrix  $A$  lässt sich mit dem folgenden Vorgehen auf Diagonalgestalt bringen (was eine Hälfte von 6.7 zeigt).

**6.10** Sei  $A \in \mathbb{R}^{n \times n}$  diagonalisierbar und  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine Basis für  $\mathbb{R}^n$  aus Eigenvektoren zu  $\phi_A$ ; es gibt also reelle Zahlen  $\lambda_1, \dots, \lambda_n$  derart, dass

$$A\mathbf{b}_j = \lambda_j\mathbf{b}_j$$

für alle  $j \in \{1, \dots, n\}$ . Dann ist die Matrix

$$S = (\mathbf{b}_1 \ \dots \ \mathbf{b}_n) \in \mathbb{R}^{n \times n}$$

mit den Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$  invertierbar. Setzen wir  $D := \text{diag}(\lambda_1, \dots, \lambda_n)$ , so ist

$$S^{-1}AS = D. \tag{37}$$

Zum Nachweis benutzen wir die Standard-Basisvektoren  $\mathbf{e}_1, \dots, \mathbf{e}_n$  für  $\mathbb{R}^n$ . Für jedes  $j \in \{1, \dots, n\}$  ist  $S\mathbf{e}_j = \mathbf{b}_j$ , die  $j$ te Spalte von  $S$ . Die  $j$ te Spalte der linken Matrix in (37) ist also

$$S^{-1}AS\mathbf{e}_j = S^{-1}A\mathbf{b}_j = \lambda_j S^{-1}\mathbf{b}_j = \lambda_j \mathbf{e}_j.$$

Dies stimmt überein mit der  $j$ ten Spalte  $D\mathbf{e}_j$  der Matrix  $D$ .

**Beispiel 6.11** Zeigen Sie, dass die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

diagonalisierbar ist. Finden Sie zudem eine invertierbare Matrix  $S \in \mathbb{R}^{2 \times 2}$  und eine Diagonalmatrix  $D \in \mathbb{R}^{2 \times 2}$  derart, dass  $S^{-1}AS = D$ .

Lösung. Die Eigenwerte von  $A$  sind die Nullstellen des charakteristischen Polynoms,

$$0 = \det \left( \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) = \det \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 4.$$

Es ist also  $1 - \lambda = \pm 2$ , die Eigenwerte sind folglich  $\lambda_1 = -1$ ,  $\lambda_2 = 3$ .

Wir suchen nun einen Eigenvektor zum Eigenwert  $-1$  und einen zum Eigenwert  $3$ . Wir werden sehen, dass die beiden linear unabhängig sind und somit eine Basis von Eigenvektoren bilden.

Einen Eigenvektor  $\mathbf{b}_1 \neq \mathbf{0}$  zum Eigenwert  $-1$  setzen wir an in der Form

$$\mathbf{b}_1 = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Er hat  $A\mathbf{b}_1 = (-1)\mathbf{b}_1$  zu erfüllen, also  $(A - (-1)\mathbf{1}_2)\mathbf{b}_1 = \mathbf{0}$ , also

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also  $x + y = 0$ . Wir können somit  $x = 1$ ,  $y = -1$  wählen. Einen Eigenvektor  $\mathbf{b}_2 \neq \mathbf{0}$  zum Eigenwert  $3$  setzen wir an in der Form

$$\mathbf{b}_2 = \begin{pmatrix} u \\ v \end{pmatrix}.$$

Er hat  $(A - 3\mathbf{1}_2)\mathbf{b}_2 = \mathbf{0}$  zu erfüllen, also

$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also  $u - v = 0$ . Wir können somit  $u = 1, v = 1$  wählen. Mit den Eigenvektoren

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{und} \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{setzen wir} \quad S := (\mathbf{b}_1 \ \mathbf{b}_2) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Wie in (37) gezeigt, ist dann

$$S^{-1}AS = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

die Diagonalmatrix mit den zu  $\mathbf{b}_1$  bzw.  $\mathbf{b}_2$  gehörigen Eigenwerten auf der Diagonalen.

**6.12** Wir zeigen nun die verbleibende Aussage von 6.7: Ist  $A \in \mathbb{R}^{n \times n}$  und gibt es eine invertierbare Matrix  $S \in \mathbb{R}^{n \times n}$  derart, dass  $D := S^{-1}AS$  eine Diagonalmatrix ist, so ist  $A$  diagonalisierbar. Sei  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Wir zeigen zudem:

*Die Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$  von  $S$  bilden eine Basis von Eigenvektoren für  $A$ . Die zugehörigen Eigenwerte sind die Diagonaleinträge  $\lambda_1, \dots, \lambda_n$  von  $D$ .*

Zum Nachweis multiplizieren wir  $S^{-1}AS = D$  von links mit  $S$  und erhalten

$$AS = SD.$$

Anwenden der linken und rechten Seite auf den Standard-Einheitsvektor  $\mathbf{e}_j$  liefert

$$A\mathbf{b}_j = ASe_j = SDe_j = \lambda_j S\mathbf{e}_j = \lambda_j \mathbf{b}_j.$$

Also ist  $\mathbf{b}_j$  Eigenvektor zum Eigenwert  $\lambda_j$ . Da  $S$  invertierbar ist, bilden  $\mathbf{b}_1, \dots, \mathbf{b}_n$  eine Basis für  $\mathbb{R}^n$ . Also ist  $A$  diagonalisierbar.

Nicht jede quadratische Matrix ist diagonalisierbar.

**Beispiel 6.13** Wir zeigen, dass die Matrix

$$A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

nicht diagonalisierbar ist. Das charakteristische Polynom ist

$$\det(A - \lambda \mathbf{1}_2) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2$$

mit der doppelten Nullstelle 1. Es ist also  $\lambda = 1$  der einzige Eigenwert. Wäre  $A$  diagonalisierbar, so gäbe es nach 6.7 eine invertierbare Matrix  $S \in \mathbb{R}^{2 \times 2}$  und eine Diagonalmatrix  $D \in \mathbb{R}^{2 \times 2}$  derart, dass  $S^{-1}AS = D$ . Nach 6.12 sind alle Diagonaleinträge von  $D$  Eigenwerte von  $A$ , sie müssten also alle gleich 1 sein. Somit wäre  $D = \mathbf{1}_2$  und folglich

$$A = SDS^{-1} = S\mathbf{1}_2S^{-1} = SS^{-1} = \mathbf{1}_2,$$

Widerspruch. Also kann  $A$  nicht diagonalisierbar sein.

**Definition 6.14** Es sei  $E$  ein endlich-dimensionaler Vektorraum und  $\phi: E \rightarrow E$  ein Endomorphismus. Für  $\lambda \in \mathbb{R}$  ist

$$E_\lambda(\phi) := \ker(\phi - \lambda \text{id}_E)$$

ein Untervektorraum von  $E$ . Alle von Null verschiedenen Vektoren  $\mathbf{v} \in E_\lambda(\phi)$  sind dann Eigenvektoren für  $\phi$  zum Eigenwert  $\lambda$ . Also ist  $E_\lambda(\phi) \neq \{\mathbf{0}\}$  genau dann, wenn  $\lambda$  ein Eigenwert für  $\phi$  ist. In diesem Fall nennen wir  $E_\lambda(\phi)$  den *Eigenraum* von  $\phi$  zum Eigenwert  $\lambda$ .

Ist  $E = \mathbb{R}^n$  und  $\phi = \phi_A$  für eine Matrix  $A \in \mathbb{R}^{n \times n}$ , schreiben wir einfach  $E_\lambda(A)$  statt  $E_\lambda(\phi_A)$ .

**6.15** *Es sei  $A \in \mathbb{R}^{n \times n}$  und es seien  $\lambda_1, \dots, \lambda_m$  die paarweise verschiedenen reellen Eigenwerte für  $A$ . Genau dann ist  $A$  diagonalisierbar, wenn*

$$\sum_{j=1}^m \dim E_{\lambda_j}(A) = \dim(\mathbb{R}^n) = n.$$

**Beispiel 6.16** Wir wollen klären, ob die Matrix

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

diagonalisierbar ist. Hierzu berechnen wir die Eigenwerte für  $A$ , die Eigenräume und ihre Dimensionen. Diagonalisierbarkeit liegt vor, wenn die Summe der Dimensionen gleich 3 ist,

Das charakteristische Polynom ist

$$\det(A - \lambda \mathbf{1}_3) = \det \begin{pmatrix} -\lambda & 2 & 1 \\ 2 & -\lambda & -1 \\ 0 & 0 & 2 - \lambda \end{pmatrix} = (2 - \lambda)(\lambda^2 - 4) = -(\lambda - 2)^2(\lambda + 2),$$

wobei die Determinante nach der letzten Zeile entwickelt wurde. Wir haben  $\lambda_1 = 2$  als doppelte Nullstelle,  $\lambda_2 = -2$  als einfache Nullstelle; dies sind die Eigenwerte. Der Eigenraum zum Eigenwert  $\lambda_1 = 2$  besteht aus allen Lösungen des homogenen linearen Gleichungssystems  $(A - 2\mathbf{1}_3)\mathbf{x} = \mathbf{0}$ , also

$$\begin{pmatrix} -2 & 2 & 1 \\ 2 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0},$$

mit den Komponenten  $x, y, z$  von  $\mathbf{x}$ . Durch Addieren der ersten Zeile zur zweiten bringen wir dies auf die Zeilenstufenform

$$\begin{pmatrix} -2 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$$

und erhalten mit den freien Variablen  $z = s$ ,  $y = t$  die Bedingung  $x = y + z/2 = t + s/2$  und somit Eigenraum  $E_2(A)$  der Vektoren

$$\begin{pmatrix} t + s/2 \\ t \\ s \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix}.$$

Die zwei gezeigten Spaltenvektoren sind linear unabhängig, bilden also eine Basis des  $E_2(A)$ . Somit ist  $\dim E_2(A) = 2$ . Für  $\lambda_2 = -2$  besteht der Eigenraum  $E_{-2}(A)$  aus allen Vektoren  $\mathbf{x}$  mit  $(A + 2\mathbf{1}_3)\mathbf{x} = \mathbf{0}$ , also

$$\begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}.$$

Subtrahieren der ersten Zeile von der zweiten führt auf

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$$

und zweimaliges Addieren der zweiten zur dritten Zeile schließlich auf die Zeilenstufenform

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}.$$

Es ist also  $z = 0$  und mit  $y = s$  als freier Variable haben wir die Lösungen

$$s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

mit  $s \in \mathbb{R}$ , die einen eindimensionalen Eigenraum  $E_{-2}(A)$  bilden. Da

$$\dim E_2(A) + \dim E_{-2}(A) = 2 + 1 = 3,$$

ist  $A$  diagonalisierbar.

**Beispiel 6.17** Zeige durch Berechnen der Eigenräume, dass die Matrix

$$A := \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

nicht diagonalisierbar ist.

Lösung: Das charakteristische Polynom von  $A$  ist

$$\det(A - \lambda \mathbf{1}_2) = \det \begin{pmatrix} 1 - \lambda & 1 \\ -1 & 3 - \lambda \end{pmatrix} = (1 - \lambda)(3 - \lambda) + 1 = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2;$$

es hat die doppelte Nullstelle  $\lambda = 2$ . Der Eigenraum  $E_2(A)$  zum Eigenwert 2 besteht aus allen Lösungen

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

des homogenen linearen Gleichungssystems  $(A - 2\mathbf{1}_2)\mathbf{x} = \mathbf{0}$ , also

$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Durch Subtrahieren der ersten von der zweiten Zeile lässt sich dieses auf die Zeilenstufenform

$$\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

bringen. Mit  $y = s \in \mathbb{R}$  als freier Variablen haben wir  $x = y = s$  und somit einen eindimensionalen Lösungsraum (und Eigenraum) bestehend aus den Vektoren

$$s \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

mit  $s \in \mathbb{R}$ . Die Summe der Dimensionen aller Eigenräume ist also 1 und somit  $< 2 = \dim(\mathbb{R}^2)$ . Nach 6.15 ist  $A$  nicht diagonalisierbar.

Um 6.15 zu begründen, nutzt uns:

**6.18** Für  $j \in \{1, \dots, m\}$  sei  $\mathbf{b}_{j,1}, \dots, \mathbf{b}_{j,n_j}$  eine Basis für  $E_{\lambda_j}(A)$ . Dann sind die Vektoren  $\mathbf{b}_{j,i}$  für  $j \in \{1, \dots, m\}$  und  $i \in \{1, \dots, n_j\}$  in  $\mathbb{R}^n$  linear unabhängig. Folglich ist

$$\sum_{j=1}^m \dim(E_{\lambda_j}(A)) = \sum_{j=1}^m n_j \leq n.$$

**6.19** Begründung für 6.18: Wir zeigen per Induktion nach  $k \in \{1, \dots, m\}$ , dass die Vektoren  $\mathbf{b}_{j,i}$  für  $j \in \{1, \dots, k\}$  und  $i \in \{1, \dots, n_j\}$  linear unabhängig sind (Aussage (\*)). Für  $k = 1$  ist das wahr, denn  $\mathbf{b}_{1,1}, \dots, \mathbf{b}_{1,n_1}$  ist eine Basis für  $E_{\lambda_1}(A)$ . Ist  $k \in \{2, \dots, m\}$  und gilt die Aussage (\*) für  $k - 1$  an Stelle von  $k$ , so seien  $t_{ij} \in \mathbb{R}$  mit

$$\sum_{j=1}^k \sum_{i=1}^{n_j} t_{ji} \mathbf{b}_{ji} = \mathbf{0}.$$

Wären nicht alle  $t_{ji} = 0$ , so muss ein  $t_{ki} \neq 0$  sein. Andernfalls wäre nämlich  $\sum_{j=1}^{k-1} \sum_{i=1}^{n_j} t_{ji} \mathbf{b}_{ji} = \mathbf{0}$  eine nicht-triviale Linearkombination mit Summe  $\mathbf{0}$ ,

im Widerspruch zur linearen Unabhängigkeit der  $\mathbf{b}_{ji}$  mit  $j \leq k-1$ . Setzen wir

$$\mathbf{w}_j := \sum_{i=1}^{n_j} t_{ji} \mathbf{b}_{ji}$$

für  $j \in \{1, \dots, k\}$ , so ist als  $\mathbf{w}_k \neq \mathbf{0}$ . Nun ist  $\mathbf{w}_j \in E_{\lambda_j}(A)$  für  $j \in \{1, \dots, k\}$ , somit

$$A\mathbf{w}_k = \lambda_k \mathbf{w}_k. \quad (38)$$

Andererseits ist  $\mathbf{w}_k = -\sum_{j=1}^{k-1} \mathbf{w}_j$ , also

$$A\mathbf{w}_k = -\sum_{j=1}^{k-1} A\mathbf{w}_j = -\sum_{j=1}^{k-1} \lambda_j \mathbf{w}_j.$$

Wir können (38) noch umschreiben zu;

$$A\mathbf{w}_k = -\sum_{j=1}^{k-1} \lambda_k \mathbf{w}_j.$$

Subtraktion der vorigen zwei Gleichungen liefert

$$\mathbf{0} = \sum_{j=1}^{k-1} (\lambda_k - \lambda_j) \mathbf{w}_j = \sum_{j=1}^{k-1} \sum_{i=1}^{n_j} t_{ji} (\lambda_k - \lambda_j) \mathbf{b}_{ji}.$$

Da  $\lambda_k - \lambda_j \neq 0$ , folgt  $t_{ji} = 0$  für alle  $j \in \{1, \dots, k-1\}$  und  $i \in \{1, \dots, n_j\}$ . Somit wäre  $\mathbf{w}_k = -\sum_{j=1}^{k-1} \sum_{i=1}^{n_j} t_{ji} \mathbf{b}_{ji} = \mathbf{0}$ , Widerspruch.

**6.20** Begründung für 6.15: Ist  $\sum_{j=1}^m \dim E_{\lambda_j}(A) = n$ , so wählen wir für jedes  $j \in \{1, \dots, m\}$  eine Basis  $\mathbf{b}_{j1}, \dots, \mathbf{b}_{jn_j}$  für  $E_{\lambda_j}(A)$ . Dann ist also  $n_j = \dim E_{\lambda_j}(A)$ . Nach 6.18 sind die Vektoren  $\mathbf{b}_{ji}$  für  $j \in \{1, \dots, m\}$  und  $i \in \{1, \dots, n_j\}$  linear unabhängig. Ihre Anzahl ist per Voraussetzung  $\sum_{j=1}^m n_j = n = \dim(\mathbb{R}^n)$ ; die genannten linear unabhängigen Vektoren bilden somit eine Basis für  $\mathbb{R}^n$ . Da jeder der Basisvektoren ein Eigenvektor für  $A$  ist, ist  $A$  diagonalisierbar.

Ist  $A$  diagonalisierbar, so finden wir eine Basis für  $\mathbb{R}^n$  aus Eigenvektoren für  $A$ . Für  $j \in \{1, \dots, m\}$  seien  $\mathbf{v}_{ji}$  mit  $i \in \{1, \dots, \ell_j\}$  diejenigen der Basisvektoren, die Eigenvektor zum Eigenwert  $\lambda_j$  sind. Diese sind linear unabhängig und liegen in  $E_{\lambda_j}(A)$ . Für ihre Anzahl gilt also

$$\ell_j \leq \dim E_{\lambda_j}(A).$$

Wir folgern, dass

$$n = \sum_{j=1}^m \ell_j \leq \sum_{j=1}^m \dim E_{\lambda_j}(A).$$

Da die rechte Seite nach 6.18  $\leq n$  ist, muss Gleichheit herrschen:  $n = \sum_{j=1}^m \dim E_{\lambda_j}(A)$ .

### Diagonalisieren über $\mathbb{C}$

In einem Vektorraum  $V$  über  $\mathbb{C}$  kann man nicht nur mit reellen Skalaren multiplizieren, sondern hat eine Multiplikation

$$\mathbb{C} \times V \rightarrow V, \quad (\lambda, \mathbf{v}) \mapsto \lambda \mathbf{v}$$

gegeben. In den Axiomen für Vektorräume ersetze man einfach  $\mathbb{R}$  durch  $\mathbb{C}$ . Zum Beispiel ist der Raum  $\mathbb{C}^n$  der Spaltenvektoren

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

mit  $z_1, \dots, z_n \in \mathbb{C}$  ein komplexer Vektorraum. Lineare Gleichungssysteme, Matrizen  $A \in \mathbb{C}^{m \times n}$ , komplex lineare Abbildungen, lineare Unabhängigkeit, Basen usw. definiert man im Komplexen wie im Reellen – wir ersetzen nur überall  $\mathbb{R}$  durch  $\mathbb{C}$ . Ebenso können wir von Endomorphismen eines endlich-dimensionalen komplexen Vektorraums sprechen, von Eigenvektoren (wobei nun Eigenwerte  $\lambda \in \mathbb{C}$  erlaubt sind) und von Diagonalisierbarkeit. In allen bisherigen Definitionen und Ergebnissen kann einfach  $\mathbb{R}$  durch  $\mathbb{C}$  ersetzt werden.

**Beispiel 6.21** Wir betrachten die Matrix  $A \in \mathbb{R}^{2 \times 2}$ ,

$$A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Diese Matrix ist über  $\mathbb{R}$  nicht diagonalisierbar, denn das charakteristische Polynom

$$\det(A - \lambda \mathbf{1}_2) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$$

hat keine reellen Nullstellen; es gibt somit keine reellen Eigenwerte und somit in  $\mathbb{R}^2$  keine Eigenvektoren. Jedoch ist  $A$  über  $\mathbb{C}$  diagonalisierbar: Es sind  $\lambda_1 = i$  und  $\lambda_2 = -i$  komplexe Nullstellen des charakteristischen Polynoms und somit komplexe Eigenwerte. Eigenvektoren zum Eigenwert  $i$  sind nicht-verschwindende Lösungen zu

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Mit dem komplexen Parameter  $s \in \mathbb{C}$  ist also  $ix = y = s$  verlangt und somit  $x = (1/i)s$ . Mit  $s = 1$  erhalten wir einen Eigenvektor

$$\begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Analog erhalten wir als einen Eigenvektor zum Eigenwert  $-i$

$$\begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Da dieser und der vorige über  $\mathbb{C}$  linear unabhängig sind, haben wir eine Basis aus Eigenvektoren für  $\mathbb{C}^2$ . Also ist  $A$  über  $\mathbb{C}$  diagonalisierbar.

## Darstellungsmatrizen für lineare Abbildungen

**Definition 6.22** Es sei  $\phi: V \rightarrow W$  eine lineare Abbildung zwischen endlich-dimensionalen (reellen) Vektorräumen. Ist  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  eine Basis für  $V$  und  $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$  eine Basis für  $W$ , so ist für jedes  $j \in \{1, \dots, n\}$

$$\phi(\mathbf{b}_j) = \sum_{i=1}^m a_{ij} \mathbf{c}_i$$

mit eindeutig festgelegten  $a_{1j}, \dots, a_{mj} \in \mathbb{R}$ . Man nennt

$$M_{CB}(\phi) := (a_{ij}) \in \mathbb{R}^{m \times n}$$

die *darstellende Matrix* für  $\phi$  bezüglich den Basen  $B$  von  $V$  und  $C$  von  $W$ .

**Beispiel 6.23** Es sei  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung. Ist

$$A := (a_{ij}) := M_{CB}(\phi)$$

die darstellende Matrix für  $\phi$  bezüglich der Standard-Basis  $B := (\mathbf{e}_1, \dots, \mathbf{e}_n)$  von  $\mathbb{R}^n$  und der entsprechenden Standard-Basis  $C := (\mathbf{f}_1, \dots, \mathbf{f}_m)$  von  $\mathbb{R}^m$ , so ist für jedes  $j \in \{1, \dots, m\}$

$$\phi(\mathbf{e}_j) = \sum_{i=1}^m a_{ij} \mathbf{f}_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Die  $j$ te Spalte von  $A$  ist also gleich dem Bild  $\phi(\mathbf{e}_j)$  des  $j$ ten Standard-Einheitsvektors.

**6.24** *Es sei  $\phi: V \rightarrow V$  ein Endomorphismus eines endlich-dimensionalen Vektorraums  $V$ . Eine Basis  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  von  $V$  besteht genau dann aus Eigenvektoren, wenn  $M_{BB}(\phi)$  eine Diagonalmatrix ist.*

Für jedes  $j \in \{1, \dots, n\}$  ist

$$\phi(\mathbf{b}_j) = \sum_{i=1}^n a_{ij} \mathbf{b}_i$$

mit eindeutigen  $a_{1j}, \dots, a_{nj} \in \mathbb{R}$ . Genau dann ist  $\phi(\mathbf{b}_j) = \lambda_j \mathbf{b}_j$  für ein  $\lambda_j \in \mathbb{R}$ , wenn  $a_{ij} = 0$  für alle  $i \in \{1, \dots, n\}$  mit  $i \neq j$ . Dies gilt genau dann für alle  $j$  (so dass  $B$  eine Basis aus Eigenvektoren ist), wenn  $A = (a_{ij})$  eine Diagonalmatrix ist.

### Mehr zu Darstellungsmatrizen\*

Die folgenden Rechenregeln wurden am 20.5. für die Allgemeinbildung erwähnt.

**6.25** *Es seien  $U, V$  und  $W$  endlich-dimensionale Vektorräumen mit Basen  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ ,  $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$  bzw.  $D = (\mathbf{d}_1, \dots, \mathbf{d}_\ell)$ . Dann gilt:*

(a) *Es ist  $M_{BB}(\text{id}_U) = \mathbf{1}_n$ .*

(b) *Sind  $\phi: U \rightarrow V$  und  $\psi: V \rightarrow W$  lineare Abbildungen, so ist*

$$M_{DB}(\psi \circ \phi) = M_{DC}(\psi) M_{CB}(\phi) \quad (\text{Produkt der Matrizen})$$

(c) *Ist  $\phi: U \rightarrow V$  ein Isomorphismus, so ist  $M_{CB}(\phi) \in \mathbb{R}^{n \times n}$  eine invertierbare Matrix mit inverser Matrix*

$$(M_{CB}(\phi))^{-1} = M_{BC}(\phi^{-1}).$$

Weiter gilt:

- (d) Sei  $C = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  die Standardbasis für  $\mathbb{R}^n$  und  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  eine weitere Basis. Dann ist

$$M_{CB}(\text{id}_{\mathbb{R}^n})$$

die Matrix mit den Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$ .

- (e) Sei  $A \in \mathbb{R}^{m \times n}$  und  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \mapsto A\mathbf{x}$  die zugehörige lineare Abbildung. Für die Standardbasis und  $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  von  $\mathbb{R}^n$  und die Standardbasis  $C = (\mathbf{f}_1, \dots, \mathbf{f}_m)$  für  $\mathbb{R}^m$  gilt dann

$$M_{CB}(\phi_A) = A.$$

Nachweis: (a) Es ist  $\text{id}_U(\mathbf{b}_j) = \mathbf{b}_j = \sum_{i=1}^n a_{ij} \mathbf{b}_i$  mit  $a_{jj} = 1$ ,  $a_{ij} = 0$  für  $i \neq j$ . Also ist  $M_{BB}(\text{id}_U) = (\delta_{ij}) = \mathbf{1}_n$ .

(b) Sei  $M_{CB}(\phi) = (b_{jk})$ ,  $M_{DC}(\psi) = (a_{ij})$  und  $M_{DB}(\psi \circ \phi) = (c_{ik})$ . Für  $k \in \{1, \dots, n\}$  ist

$$\phi(\mathbf{b}_k) = \sum_{j=1}^m b_{jk} \mathbf{c}_j,$$

also

$$\psi(\phi(\mathbf{b}_k)) = \sum_{j=1}^m b_{jk} \psi(\mathbf{c}_j) = \sum_{i=1}^{\ell} \sum_{j=1}^m a_{ij} b_{jk} \mathbf{d}_i.$$

Somit ist  $c_{ik} = \sum_{j=1}^m a_{ij} b_{jk}$ , also  $M_{DB}(\psi \circ \phi) = M_{DC}(\psi) M_{CB}(\phi)$ .

(c) Nach (a) und (b) ist

$$\mathbf{1}_n = M_{BB}(\text{id}_U) = M_{BB}(\phi^{-1} \circ \phi) = M_{BC}(\phi^{-1}) M_{CB}(\phi).$$

Analog ist

$$\mathbf{1}_n = M_{CB}(\phi) M_{BC}(\phi^{-1}).$$

Also ist die Matrix  $M_{CB}(\phi)$  invertierbar mit Inverser  $M_{BC}(\phi^{-1})$ .

(d) Ist

$$\mathbf{b}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

so ist  $\text{id}_{\mathbb{R}^n}(\mathbf{b}_j) = \mathbf{b}_j = \sum_{i=1}^n a_{ij} \mathbf{e}_i$  und somit  $M_{CB}(\text{id}_{\mathbb{R}^n}) = (a_{ij})$ , mit  $j$ ter Spalte  $\mathbf{b}_j$ .

(e) Es ist  $\phi_A(\mathbf{e}_j) = A\mathbf{e}_j = \sum_{i=1}^m a_{ij} \mathbf{f}_i$ , also  $M_{CB}(\phi_A) = A$ .

**Beispiel 6.26** Sei  $A \in \mathbb{R}^{n \times n}$  diagonalisierbar und  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  eine Basis aus Eigenvektoren mit den zugehörigen reellen Eigenwerten  $\lambda_1, \dots, \lambda_n$ . Wegen

$$A\mathbf{b}_j = \lambda_j\mathbf{b}_j$$

ist dann  $M_{BB}(A_\phi) = \text{diag}(\lambda_1, \dots, \lambda_n) =: D$ . Für die Standardbasis  $C = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  von  $\mathbb{R}^n$  ist zudem  $M_{CC}(\phi_A) = A$ , nach 6.25 (e). Schließlich ist nach 6.25 (d)

$$M_{CB}(\text{id}_{\mathbb{R}^n}) =: S$$

die Matrix mit den Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$ . Nach 6.25 (b) und (c) ist wegen  $\phi_A = \text{id}_{\mathbb{R}^n} \circ \phi_A \circ \text{id}_{\mathbb{R}^n}$  und  $\text{id}_{\mathbb{R}^n} = (\text{id}_{\mathbb{R}^n})^{-1}$  nun

$$A = M_{CC}(\phi_A) = M_{CB}(\text{id}_{\mathbb{R}^n})M_{BB}(\phi_A)M_{BC}(\text{id}_{\mathbb{R}^n}) = SDS^{-1},$$

bzw.  $D = S^{-1}AS$ . Dies liefert eine neue Begründung zu den oben diskutierten Formeln.

## Skalarprodukt und Orthogonalität

**6.27** In Analogie zum in der Mathematik I diskutierten Skalarprodukt auf  $\mathbb{R}^2$  und  $\mathbb{R}^3$  definieren wir für Spaltenvektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

in  $\mathbb{R}^n$  ihr Skalarprodukt (inneres Produkt) als

$$\mathbf{x} \cdot \mathbf{y} := \sum_{j=1}^n x_j y_j \in \mathbb{R}.$$

Eine andere Notation ist  $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x} \cdot \mathbf{y}$ . Ist  $\mathbf{x} \cdot \mathbf{y} = 0$ , nennen wir die Vektoren  $\mathbf{x}$  und  $\mathbf{y}$  zueinander orthogonal.

**6.28** Wir erinnern an die transponierte Matrix  $A^T \in \mathbb{R}^{n \times m}$  einer Matrix  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ : Es ist  $A^T = (a_{ji})$ . Beispielsweise ist für einen Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

die transponierte Matrix der Zeilenvektor  $\mathbf{x}^T = (x_1 \dots x_n)$ . Für alle  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  ist

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

im Sinne eines Matrixprodukts von Zeilenmatrix mal Spaltenmatrix.

**6.29** Das Skalarprodukt von

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

in  $\mathbb{C}^n$  definieren wir als

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{j=1}^n \overline{x_j} y_j \in \mathbb{C}$$

unter Benutzung der komplexen Konjugation.

**6.30** Für eine Matrix  $A = (a_{jk}) \in \mathbb{C}^{m \times n}$  definieren wir die hermitesch konjugierte Matrix als

$$A^* := (\overline{a_{kj}}) \in \mathbb{C}^{n \times m}.$$

Wir bilden also die transponierte Matrix und ersetzen anschließend jeden Matrixeintrag durch die dazu komplex konjugierte Zahl. Für alle  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$  ist dann

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y}.$$

### Besondere Arten von Matrizen

**6.31** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *orthogonale Matrix*, wenn  $A^T A = \mathbf{1}_n$ . Wegen  $\det(A)^2 = \det(A^T) \det(A) = \det(A^T A) = \det(\mathbf{1}_n) = 1$  ist dann  $A$  invertierbar und  $A^{-1} = A^T$ . Sind  $\mathbf{b}_1, \dots, \mathbf{b}_n$  die Spalten von  $A$ , so ist

$$A^T A = (\mathbf{b}_j \cdot \mathbf{b}_k)_{j,k=1}^n$$

gleich  $\mathbf{1}_n = (\delta_{jk})_{j,k=1}^n$  genau dann, wenn  $\mathbf{b}_j \cdot \mathbf{b}_k = \delta_{jk}$  (Kronecker-Delta) für alle  $j, k \in \{1, \dots, n\}$ . Also gilt:

*Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist genau dann orthogonal, wenn ihre Spalten eine Orthonormalbasis von  $\mathbb{R}^n$  bilden.*

**6.32** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *symmetrische Matrix*, wenn  $A = A^T$ .

**6.33** Es lässt sich zeigen, dass jede symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  über  $\mathbb{R}$  diagonalisierbar ist und es eine Orthonormalbasis  $\mathbf{b}_1, \dots, \mathbf{b}_n$  von Eigenvektoren in  $\mathbb{R}^n$  gibt. Die Matrix  $S$  mit den Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$  ist dann eine orthogonale Matrix und es ist

$$S^{-1}AS = S^TAS =: D$$

eine Diagonalmatrix mit reellen Diagonaleinträgen.

Die entsprechenden Begriffe im Komplexen sind wie folgt:

**6.34** Eine Matrix  $A \in \mathbb{C}^{n \times n}$  heißt *unitäre Matrix*, wenn  $A^*A = \mathbf{1}_n$ . Wegen  $|\det(A)|^2 = \overline{\det A} \det A = \det(A^*) \det(A) = \det(A^*A) = \det(\mathbf{1}_n) = 1$  ist dann  $A$  invertierbar und  $A^{-1} = A^*$ . Sind  $\mathbf{b}_1, \dots, \mathbf{b}_n$  die Spalten von  $A$ , so ist

$$A^*A = (\langle \mathbf{b}_j, \mathbf{b}_k \rangle)_{j,k=1}^n$$

gleich  $\mathbf{1}_n = (\delta_{jk})_{j,k=1}^n$  genau dann, wenn  $\langle \mathbf{b}_j, \mathbf{b}_k \rangle = \delta_{jk}$  (Kronecker-Delta) für alle  $j, k \in \{1, \dots, n\}$ . Also gilt:

*Eine Matrix  $A \in \mathbb{C}^{n \times n}$  ist genau dann unitär, wenn ihre Spalten eine Orthonormalbasis von  $\mathbb{C}^n$  bilden.*

**6.35** Eine Matrix  $A \in \mathbb{C}^{n \times n}$  heißt *hermitesche Matrix*, wenn  $A = A^*$ .

**6.36** Es lässt sich zeigen, dass jede hermitesche Matrix  $A \in \mathbb{C}^{n \times n}$  über  $\mathbb{C}$  diagonalisierbar ist und es eine Orthonormalbasis  $\mathbf{b}_1, \dots, \mathbf{b}_n$  von Eigenvektoren in  $\mathbb{C}^n$  gibt. Alle Eigenwerte sind reell. Die Matrix  $S$  mit den Spalten  $\mathbf{b}_1, \dots, \mathbf{b}_n$  ist dann eine unitäre Matrix und es ist

$$S^{-1}AS = S^*AS =: D$$

eine Diagonalmatrix mit reellen Diagonaleinträgen.

Eine weitere Rechenregel ist manchmal nützlich.

**6.37** Sind  $A \in \mathbb{R}^{\ell \times m}$  und  $B \in \mathbb{R}^{m \times n}$ , so ist

$$(AB)^T = B^T A^T,$$

wie man sofort nachrechnet.

Daraus folgt:

**6.38** Für alle  $A \in \mathbb{R}^{n \times n}$  und  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  ist

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^T \mathbf{y} \rangle.$$

Es ist nämlich

$$\langle A\mathbf{x}, \mathbf{y} \rangle = (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \langle \mathbf{x}, A^T \mathbf{y} \rangle.$$

**6.39** Analog gilt

$$(AB)^* = B^* A^*$$

für alle  $A \in \mathbb{C}^{\ell \times m}$  und  $B \in \mathbb{C}^{m \times n}$ . Man folgert analog zum Obigen, dass

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^* \mathbf{y} \rangle$$

für alle  $A \in \mathbb{C}^{n \times n}$  und  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ .

## 7 Abstände in $\mathbb{R}^n$ , konvergente Folgen und stetige Funktionen

Im Folgenden schreiben wir einfach  $x$  statt  $\mathbf{x}$  für Elemente von  $\mathbb{R}^n$ , so dass also

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \dots x_n)^T.$$

Wir schreiben auch einfach  $x = (x_1, \dots, x_n)$ . Wann immer Matrizen und Vektoren multipliziert werden, müssen wir präzise sein: Dann ist  $x$  als Spaltenvektor zu lesen und  $x^T$  der zugehörige Zeilenvektor.

**Definition 7.1** Die *Länge* eines Vektors  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  ist definiert als

$$|x| := \sqrt{(x_1)^2 + \dots + (x_n)^2}.$$

Es ist also

$$|x|^2 = x_1^2 + \dots + x_n^2 = x \cdot x,$$

das Skalarprodukt des Spaltenvektors  $x$  mit sich selbst.

Für den Nullvektor  $0 \in \mathbb{R}^n$  ist  $|0| = 0$ . Weiter ist für alle  $t \in \mathbb{R}$  und  $x \in \mathbb{R}^n$

$$|tx| = \sqrt{t^2 x_1^2 + \dots + t^2 x_n^2} = \sqrt{t^2} \sqrt{x_1^2 + \dots + x_n^2} = |t| |x|. \quad (39)$$

Insbesondere lässt sich jeder von Null verschiedene Vektor  $x \in \mathbb{R}^n$  *normieren*: Der Vektor  $\frac{1}{|x|}x$  hat Länge

$$\left| \frac{1}{|x|}x \right| = \frac{1}{|x|} |x| = 1,$$

ist also ein *Einheitsvektor*.

**7.2** Für alle  $x, y \in \mathbb{R}^n$  gilt die *Cauchy-Schwarzsche Ungleichung*

$$|x \cdot y| \leq |x| |y|$$

für den Betrag des Skalarprodukts  $x \cdot y = \langle x, y \rangle$  der Spaltenvektoren  $x$  und  $y$ .

Begründung: Ist  $x = 0$  oder  $y = 0$ , steht rechts und links Null und die Ungleichung ist trivial. Sei nun  $x \neq 0$  und  $y \neq 0$ . Teilen durch  $|x| |y|$  zeigt, dass die Ungleichung zu

$$\left| \frac{1}{|x|}x \cdot \frac{1}{|y|}y \right| \leq 1$$

äquivalent ist. Wir brauchen also nur den Fall zu behandeln, dass  $x$  und  $y$  Einheitsvektoren sind, also  $|x| = |y| = 1$ . Für alle  $t \in \mathbb{R}$  ist wegen  $x \cdot y = y \cdot x$  nun

$$\begin{aligned} 0 &\leq (x + ty) \cdot (x + ty) = x \cdot x + x \cdot ty + (tx) \cdot y + (ty) \cdot (ty) \\ &= |x|^2 + 2tx \cdot y + t^2|y|^2 = 1 + 2tx \cdot y + t^2 \\ &= (t + x \cdot y)^2 + 1 - (x \cdot y)^2. \end{aligned}$$

Das Minimum der letzten Zeile wird für  $t = -x \cdot y$  angenommen (dann ist das erste Quadrat gleich Null) und hat den Wert  $1 - (x \cdot y)^2$ . Es muss also

$$0 \leq 1 - (x \cdot y)^2$$

sein, woraus  $(x \cdot y)^2 \leq 1$  folgt. Somit ist  $|x \cdot y| \leq 1 = |x| |y|$ .

**7.3** Für alle  $x, y \in \mathbb{R}^n$  gilt die Dreiecksungleichung:

$$|x + y| \leq |x| + |y|.$$

Nach der Cauchy-Schwarzschen Ungleichung ist nämlich  $|x \cdot y| \leq |x| |y|$ , woraus

$$x \cdot y \leq |x \cdot y| \leq |x| |y|$$

folgt. Somit ist

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) = x \cdot x + x \cdot y + y \cdot x + y \cdot y \\ &= |x|^2 + 2x \cdot y + |y|^2 \leq |x|^2 + 2|x| \cdot |y| + |y|^2 \\ &= (|x| + |y|)^2. \end{aligned}$$

Da die Quadratwurzel eine monoton wachsende Funktion ist, bleibt die vorige Ungleichung bestehen, wenn wir die Wurzel ziehen:

$$|x + y| \leq |x| + |y|.$$

**Definition 7.4** Für  $x, y \in \mathbb{R}^n$  nennen wir  $|x - y|$  den *Abstand* von  $x$  und  $y$ . Gegeben  $x \in \mathbb{R}^n$  und eine reelle Zahl  $r > 0$  nennen wir

$$B_r(x) := \{y \in \mathbb{R}^n : |y - x| < r\}$$

die *offene Kugel vom Radius  $r$  um den Mittelpunkt  $x$* . Wir nennen

$$\overline{B}_r(x) := \{y \in \mathbb{R}^n : |y - x| \leq r\}$$

die *abgeschlossene Kugel vom Radius  $r$  um  $x$* .

Die abgeschlossene Kugel enthält also auch alle Punkte  $y$  mit Abstand  $|y - x| = r$  von  $x$ , also den Rand der Kugel (die Sphäre vom Radius  $r$  um  $x$ ). Die offene Kugel enthält die Randpunkte nicht.

**Definition 7.5** Wir sagen, eine Folge  $(x_k)_{k \in \mathbb{N}}$  von Punkten  $x_k \in \mathbb{R}^n$  *konvergiert* gegen ein  $x \in \mathbb{R}^n$ , wenn für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  existiert derart, dass für alle  $k \geq N$

$$|x - x_k| < \varepsilon.$$

Wir schreiben dann

$$\lim_{k \rightarrow \infty} x_k = x$$

oder auch:  $x_k \rightarrow x$  für  $k \rightarrow \infty$ .

Wir im Falle von Folgen in  $\mathbb{R}$  in der Mathematik 1 sieht man, dass Grenzwerte von Folgen in  $\mathbb{R}^n$  immer eindeutig sind.

**7.6** Eine Folge  $(x_k)_{k \in \mathbb{N}}$  von Elementen  $x_k = (x_{k,1}, \dots, x_{k,n}) \in \mathbb{R}^n$  konvergiert genau dann gegen  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ , wenn jede der Komponenten konvergiert, also

$$x_{j,k} \rightarrow y_j \quad \text{für } k \rightarrow \infty$$

für alle  $j \in \{1, \dots, n\}$ .

**Beispiel 7.7**  $x_k = (\frac{1}{k}, 2 + \frac{1}{k^2})$  konvergiert in  $\mathbb{R}^2$  gegen  $(0, 2)$  für  $k \rightarrow \infty$ , denn

$$\frac{1}{k} \rightarrow 0 \quad \text{und} \quad 2 + \frac{1}{k^2} \rightarrow 2.$$

Begründung für 7.6: Sei  $\varepsilon > 0$ . Gilt  $x_k \rightarrow y$ , so gibt es ein  $N \in \mathbb{N}$  derart, dass  $|y - x_k| < \varepsilon$  für alle  $k \geq N$ . Für die  $j$ te Komponente haben wir für alle  $k \geq N$  dann

$$|y_j - x_{k,j}| = \sqrt{(y_j - x_{k,j})^2} \leq \sqrt{(y_1 - x_{k,1})^2 + \cdots + (y_n - x_{k,n})^2} = |y - x_k| < \varepsilon,$$

also  $|y_j - x_{k,j}| < \varepsilon$ . Somit gilt  $x_{k,j} \rightarrow y_j$ .

Konvergieren umgekehrt alle Komponenten, so gibt es für jedes  $j \in \{1, \dots, n\}$  ein  $N_j \in \mathbb{N}$  derart, dass  $|y_j - x_{k,j}| < \varepsilon/n$  für alle  $k \geq N_j$ . Für alle  $k \geq N := \max\{N_1, \dots, N_n\}$  ist dann

$$|y - x_k| = \left| \sum_{j=1}^n (y_j - x_{k,j})e_j \right| \leq \sum_{j=1}^n |(y_j - x_{k,j})e_j| = \sum_{j=1}^n |y_j - x_{k,j}| < n\varepsilon/n = \varepsilon$$

mit den Standard-Basisvektoren  $e_1, \dots, e_n \in \mathbb{R}^n$ . Also gilt  $x_k \rightarrow y$ .

Wir betrachten nun Funktionen  $f: U \rightarrow \mathbb{R}^m$ , die auf einer Teilmenge  $U \subseteq \mathbb{R}^n$  definiert sind. Jedem  $x = (x_1, \dots, x_n) \in U$  wird also ein Spaltenvektor

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

in  $\mathbb{R}^m$  zugeordnet. Die so erhaltenen reellwertigen Funktionen

$$f_1, \dots, f_m: U \rightarrow \mathbb{R}$$

werden die *Komponenten* von  $f$  genannt. Man schreibt auch

$$f = (f_1, \dots, f_m).$$

Der *Funktionsgraph* von  $f: U \rightarrow \mathbb{R}^m$  ist

$$\text{graph}(f) := \{(x, f(x)) : x \in U\} \subseteq \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}.$$

**Beispiel 7.8**  $n = 1, m = 2$ : Der Graph der Funktion

$$f: \mathbb{R} \rightarrow \mathbb{R}^2, \quad x \mapsto (\cos x, \sin x)$$

ist eine Schraubenlinie um die  $x$ -Achse (also eine Kurve im Raum). Diese sieht aus wie ein unendlich langer Korkenzieher.

Das Bild von  $f$  ist der Einheitskreis

$$\mathbb{S}_1 := \{(x_1, x_2) \in \mathbb{R}^2: x_1^2 + x_2^2 = 1\},$$

also eine Kurve in der Ebene. Der Kreis wird von den Punkten  $f(x)$  wieder und wieder durchlaufen.

**Beispiel 7.9**  $n = 2, m = 1$ : Wir betrachten die abgeschlossene Einheitskreisscheibe

$$U := \overline{B}_1(0) = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\}$$

und die Funktion  $f: U \rightarrow \mathbb{R}, (x, y) \mapsto \sqrt{1 - x^2 - y^2}$ . Für jeden Punkt  $(x, y, f(x, y))$  im Graphen ist dann

$$|(x, y, f(x, y))| = \sqrt{x^2 + y^2 + (1 - x^2 - y^2)} = \sqrt{1} = 1,$$

er liegt also auf der Einheitskugel

$$\mathbb{S}_2 := \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 = 1\}.$$

Da  $f(x, y) \geq 0$ , bekommen wir nur die obere Hälfte der Einheitskugel, die *obere Hemisphäre*. Der Graph von  $f$  ist also eine Fläche im Raum.

**Beispiel 7.10**  $n = m = 2$ : Wir betrachten die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(x, y) = (-y, x).$$

Den Graphen  $\text{graph}(f) \subseteq \mathbb{R}^4$  können wir nicht mehr zeichnen! Ersatz: Zeichne an der Stelle  $(x, y)$  den Vektor  $f(x, y)$ . Zum Beispiel ist

$$f(1, 0) = (0, 1), f(0, 1) = (-1, 0) \text{ und } f(\sqrt{2}/2, \sqrt{2}/2) = (-\sqrt{2}/2, \sqrt{2}/2);$$

wir können diese Vektoren jeweils an der Stelle  $(x, y)$  in der Zeichenebene einzeichnen (siehe Vorlesung).

Beachten Sie, dass  $(-y, x)$  durch Drehen um 90 Grad aus  $(x, y)$  hervorgeht.

In diesem Beispiel haben wir  $f$  als ein sogenannte *Vektorfeld* interpretiert: Jedem Punkt  $(x, y)$  der Ebene wird ein Vektor  $f(x, y) \in \mathbb{R}^2$  zugeordnet.

**Beispiel 7.11**  $n = m$ : Auch bijektive Abbildungen sind interessant (etwa beim Studium von Koordinatentransformationen beim Wechsel des Koordinatensystems), etwa bijektive Abbildungen

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Ein typisches Beispiel ist die zu einer invertierbaren Matrix  $A \in \mathbb{R}^{n \times n}$  gehörige lineare Abbildung  $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x \mapsto Ax$ . Wir können z.B. die Matrix

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

nehmen mit  $\alpha \in \mathbb{R}$ ; die lineare Abbildung  $\phi_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  ist dann eine Drehung der Ebene um den Ursprung im Gegenuhrzeigersinn um den Winkel  $\alpha$ . Nehmen wir

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$$

mit  $\alpha \in \mathbb{R}$ , so ist  $\phi_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $(x, y) \mapsto (x + \alpha y, y)$  eine *Scherung* der Ebene längs der  $x$ -Achse. Der Standard-Basisvektor  $e_1$  und jeder Punkt auf der  $x$ -Achse werden festgehalten. Für einen allgemeinen Punkt  $(x, y) \in \mathbb{R}^2$  bleibt die  $y$ -Komponente unverändert, die  $x$ -Komponente wird um  $\alpha y$ , also proportional zu  $y$ , abgeändert.

Eine Skizze wurde in der Vorlesung für  $\alpha = 1$  gegeben, so dass also  $f(x, y) = (x + y, y)$ .

Die Beispiele zeigen, welche Objekte wir nun angehen wollen: Kurven, Vektorfelder, bijektive Transformationen und (in Mathematik 3) Flächen. Wir wollen praxisrelevante Größen berechnen wir die Länge einer Kurve oder den Flächeninhalt einer Fläche. Dies gelingt natürlich nur für genügend schöne Funktionen  $f$ . Typischerweise wird  $f$  mindestens differenzierbar sein in einem geeigneten Sinn. Insbesondere ist  $f$  dann stetig und wir beginnen mit einer Diskussion dieser Eigenschaft.

**Definition 7.12** Eine Funktion  $f: U \rightarrow \mathbb{R}^m$  auf einer Teilmenge  $U \subseteq \mathbb{R}^n$  heißt *stetig* an einer Stelle  $x \in U$ , wenn

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

für jede Folge  $(x_k)_{k \in \mathbb{N}}$  von Punkten  $x_k \in U$  mit  $\lim_{k \rightarrow \infty} x_k = x$ . Die Funktion  $f$  heißt *stetig*, wenn sie an jeder Stelle  $x \in U$  stetig ist, also

$$f\left(\lim_{k \rightarrow \infty} x_k\right) = \lim_{k \rightarrow \infty} f(x_k)$$

für jede Folge  $(x_k)_{k \in \mathbb{N}}$  in  $U$ , die gegen einen Punkt aus  $U$  konvergiert.

**Beispiel 7.13** Für jedes  $j \in \{1, \dots, n\}$  ist die *Projektion*

$$\text{pr}_j: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto x_j$$

auf die  $j$ te Komponente stetig. Ist nämlich  $(x_k)_{k \in \mathbb{N}}$  eine konvergente Folge in  $\mathbb{R}^n$  mit Limes  $y \in \mathbb{R}^n$ , so konvergiert nach 7.6 die  $j$ te Komponente von  $x_k$  für  $k \rightarrow \infty$  gegen die  $j$ te Komponente von  $y$ , also

$$\lim_{k \rightarrow \infty} \text{pr}_j(x_k) = \text{pr}_j(y).$$

Also ist  $\text{pr}_j$  stetig.

**Beispiel 7.14** Sei  $U \subseteq \mathbb{R}^n$  eine Teilmenge. Dann ist die Inklusionsabbildung

$$f: U \rightarrow \mathbb{R}^n, \quad x \mapsto x$$

stetig. Für jede Folge  $(x_k)_{k \in \mathbb{N}}$  in  $U$ , die gegen ein  $x \in U$  konvergiert, konvergiert nämlich  $f(x_k) = x_k$  gegen  $x = f(x)$ , es ist also  $f(x) = \lim_{k \rightarrow \infty} f(x_k)$  erfüllt.

Wie im Falle von Funktionen einer reellen Variablen (in der Mathematik 1) sind aus stetigen Funktionen zusammengesetzte Funktionen wieder stetig. Insbesondere sind Verknüpfungen stetiger Funktionen stetig:

**7.15** *Es seien  $f: U \rightarrow \mathbb{R}^m$  und  $g: V \rightarrow \mathbb{R}^\ell$  Funktionen auf Teilmengen  $U \subseteq \mathbb{R}^n$  und  $V \subseteq \mathbb{R}^m$  derart, dass  $f(U) \subseteq V$ . Also ist die Komposition*

$$g \circ f: U \rightarrow \mathbb{R}^\ell, \quad y \mapsto g(f(y))$$

*definiert. Dann gilt: Ist  $f$  stetig an einer Stelle  $x \in U$  und  $g$  stetig an der Stelle  $f(x)$ , so ist  $g \circ f$  stetig an der Stelle  $x$ .*

Insbesondere gilt: Sind  $f$  und  $g$  stetig, so auch  $g \circ f$ .

Beweis: Sei  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $U$  mit  $x_k \rightarrow x$  für  $k \rightarrow \infty$ . Da  $f$  an der Stelle  $x$  stetig ist, folgt  $f(x_k) \rightarrow f(x)$  für  $k \rightarrow \infty$ , wobei  $f(x_k) \in f(U) \subseteq V$ . Da  $g$  an der Stelle  $f(x)$  stetig ist, folgt  $g(f(x_k)) \rightarrow g(f(x))$ .

**Beispiel 7.16** Für jedes  $k \in \mathbb{N}_0$  und  $j \in \{1, \dots, n\}$  ist das Monom

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x = (x_1, \dots, x_n) \mapsto (x_j)^k$$

stetig, dann es ist  $f = g \circ \text{pr}_j$  mit der stetigen Projektion  $\text{pr}_j: \mathbb{R}^n \rightarrow \mathbb{R}$  auf die  $j$ te Komponente und dem Monom  $g: \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t^k$  in einer Variablen. Aus der Mathematik 1 wissen wir, dass  $g$  stetig ist.

Für alle  $n, m \in \mathbb{N}$  gilt:

**Beispiel 7.17** Jede lineare Abbildung  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist stetig.

Zum Beispiel ist die identische Abbildung

$$\text{id}_{\mathbb{R}^n}: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto x$$

immer stetig (siehe auch 7.14).

Das folgende Hilfsmittel nutzt uns im Beweis und wird auch später noch wiederholt benutzt.

**Definition 7.18** Die *Operatornorm* einer Matrix  $A \in \mathbb{R}^{m \times n}$  ist die reelle Zahl

$$\|A\|_{\text{op}} := \sup\{|Ax|: x \in \mathbb{R}^n \text{ mit } |x| \leq 1\}.$$

Man beachte, dass  $\|A\|_{\text{op}} < \infty$ . Schreiben wir  $A = (a_{ij})$ , so ist für jeden Vektor  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  mit  $|x| \leq 1$  nämlich  $|x_j| \leq |x| \leq 1$  für alle  $j \in \{1, \dots, n\}$ , somit unter Benutzung der Standard-Basisvektoren  $e_1, \dots, e_n$

für  $\mathbb{R}^m$

$$\begin{aligned}
 |Ax| &= \left| \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e_i \right| \\
 &\leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij} x_j e_i| = \sum_{i=1}^m \sum_{j=1}^n |a_{ij} x_j| \underbrace{|e_i|}_{=1} \\
 &= \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| |x_j| \\
 &\leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| =: M.
 \end{aligned}$$

Also ist  $\|A\|_{\text{op}} \leq M < \infty$ .

**7.19** Für jedes Matrix  $A \in \mathbb{R}^{m \times n}$  und jeden Spaltenvektor  $x \in \mathbb{R}^n$  gilt

$$|Ax| \leq \|A\|_{\text{op}} |x|. \quad (40)$$

Ist  $x = 0$  so sind nämlich beide Seiten 0 und somit gleich. Ist  $x \neq 0$ , so ist  $\frac{1}{|x|}x$  ein Einheitsvektor, also gleich einem  $y \in \mathbb{R}^n$  mit  $|y| \leq 1$ . Folglich ist

$$\left| A \frac{1}{|x|} x \right| \leq \sup\{|Ay| : y \in \mathbb{R}^n \text{ mit } |y| \leq 1\} = \|A\|_{\text{op}}.$$

Die linke Seite ist nach 39 gleich  $\frac{1}{|x|}|Ax|$ . Multiplikation beider Seiten mit  $|x|$  führt auf (40).

**7.20** Beweis für 7.17. Jede lineare Abbildung  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist von der Form  $\phi = \phi_A$  für eine Matrix  $A \in \mathbb{R}^{m \times n}$ . Sei  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $\mathbb{R}^n$ , die gegen ein  $x \in \mathbb{R}^n$  konvergiert. Gegeben  $\varepsilon > 0$  existiert ein  $N \in \mathbb{N}$  derart, dass für alle natürlichen Zahlen  $k \geq N$

$$|x - x_k| < \frac{\varepsilon}{\|A\|_{\text{op}} + 1}.$$

Für alle  $k \geq N$  ist dann

$$\|\phi_A(x) - \phi_A(x_k)\| = |\phi_A(x - x_k)| = |A(x - x_k)| \leq \|A\|_{\text{op}} |x - x_k| \leq \frac{\|A\|_{\text{op}}}{\|A\|_{\text{op}} + 1} \varepsilon < \varepsilon.$$

Also gilt  $\phi_A(x_k) \rightarrow \phi_A(x)$  für  $k \rightarrow \infty$  und die Stetigkeit ist bewiesen.

**7.21** Eine Funktion  $f = (f_1, \dots, f_m): U \rightarrow \mathbb{R}^m$  auf einer Teilmenge  $U \subseteq \mathbb{R}^n$  ist genau dann stetig an einer Stelle  $x \in U$ , wenn jede der Komponenten  $f_1, \dots, f_m: U \rightarrow \mathbb{R}$  an der Stelle  $x$  stetig ist. Insbesondere ist  $f$  genau dann stetig, wenn jede der Komponenten stetig ist.

Ist nämlich  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $U$  mit  $x_k \rightarrow x$  für  $k \rightarrow \infty$ , so gilt nach 7.6

$$f(x_k) = (f_1(x_k), \dots, f_m(x_k)) \rightarrow (f_1(x), \dots, f_m(x)) = f(x)$$

genau dann für  $k \rightarrow \infty$ , wenn  $f_j(x_k) \rightarrow f_j(x)$  für alle  $j \in \{1, \dots, m\}$ . Ersteres bedeutet Stetigkeit von  $f$  an der Stelle  $x$ , letzteres Stetigkeit der Komponenten  $f_1, \dots, f_m$ .

Dies hat eine nützliche Anwendung.

**7.22** Es sei  $U \subseteq \mathbb{R}^n$  eine Teilmenge und  $x \in U$ . Jede der Funktionen  $f: U \rightarrow \mathbb{R}^m$ ,  $g: U \rightarrow \mathbb{R}^m$  und  $h: U \rightarrow \mathbb{R}$  sei stetig an der Stelle  $x$ . Dann sind auch die Funktionen

$$f + g: U \rightarrow \mathbb{R}^m, \quad y \mapsto f(y) + g(y)$$

und

$$hf: U \rightarrow \mathbb{R}^m, \quad y \mapsto h(y)f(y)$$

stetig an der Stelle  $x$ . Falls  $0 \notin h(U)$ , ist zudem

$$\frac{f}{h}: U \rightarrow \mathbb{R}^m, \quad y \mapsto \frac{1}{h(y)}f(y)$$

stetig an der Stelle  $x$ . Weiter ist jede konstante Funktion  $U \rightarrow \mathbb{R}^m$  stetig.

Bevor wir dies zeigen, geben wir eine Anwendung.

**Definition 7.23** Eine Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  heißt *affin-linear*, wenn sie von der Form  $f(x) = \phi(x) + c$  ist mit einer linearen Abbildung  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  und einer Konstanten  $c \in \mathbb{R}^m$ .

**Beispiel 7.24** Jede *affin-lineare* Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist stetig.

Es ist nämlich  $f = \phi + g$  mit einer linearen Abbildung  $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$  (die nach Beispiel 7.17 stetig ist) und einer konstanten Abbildung  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$  (die nach 7.22 stetig ist). Nach 7.22 ist  $f$  stetig.

**7.25** Beweis für 7.22. Nach 7.21 dürfen wir annehmen, dass  $m = 1$  ist, also alle Funktionen reellwertig sind. Sei  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $U$  mit  $x_k \rightarrow x$  für  $k \rightarrow \infty$ . Wegen Stetigkeit an der Stelle  $x$  sind dann  $(f(x_k))_{k \in \mathbb{N}}$ ,  $(g(x_k))_{k \in \mathbb{N}}$  und  $(h(x_k))_{k \in \mathbb{N}}$  konvergente Folgen in  $\mathbb{R}$  mit Grenzwert  $f(x)$ ,  $g(x)$  bzw.  $h(x)$ . Somit gilt nach den Rechenregeln für Grenzwerte von Folgen aus der Mathematik 1:

$$f(x_k) + g(x_k) \rightarrow f(x) + g(x) \quad \text{für } k \rightarrow \infty,$$

$h(x_k)f(x_k) \rightarrow h(x)f(x)$  und im Falle  $0 \notin h(U)$  zudem

$$\frac{1}{h(x_k)}f(x_k) \rightarrow \frac{1}{h(x)}f(x).$$

Also sind  $f + g$ ,  $hf$  und  $\frac{f}{h}$  stetig an der Stelle  $x$ . Ist  $f$  konstant mit Funktionswert  $c$ , so gilt  $f(x_k) = c \rightarrow c = f(x)$ , was Stetigkeit an der Stelle  $x$  zeigt.

Bevor wir inhaltlich fortfahren, schauen wir noch zwei Beispiele für affin-lineare Abbildungen an.

**Beispiel 7.26** Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto \frac{1}{2}x + 1$  ist affin-linear. Ihr Graph ist die Gerade  $y = \frac{1}{2}x + 1$ . Diese läuft durch den Punkt  $(0, 1)$  und enthält auch den Punkt  $(1, \frac{3}{2})$ , enthält also alle Punkte mit Ortsvektor

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} \quad \text{mit } t \in \mathbb{R}.$$

**Beispiel 7.27** Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto \frac{1}{2}x + y + 2$  ist affin-linear. Ihr Graph ist die Ebene  $z = \frac{1}{2}x + y + 2$ . Diese läuft durch den Punkt  $(0, 0, 2)$  und enthält auch die Punkte  $(1, 0, \frac{5}{2})$  und  $(0, 1, 3)$ , enthält also alle Punkte mit Ortsvektor

$$\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{mit } s, t \in \mathbb{R}.$$

**Definition 7.28** Sei  $n \in \mathbb{N}$ . Ein  $n$ -Tupel  $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$  von Zahlen  $\alpha_1, \dots, \alpha_n \in \mathbb{N}_0$  wird *Multiindex* genannt. Die Zahl

$$|\alpha| := \alpha_1 + \dots + \alpha_n \in \mathbb{N}_0$$

wird die *Länge* von  $\alpha$  genannt.<sup>9</sup> Die Funktion

$$\mathbb{R}^n \rightarrow \mathbb{R}, \quad x = (x_1, \dots, x_n) \mapsto x^\alpha := (x_1)^{\alpha_1} \cdots (x_n)^{\alpha_n}$$

ist das zugehörige *Monom*. Eine Funktion  $p: \mathbb{R}^n \rightarrow \mathbb{R}$  der Form

$$x \mapsto \sum_{|\alpha| \leq k} a_\alpha x^\alpha$$

mit  $k \in \mathbb{N}_0$  und reellen Koeffizienten  $a_\alpha$  wird *Polynom in  $n$  Variablen* genannt. Die Summation erfolgt dabei über alle Multiindizes  $\alpha \in (\mathbb{N}_0)^n$  mit Länge  $|\alpha| \leq k$ . Ist  $p$  nicht die konstante Nullfunktion, nennen wir das minimal mögliche  $n$  den *Grad* des Polynoms  $p$ . Der Grad des Nullpolynoms ist definiert als  $-\infty$ .

**Beispiel 7.29** Jedes Polynom  $p: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \mapsto \sum_{|\alpha| \leq k} a_\alpha x^\alpha$  ist stetig.

Nach 7.22 brauchen wir nur zu zeigen, dass jeder der Summanden  $\mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \mapsto a_\alpha x^\alpha$  stetig ist. Da die konstante Funktion  $\mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \mapsto a_\alpha$  stetig ist (siehe 7.22) und nach 7.16 jedes der Monome

$$x \mapsto (x_j)^{\alpha_j}$$

stetig ist für  $j \in \{1, \dots, n\}$ , ist

$$x \mapsto a_\alpha x^\alpha = a_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

nach 7.22 stetig, als Produkt von  $n + 1$  stetigen Funktionen.

Wir erwähnen weitere Multiindexnotation zur späteren Benutzung:

**Definition 7.30** Gegeben  $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$  sei

$$\alpha! := \alpha_1! \cdots \alpha_n!$$

das Produkt der Fakultäten. Gegeben  $\alpha, \beta \in (\mathbb{N}_0)^n$  schreiben wir  $\beta \leq \alpha$ , wenn  $\beta_j \leq \alpha_j$  für alle  $j \in \{1, \dots, n\}$ . Ist  $\beta \leq \alpha$ , so definieren wir den Binomialkoeffizienten

$$\binom{\alpha}{\beta} := \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}$$

als Produkt von gewöhnlichen Binomialkoeffizienten, wie in der Mathematik 1.

<sup>9</sup>Vorsicht: Dies ist nicht die (euklidische) Länge  $\sqrt{\alpha_1^2 + \cdots + \alpha_n^2}$  des Vektors mit den Komponenten  $\alpha_1, \dots, \alpha_n$ . Die Doppelung der Notation ist unschön, aber üblich.

**Beispiel 7.31** Es ist  $(3, \sqrt{2})^{(4,2)} = 3^4(\sqrt{2})^2 = 81 \cdot 2 = 162$ . Weiter ist

$$|(2, 4, 1)| = 2 + 4 + 1 = 7,$$

$(4, 2)! = 4!2! = 24 \cdot 2 = 48$  und

$$\binom{(6, 5)}{(2, 3)} := \binom{6}{2} \binom{5}{3} = \frac{6!}{2!(6-2)!} \frac{5!}{3!(5-3)!} = \frac{6 \cdot 5}{2} \cdot \frac{5 \cdot 4}{2} = 15 \cdot 10 = 150.$$

Multiindizes in  $(\mathbb{N}_0)^2$  mit Länge  $\leq 3$  sind

$$(3, 0), (2, 1), (1, 2), (0, 3) \quad (\text{mit Länge } 3),$$

$$(2, 0), (1, 1), (0, 2) \quad \text{mit Länge } 2$$

sowie  $(1, 0)$  und  $(0, 1)$  mit Länge 1 und  $(0, 0)$  mit Länge 0. Die Multiindizes  $\beta \in (\mathbb{N}_0)^2$  mit  $\beta \leq (1, 2)$  sind

$$(0, 0), (1, 0), (0, 1), (1, 1), (0, 2) \text{ und } (1, 2).$$

Für Monome in mehreren Variablen lässt sich der binomische Lehrsatz wie folgt verallgemeinern.

**7.32** Für alle  $x, y \in \mathbb{R}^n$  und  $\alpha \in (\mathbb{N}_0)^n$  gilt

$$(x + y)^\alpha = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^\beta y^{\alpha - \beta}. \quad (41)$$

Beweis: Für jedes  $j \in \{1, \dots, n\}$  ist  $(x_j + y_j)^{\alpha_j} = \sum_{\beta_j=0}^{\alpha_j} \binom{\alpha_j}{\beta_j} x_j^{\beta_j} y_j^{\alpha_j - \beta_j}$  nach dem binomischen Lehrsatz. Einsetzen für alle  $j$  und Ausmultiplizieren liefert

$$\begin{aligned} (x + y)^\alpha &= (x_1 + y_1)^{\alpha_1} \cdots (x_n + y_n)^{\alpha_n} \\ &= \sum_{\beta_1=0}^{\alpha_1} \cdots \sum_{\beta_n=0}^{\alpha_n} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} x_1^{\beta_1} \cdots x_n^{\beta_n} y_1^{\alpha_1 - \beta_1} \cdots y_n^{\alpha_n - \beta_n}. \end{aligned}$$

Umgeschrieben mit Multiindex-Notationen ist dies genau (41).

Wir erwähnen, dass sich Stetigkeit an einer Stelle  $x$  auch wie folgt umformulieren lässt.

**7.33** Sei  $U \subseteq \mathbb{R}^n$ . Eine Funktion  $f: U \rightarrow \mathbb{R}^m$  ist genau dann stetig an einer Stelle  $x \in U$ , wenn für jedes  $\varepsilon > 0$  ein  $\delta > 0$  existiert derart, dass

$$|f(y) - f(x)| < \varepsilon$$

für alle  $y \in U$  mit  $|y - x| < \delta$ .

Nachweis: Ist die Bedingung erfüllt und  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $U$  mit  $x_k \rightarrow x$  für  $k \rightarrow \infty$ , so gibt es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  derart, dass  $|f(y) - f(x)| < \varepsilon$  für alle  $y \in U$  mit  $|y - x| < \delta$ . Da  $x_k \rightarrow x$ , existiert ein  $N \in \mathbb{N}$  derart, dass  $|x - x_k| < \delta$  für alle  $k \geq N$  und somit

$$|f(x) - f(x_k)| < \varepsilon \quad \text{für alle } k \geq N.$$

Also gilt  $f(x_k) \rightarrow f(x)$  für  $k \rightarrow \infty$  und somit ist  $f$  stetig an der Stelle  $x$ .

Ist die genannte Bedingung verletzt, so existiert ein  $\varepsilon > 0$  derart, dass für jedes  $\delta > 0$  ein  $y \in U$  existiert mit  $|y - x| < \delta$  und  $|f(x) - f(y)| \geq \varepsilon$ . Für  $k \in \mathbb{N}$  wenden wir dies mit  $\delta = 1/k$  an und finden ein  $x_k := y \in U$  mit  $|x_k - x| < 1/k$  derart, dass  $|f(x) - f(x_k)| \geq \varepsilon$ . Dann gilt  $x_k \rightarrow x$  aber  $f(x_k)$  konvergiert nicht gegen  $f(x)$ ; also ist  $f$  nicht stetig an der Stelle  $x$ .

Auch mittels 7.33 hätte man 7.24 begründen können:

**Beispiel 7.34** Jede affin-lineare Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist stetig.

Neuer Beweis: Es ist  $f(x) = Ax + c$  mit einer Matrix  $A \in \mathbb{R}^{m \times n}$  und einem Vektor (einer Konstanten)  $c \in \mathbb{R}^m$ . Um Stetigkeit an einer Stelle  $x \in \mathbb{R}^n$  zu zeigen, sei  $\varepsilon > 0$ . Wir setzen  $\delta := \frac{\varepsilon}{\|A\|_{\text{op}} + 1}$ . Für alle  $y \in \mathbb{R}^n$  mit  $|y - x| < \delta$  ist dann

$$\begin{aligned} |f(y) - f(x)| &= |Ay + c - Ax - c| = |Ay - Ax| = |A(y - x)| \\ &\leq \|A\|_{\text{op}} |y - x| \leq \frac{\|A\|_{\text{op}}}{\|A\|_{\text{op}} + 1} \varepsilon < \varepsilon. \end{aligned}$$

Nach 7.33 ist  $f$  also stetig an der Stelle  $x$ . Da  $x$  beliebig war, ist  $f$  stetig.

In der Mathematik 1 hatten wir besonders schöne Resultate für Funktionen  $f: ]a, b[ \rightarrow \mathbb{R}$  auf einem offenen Intervall und Funktionen  $f: [a, b] \rightarrow \mathbb{R}$  auf einem beschränkten, abgeschlossenen Intervall.<sup>10</sup> Auch im Fall von Funktionen mehrerer reeller Variablen wird der Definitionsbereich  $U \subseteq \mathbb{R}^n$  einer Funktion  $f: U \rightarrow \mathbb{R}^m$  meist “schön” angenommen, z.B. offen oder abgeschlossen. Diese Begriffe wollen wir nun kennenlernen.

<sup>10</sup>Letztere werden wir bald “kompakte Intervalle” nennen.

**Definition 7.35** Eine Teilmenge  $U \subseteq \mathbb{R}^n$  heißt *offen*, wenn sie um jeden Punkt  $x \in U$  eine ganze Kugel enthält, also ein  $\varepsilon > 0$  existiert mit  $B_\varepsilon(x) \subseteq U$ .

Um jeden Punkt  $x \in U$  haben wir also “etwas Platz”.

**Beispiel 7.36** Für jedes  $y \in \mathbb{R}^n$  und  $r > 0$  ist die Kugel  $B_r(y)$  offen in  $\mathbb{R}^n$ .

Ist nämlich  $x \in B_r(y)$ , so ist  $|x - y| < r$ , also  $\varepsilon := r - |x - y| > 0$ . Wir zeigen, dass  $B_\varepsilon(x) \subseteq B_r(y)$  (womit die Offenheit von  $B_r(y)$  bewiesen ist). Für jedes  $z \in B_\varepsilon(x)$  ist

$$|z - y| = |z - x + x - y| \leq |z - x| + |x - y| < \varepsilon + |x - y| = r,$$

also  $z \in B_r(y)$ . In der Tat ist also  $B_\varepsilon(x) \subseteq B_r(y)$ .

**Beispiel 7.37** Die rechte Halbebene  $H := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0\}$  ist offen.

Ist nämlich  $x = (x_1, x_2) \in H$ , so ist  $\varepsilon := x_1 > 0$ . Ist  $y = (y_1, y_2) \in B_\varepsilon(x)$ , so ist  $|y_1 - x_1| \leq |y - x| < \varepsilon = x_1$ , also

$$y_1 = x_1 + (y_1 - x_1) \geq x_1 - |y_1 - x_1| > x_1 - x_1 = 0,$$

folglich  $y \in H$ . Also ist  $B_\varepsilon(x) \subseteq H$  und somit Offenheit von  $H$  bewiesen.

**Beispiel 7.38** Für  $y \in \mathbb{R}^n$  und  $r > 0$  ist  $\overline{B}_r(y)$  nicht offen.

Schreiben wir  $y = (y_1, \dots, y_n)$ , so ist mit dem  $n$ ten Standard-Basisvektor  $e_n$

$$y + re_n \in \overline{B}_r(y),$$

da  $|y + re_n - y| = |re_n| = r$ . Wäre  $\overline{B}_r(y)$  offen, so gäbe es ein  $\varepsilon > 0$  mit  $B_\varepsilon(y + re_n) \subseteq \overline{B}_r(y)$ . Dann wäre  $y + re_n + \frac{\varepsilon}{2}e_n \in B_\varepsilon(y + re_n) \subseteq \overline{B}_r(y)$ . Jedoch ist

$$|y + re_n + (\varepsilon/2)e_n - y| = |(\varepsilon/2 + r)e_n| = (\varepsilon/2) + r > r,$$

also  $y + re_n + (\varepsilon/2)e_n \notin \overline{B}_r(y)$ , Widerspruch. Die Annahme der Offenheit von  $\overline{B}_r(y)$  muss also falsch sein.

**Definition 7.39** Eine Teilmenge  $A \subseteq \mathbb{R}^n$  heißt *abgeschlossen*, wenn sie unter Grenzwerten abgeschlossen ist, d.h. für jede Folge  $(x_k)_{k \in \mathbb{N}}$  in  $A$ , die in  $\mathbb{R}^n$  konvergiert, ist

$$\lim_{k \rightarrow \infty} x_k \in A.$$

**Beispiel 7.40** Für alle  $y \in \mathbb{R}^n$  und  $r > 0$  ist die Kugel  $\overline{B}_r(y)$  abgeschlossen in  $\mathbb{R}^n$ .

Sei nämlich  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $\overline{B}_r(y)$ , die gegen ein  $z \in \mathbb{R}^n$  konvergiert. Wir zeigen, dass  $z \in \overline{B}_r(y)$  (was den Beweis beendet), also  $|z - y| \leq r$ , also  $|z - y|^2 \leq r^2$ . Die Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$x = (x_1, \dots, x_n) \mapsto (x_1 - y_1)^2 + \dots + (x_n - y_n)^2 = \sum_{j=1}^n x_j^2 - \sum_{j=1}^n 2x_j y_j + \sum_{j=1}^n y_j^2$$

ist ein Polynom und somit stetig. Also ist

$$f(z) = f\left(\lim_{n \rightarrow \infty} x_k\right) = \lim_{k \rightarrow \infty} \underbrace{f(x_k)}_{\leq r^2} \leq r^2$$

und somit (wie zu zeigen)  $|z - y|^2 \leq r^2$ . Hierbei wurde benutzt, dass  $f(x_k) = |x_k - y|^2 \leq r^2$  für alle  $k \in \mathbb{N}$  und Ungleichungen der Art  $\leq$  bei Grenzwerten bestehen bleiben.

Wir halten zwei nützliche Eigenschaften fest.

**7.41** Vereinigungen offener Mengen sind offen: Ist  $J$  eine Menge und  $V_j \subseteq \mathbb{R}^n$  eine offene Teilmenge für alle  $j \in J$ , so ist auch

$$V := \bigcup_{j \in J} V_j = \{x \in \mathbb{R}^n : (\exists j \in J) : x \in V_j\}$$

eine offene Teilmenge von  $\mathbb{R}^n$ .

Für  $x \in V$  existiert nämlich ein  $j \in J$  mit  $x \in V_j$ . Da  $V_j$  offen ist, gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subseteq V_j$ . Dann gilt auch  $B_\varepsilon(x) \subseteq V$  und somit ist  $V$  offen.

**7.42** Eine Teilmenge  $A \subseteq \mathbb{R}^n$  ist genau dann abgeschlossen in  $\mathbb{R}^n$ , wenn das Komplement  $\mathbb{R}^n \setminus A$  in  $\mathbb{R}^n$  offen ist.

Sei  $\mathbb{R}^n \setminus A$  offen und  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $A$ , die in  $\mathbb{R}^n$  gegen ein  $x \in \mathbb{R}^n$  konvergiert. Wäre  $x \notin A$ , so wäre  $x$  in der offenen Menge  $\mathbb{R}^n \setminus A$  und es gäbe also ein  $\varepsilon > 0$  derart, dass  $B_\varepsilon(x) \subseteq \mathbb{R}^n \setminus A$ . Da  $x_k \rightarrow x$ , gibt es ein  $N \in \mathbb{N}$  mit  $|x_k - x| < \varepsilon$  für alle  $k \geq N$ . Für alle  $k \geq N$  wäre also  $x_k \in B_\varepsilon(x)$  und somit  $x_k \in \mathbb{R}^n \setminus A$ , im Widerspruch zu  $x_k \in A$ . Also muss doch  $x \in A$  sein und  $A$  ist abgeschlossen.

Ist umgekehrt  $\mathbb{R}^n \setminus A$  nicht offen, so gibt es ein  $x \in \mathbb{R}^n \setminus A$  derart, dass  $B_\varepsilon(x)$  nicht in  $\mathbb{R}^n \setminus A$  enthalten ist für alle  $\varepsilon > 0$ , also  $B_\varepsilon(x) \cap A \neq \emptyset$ . Für  $k \in \mathbb{N}$  wenden wir dies mit  $\varepsilon = 1/k$  an und finden ein  $x_k \in B_{1/k}(x) \cap A$ . Dann gilt  $x_k \in A$  und  $x_k \rightarrow x$  für  $k \rightarrow \infty$ . Da  $x \notin A$ , ist  $A$  nicht abgeschlossen.

**Definition 7.43** Eine Menge  $K \subseteq \mathbb{R}^n$  heißt *kompakt*, wenn jede Folge  $(x_k)_{k \in \mathbb{N}}$  in  $K$  eine Teilfolge hat, die gegen ein  $x \in K$  konvergiert.

Eine Teilmenge  $M \subseteq \mathbb{R}^n$  heißt *beschränkt*, wenn

$$\sup\{|x|: x \in M\} < \infty,$$

also  $M$  in einer Kugel  $\overline{B}_r(0)$  enthalten ist für ein  $r > 0$ . Zum Beispiel ist jede Kugel  $\overline{B}_r(x)$  beschränkt, denn für alle  $y \in \overline{B}_r(x)$  ist

$$|y| = |x + y - x| \leq |x| + |y - x| \leq |x| + r,$$

also  $\overline{B}_r(x) \subseteq \overline{B}_{|x|+r}(0)$ .

**7.44** (Satz von Heine-Borel). *Eine Teilmenge  $K$  von  $\mathbb{R}^n$  ist genau dann kompakt, wenn sie abgeschlossen ist und beschränkt.*

**Beispiel 7.45** *Jede Kugel  $\overline{B}_r(x) \subseteq \mathbb{R}^n$  ist kompakt.*

Nach 7.40 ist  $\overline{B}_r(x)$  nämlich abgeschlossen und da  $\overline{B}_r(x)$  zudem beschränkt ist, ist  $\overline{B}_r(x)$  nach dem Satz von Heine-Borel kompakt.

**Beispiel 7.46** *Jeder Quader der Form*

$$Q := [a_1, b_1] \times \cdots \times [a_n, b_n]$$

mit  $a_j \leq b_j$  für  $j \in \{1, \dots, n\}$  ist kompakt.

Es gibt nämlich ein  $r > 0$  derart, dass  $[a_j, b_j] \subseteq [-r, r]$  für alle  $j \in \{1, \dots, n\}$ . Dann ist  $x_j^2 \leq r^2$  für alle  $x_j \in [a_j, b_j]$  und somit

$$|x|^2 = x_1^2 + \cdots + x_n^2 \leq n r^2$$

für alle  $x = (x_1, \dots, x_n) \in Q$ . Also ist  $Q$  beschränkt. Ist  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $Q$ , die gegen ein  $x \in \mathbb{R}^n$  konvergiert, so gilt für jedes  $j \in \{1, \dots, n\}$

$$\text{pr}_j(x) = \text{pr}_j\left(\lim_{k \rightarrow \infty} x_k\right) = \lim_{k \rightarrow \infty} \text{pr}_j(x_k) \in [a_j, b_j],$$

da die Ungleichung  $a_j \leq \text{pr}_j(x_k) \leq b_j$  beim Grenzübergang  $k \rightarrow \infty$  bestehen bleibt, also  $a_j \leq \text{pr}_j(x) \leq b_j$  gilt für den Grenzwert  $\text{pr}_j(x) = \lim_{k \rightarrow \infty} \text{pr}_j(x_k)$ .

Der Vollständigkeit finden Sie hier im Skript einen Beweis für den Satz von Heine-Borel, der in der Vorlesung jedoch übersprungen wird. Die Aussage lässt sich zurückführen auf den Satz von Bolzano-Weierstraß über die Existenz konvergenter Teilfolgen in Intervallen der Form  $[a, b]$  (der aus der Mathematik 1 bekannt ist).

**7.47** Beweis des Satzes von Heine-Borel. Ist  $K \subseteq \mathbb{R}^n$  kompakt, so nimmt die stetige Funktion

$$K \rightarrow \mathbb{R}, \quad x \mapsto |x| = \sqrt{x_1^2 + \cdots + x_n^2}$$

auf  $K$  ein Maximum  $r \geq 0$  an (siehe 7.48). Also ist  $K$  beschränkt. Ist  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $K$ , die gegen ein  $x \in \mathbb{R}^n$  konvergiert, so hat  $(x_k)_{k \in \mathbb{N}}$  wegen der Kompaktheit eine Teilfolge  $(x_{k_\ell})_{\ell \in \mathbb{N}}$ , die gegen ein  $y \in K$  konvergiert. Da sie auch gegen  $x$  konvergiert und Grenzwerte eindeutig sind, folgt  $x = y \in K$ . Also ist  $K$  eine abgeschlossene Teilmenge von  $\mathbb{R}^n$ .

Sei umgekehrt  $K$  abgeschlossen in  $\mathbb{R}^n$  und beschränkt. Es gibt dann also ein  $r > 0$  mit  $K \subseteq \overline{B}_r(0)$ . Für jedes  $x \in K$  und jedes  $j \in \{1, \dots, n\}$  ist dann

$$|x_j| \leq |x| \leq r,$$

also  $x_j \in [-r, r]$ . Somit ist  $K \subseteq [-r, r]^n =: W$ . Können wir zeigen, dass der Würfel  $W$  kompakt ist, dann ist auch  $K$  kompakt. Ist nämlich  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $K$ , so ist diese auch eine Folge in  $W$  und hat somit eine Teilfolge  $(x_{k_\ell})_{\ell \in \mathbb{N}}$ , die gegen ein  $x \in W$  konvergiert. Da  $K$  in  $\mathbb{R}^n$  abgeschlossen ist und  $x_{k_\ell} \in K$  für alle  $\ell$ , folgt  $x \in K$ .

Wir zeigen nun per Induktion nach  $n \in \mathbb{N}$ , dass jeder Würfel  $W = [-r, r]^n$  kompakt ist. Im Falle  $n = 1$  hat jede Folge in  $[-r, r]^1 = [-1, 1]$  eine konvergente Teilfolge nach dem Satz von Bolzano-Weierstraß (siehe Mathematik 1); also ist  $[-r, r]$  kompakt. Sei nun  $n \geq 2$  und  $[-r, r]^{n-1}$  kompakt. Ist  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $[-r, r]^n$ , so schreiben wir  $x_k = (x_{k,1}, \dots, x_{k,n})$  in Komponenten und setzen  $x'_k := (x_{k,1}, \dots, x_{k,n-1}) \in [-r, r]^{n-1}$ . Per Induktionsvoraussetzung hat  $(x'_k)_{k \in \mathbb{N}}$  eine gegen ein  $y' = (y_1, \dots, y_{n-1}) \in [-r, r]^{n-1}$  konvergente Teilfolge  $(x_{k_\ell})_{\ell \in \mathbb{N}}$ . Nach dem Induktionsanfang hat die Folge  $(x_{k_\ell, n})_{\ell \in \mathbb{N}}$  in  $[-r, r]$  eine gegen ein  $y_n \in [-r, r]$  konvergente Teilfolge  $(x_{k_{\ell_m}, n})_{m \in \mathbb{N}}$ . Dann ist  $(x_{k_{\ell_m}})_{m \in \mathbb{N}}$  eine Teilfolge von  $(x_k)_{k \in \mathbb{N}}$ , die gegen  $(y', y_n) \in [-r, r]^n$  konvergiert. Also ist  $[-r, r]^n$  kompakt.

Wir geben eine typische Anwendung von Kompaktheit.

**7.48** (Der Satz vom Maximum). *Jede stetige Funktion  $f: K \rightarrow \mathbb{R}$  auf einer kompakten, nicht leeren Teilmenge  $K \subseteq \mathbb{R}^n$  nimmt ein Maximum und ein Minimum an.*

Beweis. Sei  $s := \sup f(K) \in \mathbb{R} \cup \{\infty\}$ . Es gibt eine Folge  $(x_k)_{k \in \mathbb{N}}$  in  $K$  derart, dass

$$f(x_k) \rightarrow s$$

für  $k \rightarrow \infty$ . Da  $K$  kompakt ist, gibt es eine Teilfolge  $(x_{k_\ell})_{\ell \in \mathbb{N}}$  von  $(x_k)_{k \in \mathbb{N}}$ , die für  $\ell \rightarrow \infty$  gegen ein  $x \in K$  konvergiert. Da  $f$  stetig ist, folgt

$$f(x_{k_\ell}) \rightarrow f(x)$$

für  $\ell \rightarrow \infty$ . Es gilt aber auch

$$f(x_{k_\ell}) \rightarrow s,$$

da  $(f(x_{k_\ell}))_{\ell \in \mathbb{N}}$  eine Teilfolge von  $(f(x_k))_{k \in \mathbb{N}}$  ist. Also ist  $s = f(x)$ . Das Supremum wird also angenommen (insbesondere ist  $s = f(x) < \infty$ ).

Drei weitere Grundbegriffe sollen einmal erwähnt werden. Wir brauchen sie aber nur gelegentlich.

**Definition 7.49** Es sei  $M \subseteq \mathbb{R}^n$  eine Teilmenge.

- (a) Der *Rand von  $M$*  ist per Definition die Menge  $\partial M$  aller  $x \in \mathbb{R}^n$  derart, dass für jedes  $\varepsilon > 0$  die Kugel  $B_\varepsilon(x)$  sowohl ein Element von  $M$  als auch ein Element des Komplements  $\mathbb{R}^n \setminus M$  enthält.
- (b) Der *Abschluss von  $M$*  ist definiert als die Menge  $\overline{M}$  aller  $x \in \mathbb{R}^n$  derart, dass für jedes  $\varepsilon > 0$  die Kugel  $B_\varepsilon(x)$  ein Element von  $M$  enthält.
- (c) Das *Innere von  $M$*  ist per Definition die Menge  $M^0$  aller  $x \in \mathbb{R}^n$  derart, dass  $B_\varepsilon(x) \subseteq M$  für ein  $\varepsilon > 0$  (insbesondere ist also  $x \in M$ ).

**Beispiel 7.50** Wir betrachten die offene Kreisscheibe  $B_1(0)$  in  $\mathbb{R}^2$ . Dann ist  $\partial B_1(0) = \mathbb{S}_1$  der Einheitskreis, der Abschluss  $\overline{B_1(0)}$  gleich der abgeschlossenen Kreisscheibe  $\overline{B_1(0)}$  und das Innere  $(B_1(0))^0$  gleich  $B_1(0)$  (siehe Skizzen aus der Vorlesung).

**Bemerkung 7.51** (a) Ist  $x \in M^0$ , so ist  $B_{\varepsilon(x)}(x) \subseteq M$  für ein  $\varepsilon(x) > 0$ . Für jedes  $y \in B_{\varepsilon(x)}(x)$  enthält  $B_{\varepsilon(x)}(x)$  (und somit  $M$ ) nach Beispiel 7.36 eine ganze Kugel um  $y$ ; also ist auch  $y \in M^0$  und somit  $B_{\varepsilon(x)}(x) \subseteq M^0$ . Wir schließen, dass

$$M^0 = \bigcup_{x \in M^0} B_{\varepsilon(x)}(x).$$

Als Vereinigung offener Mengen ist  $M^0$  nach 7.41 offen. Es ist  $M^0$  die größte offene Teilmenge von  $\mathbb{R}^n$ , die in  $M$  enthalten ist (d.h. jede andere solche offene Menge ist in  $M^0$  enthalten).

(b) Das Komplement von  $\overline{M}$  ist die Menge aller  $x \in \mathbb{R}^n$  derart, dass ein  $\varepsilon > 0$  existiert mit  $B_\varepsilon(x) \cap M = \emptyset$ , also  $B_\varepsilon(x) \subseteq \mathbb{R}^n \setminus M$ . Also ist

$$\mathbb{R}^n \setminus \overline{M} = (\mathbb{R}^n \setminus M)^0$$

gleich dem Inneren von  $\mathbb{R}^n \setminus M$  und somit offen. Nach 7.42 ist  $\overline{M}$  somit abgeschlossen. Man kann zeigen, dass  $\overline{M}$  die kleinste abgeschlossene Menge ist, die  $M$  enthält (also in jeder anderen enthalten).

(c) Es ist

$$\partial M = \overline{M} \setminus M^0. \quad (42)$$

Wenn das stimmt, so ist  $\partial M = \overline{M} \cap (\mathbb{R}^n \setminus M^0)$  und folglich nach den de Morganschen Regeln  $\mathbb{R}^n \setminus \partial M = (\mathbb{R}^n \setminus \overline{M}) \cup M^0$ . Dies ist eine offene Menge, folglich  $\partial M$  eine abgeschlossene Teilmenge von  $\mathbb{R}^n$ .

Nachweis von (42): Gegeben  $x \in \mathbb{R}^n$  gilt  $x \in \partial M$ , wenn für jedes  $\varepsilon > 0$   $B_\varepsilon(x) \cap M \neq \emptyset$  ist und  $B_\varepsilon(x) \cap (\mathbb{R}^n \setminus M) \neq \emptyset$ . Ersteres ist zu  $x \in \overline{M}$  äquivalent, zweiteres zu  $x \notin M^0$ , also  $x \in \mathbb{R}^n \setminus M^0$ .

Wir erwähnen noch eine weitere Anwendung von Kompaktheit, die uns zum Beispiel beim Studium parameterabhängiger Integrale nutzen wird. Über 7.33 hinaus lässt sich zeigen:

**7.52** *Es sei  $K \subseteq \mathbb{R}^n$  eine kompakte Teilmenge und  $f: K \rightarrow \mathbb{R}^m$  eine stetige Funktion. Dann ist  $f$  **gleichmäßig stetig** im folgenden Sinn: Für jedes  $\varepsilon > 0$  gibt es ein  $\delta > 0$  derart, dass für alle  $x, y \in K$  gilt:*

$$|x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

Andernfalls würde nämlich der Verneinung der letzten Aussage gelten, also: Es gibt ein  $\varepsilon > 0$  derart, dass für jedes  $\delta > 0$  Elemente  $x, y \in K$  existieren derart, dass  $|x - y| < \delta$  aber  $|f(x) - f(y)| \geq \varepsilon$ . Für  $k \in \mathbb{N}$  wenden wir dies mit  $\delta = \frac{1}{k}$  an und finden  $x_k, y_k \in K$  derart, dass  $|x_k - y_k| < \frac{1}{k}$  aber

$$|f(x_k) - f(y_k)| \geq \varepsilon. \quad (43)$$

Wegen der Kompaktheit von  $K$  gibt es eine Teilfolge  $(x_{k_\ell})_{\ell \in \mathbb{N}}$  von  $(x_k)_{k \in \mathbb{N}}$ , die gegen ein  $x \in K$  konvergiert. Wegen der Stetigkeit von  $f$  an der Stelle  $x$  gibt es nach 7.33 ein  $\delta > 0$  derart, dass

$$|f(y) - f(x)| < \frac{\varepsilon}{2}$$

für alle  $y \in K$  mit  $|y - x| < \delta$ . Da  $x_{k_\ell} \rightarrow x$ , gibt es ein  $N \in \mathbb{N}$  derart, dass

$$|x - x_{k_\ell}| < \frac{\delta}{2} \quad \text{für alle } \ell \geq N. \quad (44)$$

Indem wir  $N$  notfalls vergrößern, dürfen wir annehmen, dass zudem

$$\frac{1}{N} \leq \frac{\delta}{2}.$$

Für alle  $\ell \geq N$  ist  $k_\ell \geq \ell \geq N$  und somit

$$|y_{k_\ell} - x_{k_\ell}| < \frac{1}{k_\ell} \leq \frac{1}{N} \leq \frac{\delta}{2}.$$

Mit dieser Ungleichung und (44) folgt

$$|x - y_{k_\ell}| = |x - x_{k_\ell} + x_{k_\ell} - y_{k_\ell}| \leq |x - x_{k_\ell}| + |x_{k_\ell} - y_{k_\ell}| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Somit ist

$$|f(y_{k_\ell}) - f(x_{k_\ell})| \leq |f(y_{k_\ell}) - f(x)| + |f(x) - f(x_{k_\ell})| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für alle  $\ell \geq N$ . Aber  $|f(y_{k_\ell}) - f(x_{k_\ell})| \geq \varepsilon$  nach (43), Widerspruch. Die Aussage muss also doch richtig sein.

Halten wir noch eine Fassung des Satzes von Pythagoras fest.

**7.53** Seien  $v_1, \dots, v_m$  Vektoren in  $\mathbb{R}^n$ , welche paarweise orthogonal sind, also  $v_j \cdot v_k = 0$  für alle  $j, k \in \{1, \dots, m\}$  mit  $j \neq k$ . Dann ist

$$|v_1 + \dots + v_m|^2 = |v_1|^2 + \dots + |v_m|^2.$$

Es gilt nämlich

$$|v_1 + \dots + v_m|^2 = (v_1 + \dots + v_m) \cdot (v_1 + \dots + v_m) = \sum_{j=1}^m \sum_{k=1}^m v_j \cdot v_k = \sum_{j=1}^m v_j \cdot v_j = \sum_{j=1}^m |v_j|^2.$$

## 8 Kurven, Bogenlänge, Kurvenintegrale

In diesem Kapitel betrachten wir parametrisierte Kurven in  $\mathbb{R}^n$  und ordnen ihnen eine Kurvenlänge zu. Anschließend diskutieren wir zwei Arten von Integralen längs Kurven, die z.B. in der Physik wichtig sind.

### Kurven und Wege

**8.1** Eine *stetige Kurve* (oder einfach: Kurve) in  $\mathbb{R}^n$  ist eine stetige Abbildung  $\gamma: I \rightarrow \mathbb{R}^n$  auf einem nicht entarteten Intervall  $I \subseteq \mathbb{R}$ . Dann ist also

$$\gamma = (\gamma_1, \dots, \gamma_n)$$

mit den stetigen Komponenten  $\gamma_1, \dots, \gamma_n: I \rightarrow \mathbb{R}$ . Ist jede der Komponenten  $\gamma_1, \dots, \gamma_n$  sogar stetig differenzierbar,<sup>11</sup> so nennen wir  $\gamma$  eine *stetig differenzierbare Kurve* (oder kurz:  $C^1$ -Kurve). Kurven mit Definitionsbereich  $I = [a, b]$  nennen wir auch *Wege*.

Ist  $0 \neq t \in \mathbb{R}$  und  $v \in \mathbb{R}^n$ , so schreiben wir  $\frac{v}{t} := \frac{1}{t}v$ .

**8.2** Ist  $\gamma: I \rightarrow \mathbb{R}^n$  eine  $C^1$ -Kurve, so existiert für jedes  $t \in I$  die Ableitung

$$\gamma'(t) = \lim_{s \rightarrow t} \frac{\gamma(s) - \gamma(t)}{s - t}$$

und stimmt mit der komponentenweisen Ableitung  $(\gamma'_1(t), \dots, \gamma'_n(t))$  überein, d.h. für jedes Folge von Zahlen  $s_k \in I$  mit  $s_k \neq t$  gilt

$$\lim_{k \rightarrow \infty} \frac{\gamma(s_k) - \gamma(t)}{s_k - t} = (\gamma'_1(t), \dots, \gamma'_n(t)).$$

Der vorige Differenzenquotient ist nämlich gleich

$$\left( \frac{\gamma_1(s_k) - \gamma_1(t)}{s_k - t}, \dots, \frac{\gamma_n(s_k) - \gamma_n(t)}{s_k - t} \right)$$

und konvergiert nach 7.6 gegen das  $n$ -Tupel der Grenzwerte der Komponenten. Nach 7.21 ist  $\gamma' = (\gamma'_1, \dots, \gamma'_n): I \rightarrow \mathbb{R}^n$  stetig. Umgekehrt hätten wir Existenz und Stetigkeit von  $\gamma'$  als Definition einer  $C^1$ -Kurve nehmen können.

---

<sup>11</sup>Eine Funktion  $f: I \rightarrow \mathbb{R}$  hatten wir in der Mathematik 1 *stetig differenzierbar* genannt (kurz:  $C^1$ ), wenn  $f$  differenzierbar (also auch stetig) ist und  $f': I \rightarrow \mathbb{R}$  stetig.

**Beispiel 8.3** Es ist

$$\mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto (\cos(t), 2 \sin(t))$$

ein  $C^1$ -Weg, der die Ellipse  $\{(x, y): x^2 + (y/2)^2 = 1\}$  mit den Halbachsen 1 und 2 durchläuft.

### Vektorwertige Integrale und Hauptsatz

**8.4** Es seien  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall,

$$\gamma = (\gamma_1, \dots, \gamma_n): I \rightarrow \mathbb{R}^n$$

eine stetige Kurve und  $a, b \in I$ . Wir definieren das Integral der Funktion  $\gamma$  von  $a$  bis  $b$  komponentenweise:

$$\int_a^b \gamma(t) dt := \left( \int_a^b \gamma_1(t) dt, \dots, \int_a^b \gamma_n(t) dt \right) \in \mathbb{R}^n.$$

**8.5** (Eigenschaften vektorwertiger Integrale). *Es sei  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $\gamma, \eta: I \rightarrow \mathbb{R}^n$  stetige Kurven. Dann gilt:*

- (a) *Für alle  $r, s \in \mathbb{R}$  ist  $r\gamma + s\eta: I \rightarrow \mathbb{R}^n$  eine stetige Kurve und für alle  $a, b \in I$  ist*

$$\int_a^b (r\gamma(t) + s\eta(t)) dt = r \int_a^b \gamma(t) dt + s \int_a^b \eta(t) dt.$$

- (b) (Intervalladditivität). *Für alle  $a, b, c \in I$  ist*

$$\int_a^b \gamma(t) dt = \int_a^c \gamma(t) dt + \int_c^b \gamma(t) dt. \quad (45)$$

- (c) (Integralabschätzungen). *Für  $a \leq b$  in  $I$  gilt*

$$\left| \int_a^b \gamma(t) dt \right| \leq \int_a^b |\gamma(t)| dt \leq \|\gamma\|_\infty (b - a) \quad (46)$$

*mit der "Supremumsnorm"  $\|\gamma\|_\infty := \sup\{|\gamma(t)|: t \in [a, b]\}$ .*

Nachweis: (a) und (b) sind uns für reellwertige Funktionen bekannt. Im vektorwertigen Fall können wir (a) und (b) komponentenweise nachprüfen.

Zum Nachweis von (c) machen wir uns zunächst klar, dass für jeden Vektor  $v \in \mathbb{R}^n$  mit  $v \neq 0$

$$|v| = \frac{v}{|v|} \cdot v \quad (47)$$

gilt, weil  $|v| = \frac{1}{|v|}|v|^2 = \frac{1}{|v|}(v \cdot v) = \frac{v}{|v|} \cdot v$ . Zum Beweis von (c) sei

$$v := \int_a^b \gamma(t) dt,$$

also  $v = (v_1, \dots, v_n)$  mit  $v_j = \int_a^b \gamma_j(t) dt$  für  $j \in \{1, \dots, n\}$ . Ist  $v = 0$ , sind die Abschätzungen in (c) trivial. Ist  $v \neq 0$ , erhalten wir mit  $\frac{v}{|v|} = \left(\frac{v_1}{|v|}, \dots, \frac{v_n}{|v|}\right)$

$$\begin{aligned} |v| &= \frac{v}{|v|} \cdot v = \sum_{j=1}^n \frac{v_j}{|v|} \int_a^b \gamma_j(t) dt \\ &= \int_a^b \sum_{j=1}^n \frac{v_j}{|v|} \gamma_j(t) dt = \int_a^b \frac{v}{|v|} \cdot \gamma(t) dt \\ &\leq \int_a^b \left| \frac{v}{|v|} \cdot \gamma(t) \right| dt \\ &\leq \int_a^b \underbrace{\left| \frac{v}{|v|} \right|}_{=1} |\gamma(t)| dt = \int_a^b |\gamma(t)| dt, \end{aligned}$$

wobei beim Übergang zur letzten Zeile die Cauchy-Schwarzsche Ungleichung benutzt wurde. Da  $|\gamma(t)| \leq \|\gamma\|_\infty$  für alle  $t \in [a, b]$ , können wir das letzte Integral (wegen der Monotonie des Riemann-Integrals) durch  $\int_a^b \|\gamma\|_\infty dt = \|\gamma\|_\infty(b-a)$  weiter nach oben abschätzen.

Die folgende Fassung des Hauptsatzes der Differential- und Integralrechnung kann man komponentenweise nachrechnen.

**8.6** Für jede  $C^1$ -Kurve  $\gamma: I \rightarrow \mathbb{R}^n$  und  $a, b \in I$  gilt

$$\int_a^b \gamma'(t) dt = \gamma(b) - \gamma(a).$$

## Weglänge

**8.7** Wir definieren die *Weglänge* eines  $C^1$ -Weges  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  als

$$L(\gamma) := \int_a^b |\gamma'(t)| dt. \quad (48)$$

Die Weglänge wird auch *Bogenlänge* genannt oder *Kurvenlänge*. Im entarteten Fall eines Intervalls  $[a, a] = \{a\}$  setzen wir  $L(\gamma) := 0$  für Funktionen  $\gamma: [a, a] \rightarrow \mathbb{R}^n$ .

**Beispiel 8.8** Für  $\alpha \in ]0, 2\pi[$  durchläuft der Weg

$$\gamma: [0, \alpha] \rightarrow \mathbb{R}^2, \quad t \mapsto (\cos(t), \sin(t))$$

einmal im Gegenuhrzeigersinn den Kreisbogen des Einheitskreises von  $(1, 0)$  aus bis  $(\cos(\alpha), \sin(\alpha))$ . Da  $\gamma'(t) = (-\sin(t), \cos(t))$  für alle  $t \in [0, \alpha]$  mit

$$|\gamma'(t)| = \sqrt{(-\sin(t))^2 + (\cos(t))^2} = 1,$$

ist die Weglänge des Kreisbogens

$$L(\gamma) = \int_0^\alpha |\gamma'(t)| dt = \int_0^\alpha 1 dt = \alpha.$$

**Beispiel 8.9** Gegeben  $v, w \in \mathbb{R}^n$  parametrisieren wir die Verbindungsstrecke von  $v = (v_1, \dots, v_n)$  und  $w = (w_1, \dots, w_n)$  durch den Weg

$$\gamma: [0, 1] \rightarrow \mathbb{R}^n, \quad t \mapsto v + t(w - v) = (v_1 + t(w_1 - v_1), \dots, v_n + t(w_n - v_n)).$$

Dann ist  $\gamma'(t) = (w_1 - v_1, \dots, w_n - v_n) = w - v$  für alle  $t \in [0, 1]$  und somit

$$L(\gamma) = \int_0^1 |\gamma'(t)| dt = \int_0^1 |w - v| dt = |w - v|.$$

Die Länge der Verbindungsstrecke stimmt also mit dem Abstand der Punkte  $v$  und  $w$  überein.

**8.10** Machen wir uns noch plausibel, warum sich das Integral  $\int_a^b |\gamma'(t)| dt$  wirklich anschaulich als Länge des Weges  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  interpretieren lässt.

Wir unterteilen das Intervall in Teilintervalle mit Zerlegungspunkten

$$a = t_0 < t_1 < \cdots < t_m = b.$$

Sind die Abstände  $t_j - t_{j-1}$  sehr klein, können wir  $\gamma$  auf  $[t_{j-1}, t_j]$  kaum von der Verbindungsstrecke von  $\gamma(t_{j-1})$  und  $\gamma(t_j)$  unterscheiden. Die Weglänge von  $\gamma|_{[t_{j-1}, t_j]}$  sollte also etwa

$$|\gamma(t_j) - \gamma(t_{j-1})|$$

sein. Hierbei ist nach dem Hauptsatz

$$\gamma(t_j) - \gamma(t_{j-1}) = \int_{t_{j-1}}^{t_j} \gamma'(t) dt$$

und dies ist näherungsweise  $\int_{t_{j-1}}^{t_j} \gamma'(t_{j-1}) dt = (t_j - t_{j-1})\gamma'(t_{j-1})$ , da  $\gamma'$  stetig ist und sich  $\gamma'(t)$  für  $t \in [t_{j-1}, t_j]$  kaum von  $\gamma'(t_{j-1})$  unterscheidet. Also ist näherungsweise

$$L(\gamma|_{[t_{j-1}, t_j]}) \approx (t_j - t_{j-1})|\gamma'(t_{j-1})|$$

und somit näherungsweise

$$L(\gamma) = \sum_{j=1}^m L(\gamma|_{[t_{j-1}, t_j]}) \approx \sum_{j=1}^m (t_j - t_{j-1})|\gamma'(t_{j-1})|. \quad (49)$$

Der gemachte Fehler sollte umso kleiner sein, je feiner die gewählte Zerlegung ist. Nun steht aber auf der rechten Seite von (49) eine Riemannsche Summe, wie wir sie zur Definition von Riemann-Integralen benutzt haben. Betrachten wir eine Folge von Zerlegungen, deren Maschenweite gegen 0 geht, konvergieren die Riemannschen Summen gegen das Riemann-Integral  $\int_a^b |\gamma'(t)| dt$ . Der Grenzübergang in (49) führt also auf (48).

## Umparametrisieren von Wegen

Beim Umparametrisieren ändert sich die Länge von Wegen nicht:

**8.11** *Ist  $\phi: [c, d] \rightarrow [a, b]$  ein  $C^1$ -Diffeomorphismus (also bijektiv mit  $\phi$  und  $\phi^{-1}$  stetig differenzierbar), so ist*

$$L(\gamma) = L(\gamma \circ \phi)$$

für jeden  $C^1$ -Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$ .

Schreiben wir  $\gamma = (\gamma_1, \dots, \gamma_n)$ , so hat nämlich  $\gamma \circ \phi = (\gamma_1 \circ \phi, \dots, \gamma_n \circ \phi)$  die komponentenweise gebildete Ableitung

$$(\gamma \circ \phi)'(t) = (\gamma_1'(\phi(t))\phi'(t), \dots, \gamma_n'(\phi(t))\phi'(t)) = \phi'(t)\gamma'(\phi(t)) \quad (50)$$

unter Benutzung der Kettenregel der Mathematik 1 für Funktionen einer Variablen. Also ist

$$|(\gamma \circ \phi)'(t)| = |\phi'(t)| |\gamma'(\phi(t))|.$$

Erster Fall: Ist  $\phi$  orientierungserhaltend, also streng monoton wachsend, so ist  $\phi'(t) > 0$  für alle  $t \in [c, d]$ . Wir können die Betragsstriche weglassen und erhalten

$$L(\gamma \circ \phi) = \int_c^d |(\gamma \circ \phi)'(t)| dt = \int_c^d |\gamma'(\phi(t))| \phi'(t) dt = \int_a^b |\gamma'(s)| ds = L(\gamma)$$

mit der Substitution  $s = \phi(t)$ ,  $ds = \phi'(t) dt$ ,  $\phi(c) = a$ ,  $\phi(d) = b$ . Ist  $\phi$  orientierungsumkehrend, so ist  $\phi$  streng monoton fallend und  $\phi'(t) < 0$  für alle  $t \in [c, d]$ , also  $|\phi'(t)| = -\phi'(t)$ . Die Substitution  $s = \phi(t)$  mit  $\phi(c) = b$ ,  $\phi(d) = a$  (!) ergibt

$$\begin{aligned} L(\gamma \circ \phi) &= \int_c^d |(\gamma \circ \phi)'(t)| dt = - \int_c^d |\gamma'(\phi(t))| \phi'(t) dt \\ &= - \int_b^a |\gamma'(s)| ds = \int_a^b |\gamma'(s)| ds = L(\gamma). \end{aligned}$$

**8.12** Für einen  $C^1$ -Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  sind folgende Eigenschaften äquivalent:

- (a)  $|\gamma'(t)| = 1$  für alle  $t \in [a, b]$ ;
- (b) Für alle  $c \in [a, b]$  ist  $L(\gamma|_{[a,c]}) = c - a$ .

Wir nennen dann den Weg  $\gamma$  auf Bogenlänge parametrisiert.

Gilt nämlich (a), so ist für alle  $c \in [a, b]$

$$L(\gamma|_{[a,c]}) = \int_a^c |\gamma'(t)| dt = \int_a^c 1 dt = c - a,$$

wie in (b) verlangt.

Gilt (b), so betrachten wir die Funktion

$$f: [a, b] \rightarrow \mathbb{R}, \quad t \mapsto \int_a^t |\gamma'(s)| ds.$$

Da der Integrand stetig ist, ist  $f$  nach dem Hauptsatz der Differential- und Integralrechnung (siehe Mathematik 1) eine stetig differenzierbare Funktion mit Ableitung

$$f'(t) = |\gamma'(t)|$$

für alle  $t \in [a, b]$ . Per Voraussetzung (b) ist nun aber

$$f(t) = t - a$$

für alle  $t \in [a, b]$  und somit  $f'(t) = 1$ . Also ist

$$|\gamma'(t)| = f'(t) = 1$$

für alle  $t \in [a, b]$ , d.h. (a) gilt.

**8.13** Wir nennen eine  $C^1$ -Kurve  $\gamma: I \rightarrow \mathbb{R}^n$  *regulär*, wenn  $\gamma'(t) \neq 0$  für alle  $t \in I$ .

Jeder reguläre  $C^1$ -Weg lässt sich auf Bogenlänge umparametrisieren:

**8.14** Für jeden regulären  $C^1$ -Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  existiert ein  $C^1$ -Diffeomorphismus  $\phi: [0, L(\gamma)] \rightarrow [a, b]$ , so dass  $\gamma \circ \phi$  auf Bogenlänge parametrisiert ist.

Die Funktion  $f: [a, b] \rightarrow [0, L(\gamma)]$ ,  $t \mapsto \int_a^t |\gamma'(s)| ds$  ist stetig differenzierbar mit

$$f'(t) = |\gamma'(t)| > 0$$

für alle  $t \in [a, b]$ . Also ist  $f$  ein  $C^1$ -Diffeomorphismus und

$$\phi := f^{-1}: [0, L(\gamma)] \rightarrow [a, b]$$

erfüllt

$$\phi'(t) = \frac{1}{f'(\phi(t))} = \frac{1}{|\gamma'(\phi(t))|}.$$

für alle  $t \in [a, b]$ . Da

$$|(\gamma \circ \phi)'(t)| = |\gamma'(\phi(t))\phi'(t)| = |\gamma'(\phi(t))| |\phi'(t)| = |\gamma'(\phi(t))| \frac{1}{|\gamma'(\phi(t))|} = 1,$$

ist  $\gamma \circ \phi$  nach 8.12 (a) auf Bogenlänge parametrisiert.

**Beispiel 8.15** Der Weg  $\gamma: [0, \alpha] \rightarrow \mathbb{R}^2$ ,  $t \mapsto (\cos(t), \sin(t))$  aus Beispiel 8.8 ist auf Bogenlänge parametrisiert, denn wir haben dort gesehen, dass  $|\gamma'(t)| = 1$  für alle  $t \in [0, \alpha]$ .

**8.16** Wir nennen eine Kurve  $\gamma: I \rightarrow \mathbb{R}^n$  eine  $C^2$ -Kurve, wenn  $\gamma$  eine  $C^1$ -Kurve ist und die Ableitung  $\gamma': I \rightarrow \mathbb{R}^n$  eine  $C^1$ -Kurve ist. Dann ist also  $\gamma'' := (\gamma')': I \rightarrow \mathbb{R}^n$  eine stetige Kurve.

Rekursiv nennen wir eine  $C^1$ -Kurve  $\gamma$  eine  $C^{k+1}$ -Kurve, wenn  $\gamma'$  eine  $C^k$ -Kurve ist.

**8.17** Ist ein  $C^2$ -Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  auf Bogenlänge parametrisiert, so ist

$$\gamma'(t) \cdot \gamma''(t) = 0 \quad \text{für alle } t \in [a, b],$$

die erste Ableitung  $\gamma'(t) \in \mathbb{R}^n$  und zweite Ableitung  $\gamma''(t) \in \mathbb{R}^n$  stehen also aufeinander senkrecht.

Schreiben wir  $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ , so ist nämlich

$$1 = |\gamma'(t)|^2 = \gamma'(t) \cdot \gamma'(t) = \sum_{j=1}^n \gamma'_j(t) \gamma'_j(t) = \sum_{j=1}^n (\gamma'_j(t))^2$$

für alle  $t \in [a, b]$ . Dies ist eine reellwertige Funktion einer reellen Variablen. Ableiten mit der Produktregel der Mathematik 1 liefert

$$0 = \sum_{j=1}^n 2 \gamma'_j(t) \gamma''_j(t) = 2 \gamma'(t) \cdot \gamma''(t).$$

## Wegintegrale

Wir stellen nun zwei wichtige Arten von Wegintegralen vor, die auch in physikalischen Anwendungen häufig gebraucht werden.

**8.18** (Wegintegrale erster Art). Ist  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  ein  $C^1$ -Weg und

$$f: \gamma([a, b]) \rightarrow \mathbb{R}$$

stetig, so definieren wir<sup>12</sup>

$$\int_{\gamma} f := \int_{\gamma} f ds := \int_a^b f(\gamma(t)) |\gamma'(t)| dt.$$

---

<sup>12</sup>Wir benutzen die selbe Notation auch, wenn  $f$  auf einer größeren Teilmenge von  $\mathbb{R}^n$  (z.B. auf ganz  $\mathbb{R}^n$ ) definiert ist.

### Beispiel 8.19 Die Kurve

$$\gamma = (\gamma_1, \gamma_2, \gamma_3): [a, b] \rightarrow \mathbb{R}^3$$

beschreibe einen Draht im Raum, welcher an der Stelle  $\gamma(t) = x$  die Massendichte  $\rho(x)$  besitzt (im Sinne von Masse pro Weglänge).<sup>13</sup> Dann ist

$$m = \int_{\gamma} \rho ds$$

die Masse des Drahts. Sein Schwerpunkt ist

$$\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$$

mit

$$\bar{x}_k = \frac{1}{m} \int_{\gamma} \rho \operatorname{pr}_k ds = \frac{1}{m} \int_a^b \rho(\gamma(t)) \gamma_k(t) |\gamma'(t)| dt$$

für  $k \in \{1, 2, 3\}$ . Ist  $\rho$  konstant, so ist  $m = \rho \cdot L(\gamma)$  und

$$\bar{x}_k = \frac{1}{L(\gamma)} \int_a^b \gamma_k(t) |\gamma'(t)| dt.$$

**8.20** Wie im Beweis von 8.11 sieht man, dass Wegintegrale 1. Art invariant unter Umparametrisierungen sind:

Ist  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  ein  $C^1$ -Weg,  $f: \gamma([a, b]) \rightarrow \mathbb{R}$  stetig (wie in 8.18) und  $\phi: [c, d] \rightarrow [a, b]$  ein  $C^1$ -Diffeomorphismus, so ist

$$\int_{\gamma \circ \phi} f = \int_{\gamma} f.$$

**8.21** (Wegintegrale zweiter Art). Ist  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  ein  $C^1$ -Weg und

$$F: \gamma([a, b]) \rightarrow \mathbb{R}^n$$

stetig,<sup>14</sup> so definieren wir

$$\int_{\gamma} \langle F, d\vec{s} \rangle := \int_{\gamma} F \cdot d\vec{s} := \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt$$

---

<sup>13</sup>Ist der Draht an verschiedenen Stellen unterschiedlich gezogen worden oder verrostet, braucht  $\rho$  nicht konstant sein.

<sup>14</sup>Wir benutzen die selbe Notation auch, wenn  $F$  auf einer größeren Teilmenge von  $\mathbb{R}^n$  (z.B. auf ganz  $\mathbb{R}^n$ ) definiert ist.

unter Benutzung des Skalarprodukts

$$\langle x, y \rangle := x \cdot y := \sum_{j=1}^n x_j y_j$$

für  $x = (x_1, \dots, x_n)$  und  $y = (y_1, \dots, y_n)$  in  $\mathbb{R}^n$ .

**Beispiel 8.22** Es sei  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein stetiges Kraftfeld, also  $F(x) \in \mathbb{R}^3$  ein Kraftvektor, der an der Stelle  $x \in \mathbb{R}^3$  auf ein Teilchen wirkt. Sei  $\gamma(t) \in \mathbb{R}^3$  die Position des Teilchens zur Zeit  $t$ . Die bei der Bewegung des Teilchens längs einer  $C^1$ -Kurve  $\gamma: [a, b] \rightarrow \mathbb{R}^3$  aufgewandte Energie (geleistete Arbeit) ist dann

$$E = - \int_{\gamma} F \cdot d\vec{s}.$$

**8.23** Sei  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  ein  $C^1$ -Weg und  $F: \gamma([a, b]) \rightarrow \mathbb{R}^n$  stetig. Weiter sei  $\phi: [c, d] \rightarrow [a, b]$  ein  $C^1$ -Diffeomorphismus. Ist  $\phi$  orientierungserhaltend (also streng monoton wachsend), so ist

$$\int_{\gamma} F \cdot d\vec{s} = \int_{\gamma \circ \phi} F \cdot d\vec{s}.$$

Ist  $\phi$  orientierungsumkehrend (also streng monoton fallend), so ist

$$\int_{\gamma} F \cdot d\vec{s} = - \int_{\gamma \circ \phi} F \cdot d\vec{s}.$$

Nachweis: Ist  $\phi$  streng monoton wachsend, so ist überall  $\phi'(t) > 0$ . Weiter ist  $\phi(c) = a$ ,  $\phi(d) = b$  und die Substitution  $u = \phi(t)$  liefert

$$\begin{aligned} \int_{\gamma \circ \phi} \langle F, d\vec{s} \rangle &= \int_c^d \langle F(\gamma(\phi(t))), \gamma'(\phi(t))\phi'(t) \rangle dt \\ &= \int_c^d \langle F(\gamma(\phi(t))), \gamma'(\phi(t)) \rangle \phi'(t) dt \\ &= \int_a^b \langle F(\gamma(u)), \gamma'(u) \rangle du \\ &= \int_{\gamma} \langle F, d\vec{s} \rangle. \end{aligned}$$

Ist  $\gamma$  streng monoton fallend, so ist überall  $\gamma'(t) < 0$ , weiter  $\gamma(c) = b$  und  $\gamma(d) = a$ . Wir erhalten

$$\begin{aligned} \int_{\gamma \circ \phi} \langle F, d\vec{s} \rangle &= \int_c^d \langle F(\gamma(\phi(t))), \gamma'(\phi(t))\phi'(t) \rangle dt \\ &= \int_c^d \langle F(\gamma(\phi(t))), \gamma'(\phi(t)) \rangle \phi'(t) dt \\ &= \int_b^a \langle F(\gamma(u)), \gamma'(u) \rangle du = - \int_a^b \langle F(\gamma(u)), \gamma'(u) \rangle du \\ &= - \int_{\gamma} \langle F, d\vec{s} \rangle. \end{aligned}$$

**8.24** Ein Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  heißt *stückweise stetig differenzierbar* (kurz: *stückweise  $C^1$* ), wenn es einer Zerlegung

$$a = t_0 < t_1 < \dots < t_m = b$$

des Intervalls  $[a, b]$  derart gibt, dass für alle  $j \in \{1, \dots, m\}$  die Einschränkung

$$\gamma|_{[t_{j-1}, t_j]}: [t_{j-1}, t_j] \rightarrow \mathbb{R}^n$$

ein  $C^1$ -Weg ist.

**Beispiel 8.25** Zum Beispiel ist

$$\gamma: [0, 2 + \pi] \rightarrow \mathbb{R}^2, \quad t \mapsto \begin{cases} (t - 1, 0) & \text{wenn } t \in [0, 2] \\ (\cos(t - 2), \sin(t - 2)) & \text{wenn } t \in [2, 2 + \pi] \end{cases}$$

ein stückweiser  $C^1$ -Weg zur Zerlegung  $t_0 = 0, t_1 = 2, t_2 = 2 + \pi$ . Dieser läuft geradlinig von  $(-1, 0)$  nach  $(1, 0)$ , dann durchläuft er den oberen Halbkreis im Gegenuhrzeigersinn von  $(1, 0)$  nach  $(-1, 0)$ .

**Beispiel 8.26** Ein (parametrisierter) *Polygonzug* ist eine stetige Abbildung  $\gamma: [a, b] \rightarrow \mathbb{R}^n$ , die für eine Zerlegung

$$a = t_0 < t_1 < \dots < t_m = b$$

des Intervalls auf jedem der Intervalle  $[t_{j-1}, t_j]$  affin-linear ist. Für alle  $j \in \{1, \dots, m\}$  und  $t \in [t_{j-1}, t_j]$  ist also

$$\gamma(t) = \gamma(t_{j-1}) + \frac{t - t_{j-1}}{t_j - t_{j-1}}(\gamma(t_j) - \gamma(t_{j-1})).$$

Jeder Polygonzug ist ein stückweise stetig differenzierbarer Weg.

**8.27** Sind ein stückweise stetig differenzierbarer Weg  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  und die Zerlegung  $a = t_0 < \dots < t_m = b$  wie in 8.24, so definieren wir die *Bogenlänge* von  $\gamma$  als Summe der Bogenlängen der beteiligten stetig differenzierbaren Teilwege:

$$L(\gamma) := \sum_{j=1}^m L(\gamma|_{[t_{j-1}, t_j]}).$$

Ist  $f: \gamma([a, b]) \rightarrow \mathbb{R}$  stetig, so definieren wir

$$\int_{\gamma} f \, ds := \sum_{j=1}^m \int_{\gamma|_{[t_{j-1}, t_j]}} f \, ds = \sum_{j=1}^m \int_{t_{j-1}}^{t_j} f(\gamma(t)) |\gamma'(t)| \, dt.$$

Ist  $F: \gamma([a, b]) \rightarrow \mathbb{R}^n$  stetig, so definieren wir

$$\int_{\gamma} F \cdot d\vec{s} := \sum_{j=1}^m \int_{\gamma|_{[t_{j-1}, t_j]}} F \cdot d\vec{s} = \sum_{j=1}^m \int_{t_{j-1}}^{t_j} F(\gamma(t)) \cdot \gamma'(t) \, dt.$$

## 9 Differentialrechnung für reellwertige Funktionen in mehreren Variablen

Wir betrachten zunächst reellwertige Funktionen  $f: U \rightarrow \mathbb{R}$  auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ . Im folgenden Kapitel wird dann auch Differentialrechnung für vektorwertige Funktionen  $f: U \rightarrow \mathbb{R}^m$  entwickelt.

**9.1** Die Funktion  $f: U \rightarrow \mathbb{R}$  heißt *partiell differenzierbar* an einer Stelle  $x = (x_1, \dots, x_n) \in U$ , wenn die folgenden *partiellen Ableitungen* existieren:

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x) &:= \lim_{t \rightarrow 0} \frac{f(x_1 + t, x_2, \dots, x_n) - f(x_1, \dots, x_n)}{t} \\ \frac{\partial f}{\partial x_2}(x) &:= \lim_{t \rightarrow 0} \frac{f(x_1, x_2 + t, \dots, x_n) - f(x_1, \dots, x_n)}{t} \\ &\vdots \\ \frac{\partial f}{\partial x_n}(x) &:= \lim_{t \rightarrow 0} \frac{f(x_1, x_2, \dots, x_n + t) - f(x_1, \dots, x_n)}{t}. \end{aligned}$$

Mit den Standard-Basisvektoren  $e_1, \dots, e_n$  ist also

$$\frac{\partial f}{\partial x_j}(x) = \lim_{t \rightarrow 0} \frac{f(x + te_j) - f(x)}{t}$$

für  $j \in \{1, \dots, n\}$ .

**Beispiel 9.2** Wir betrachten  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x_1, x_2) \mapsto x_1 + 2x_2$ . Dann ist

$$\frac{\partial f}{\partial x_1}(x) = \lim_{t \rightarrow 0} \frac{x_1 + t + 2x_2 - (x_1 + 2x_2)}{t} = 1$$

und

$$\frac{\partial f}{\partial x_2}(x) = \lim_{t \rightarrow 0} \frac{x_1 + 2(x_2 + 2) - (x_1 + 2x_2)}{t} = 2.$$

Effizienter sieht man dies wie folgt: Die partielle Ableitung  $\frac{\partial f}{\partial x_1}(x)$  erhält man, in dem man die zweite Variable  $x_2$  festhält und nach der ersten Variablen ableitet wie in Mathematik 1. Wir betrachten also die Funktion

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x_1 \mapsto f(x_1, x_2) = x_1 + 2x_2$$

der einen reellen Variablen  $x_1$  und betrachten  $x_2$  als Konstante. Ableiten liefert 1.

Ebenso erhalten Sie  $\frac{\partial f}{\partial x_2}(x)$ , indem Sie die erste Variable festhalten (als Konstante betrachten) und nach der zweiten Variablen ableiten.

**Beispiel 9.3** Wir betrachten  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $x = (x_1, x_2) \mapsto \sin(x_1) + (x_2)^2$ . Es ist

$$\frac{\partial f}{\partial x_1}(x) = \cos(x_1), \quad \frac{\partial f}{\partial x_2}(x) = 2x_2.$$

**Bemerkung 9.4** Weitere Notationen für  $\frac{\partial f}{\partial x_j}(x)$  sind u.a.

$$\frac{\partial}{\partial x_j} f(x), \quad D_j f(x), \quad f_{x_j}(x).$$

**9.5** Sei  $f: U \rightarrow \mathbb{R}$  eine Funktion auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ . Existieren die partiellen Ableitungen  $\frac{\partial f}{\partial x_j}(x)$  für alle  $j \in \{1, \dots, n\}$  und  $x \in U$ , so nennen wir  $f$  *partiell differenzierbar*. Ist  $f$  partiell differenzierbar und sind die Funktionen

$$\frac{\partial f}{\partial x_j}: U \rightarrow \mathbb{R}, \quad x \mapsto \frac{\partial f}{\partial x_j}(x)$$

stetig, so nennt man  $f$  *stetig differenzierbar* (oder kurz: eine  $C^1$ -Funktion).

Wir werden bald sehen (siehe 9.12), dass jede  $C^1$ -Funktion insbesondere stetig ist (also eine  $C^0$ -Funktion).

**Beispiel 9.6** Die Funktion  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $x = (x_1, x_2, x_3) \mapsto x_1 e^{x_2} + 2x_3$  ist stetig differenzierbar. Die partiellen Ableitungen

$$\frac{\partial f}{\partial x_1}(x) = e^{x_2}, \quad \frac{\partial f}{\partial x_2}(x) = x_1 e^{x_2} \quad \text{und} \quad \frac{\partial f}{\partial x_3}(x) = 2$$

existieren nämlich und sind stetige Funktionen von  $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ .

**9.7** Es sei  $U \subseteq \mathbb{R}^n$  eine Teilmenge und  $f: U \rightarrow \mathbb{R}$  eine Funktion. Wir nennen einen Punkt  $x \in U$  eine *lokale Minimalstelle* von  $f$ , wenn es ein  $\varepsilon > 0$  gibt derart, dass

$$f(x) \leq f(y)$$

für alle  $y \in U$  mit  $|y - x| < \varepsilon$ . Gibt es ein  $\varepsilon > 0$  derart, dass

$$f(x) \geq f(y)$$

für alle  $y \in U$  mit  $|y - x| < \varepsilon$ , so nennt man  $x$  eine *lokale Maximalstelle* von  $f$ . Ist  $x$  eine lokale Minimalstelle oder eine lokale Maximalstelle, so nennt man  $x$  eine lokale *Extremstelle* von  $f$ .

**9.8** Es sei  $f: U \rightarrow \mathbb{R}$  eine partiell differenzierbare Funktion auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ . Wir nennen  $x \in U$  einen *kritischen Punkt* von  $f$ , wenn

$$\frac{\partial f}{\partial x_1}(x) = \cdots = \frac{\partial f}{\partial x_n}(x) = 0.$$

**9.9** Ist  $U \subseteq \mathbb{R}^n$  offen und  $f: U \rightarrow \mathbb{R}$  partiell differenzierbar, so definieren wir den *Gradienten*  $\nabla f(x) \in \mathbb{R}^n$  von  $f$  an der Stelle  $x \in U$  als den Spaltenvektor der dortigen partiellen Ableitungen:

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

Hierbei wird  $\nabla$  "Nabla" ausgesprochen. Statt  $\nabla f(x)$  schreibt man auch  $\text{grad } f(x)$ .

Es ist also  $x \in U$  genau dann ein kritischer Punkt von  $f$ , wenn  $\nabla f(x) = 0$  der Nullvektor ist.

**9.10** (Notwendige Bedingung für Extremstellen). *Es sei  $f: U \rightarrow \mathbb{R}$  eine partiell differenzierbare Funktion auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ . Ist  $x \in U$  eine lokale Extremstelle von  $f$ , so ist  $\nabla f(x) = 0$ , also  $x$  ein kritischer Punkt von  $f$ .*

Begründung: Sei etwa  $x$  eine lokale Minimalstelle. Da  $U$  offen ist, gibt es ein  $\varepsilon > 0$  mit  $B_\varepsilon(x) \subseteq U$ . Nach Verkleinern von  $\varepsilon$  dürfen wir annehmen, dass

$$f(x) \leq f(y) \quad \text{für alle } y \in B_\varepsilon(x).$$

Für  $t \in ]-\varepsilon, \varepsilon[$  ist  $|te_j| < \varepsilon$ , also  $x + te_j \in B_\varepsilon(x)$  und somit

$$f(x + te_j) - f(x) \geq 0.$$

Ist  $t \in ]0, \varepsilon[$ , liefert Teilen durch  $t$

$$\frac{f(x + te_j) - f(x)}{t} \geq 0.$$

Mit  $t \searrow 0$  folgt  $\frac{\partial f}{\partial x_j}(x) \geq 0$ . Ist  $t \in ]-\varepsilon, 0[$ , liefert Teilen durch  $t$

$$\frac{f(x + te_j) - f(x)}{t} \leq 0 :$$

und mit  $t \nearrow 0$  folgt  $\frac{\partial f}{\partial x_j}(x) \leq 0$ . Also ist  $\frac{\partial f}{\partial x_j}(x) = 0$ .

**Beispiel 9.11** Die Funktion

$$f: ]-1, 1[ \times ]-\frac{\pi}{2}, \frac{\pi}{2}[ \rightarrow \mathbb{R}, \quad (x, y) \mapsto (1 - x^2) \cos(y)$$

ist stetig differenzierbar mit

$$\nabla f(x, y) = \begin{pmatrix} -2x \cos y \\ -(1 - x^2) \sin y \end{pmatrix}.$$

Für einen kritischen Punkt  $(x, y)$  muss

$$-2x \cos(y) = 0 \quad \text{und} \quad (1 - x^2) \sin(y) = 0$$

gelten, also  $x = 0$  und  $y = 0$ . Es ist also  $(0, 0)$  der einzige kritische Punkt von  $f$ ; nur dort kann eine lokale Extremstelle vorliegen.

In der Tat  $f(0, 0) = 1$  das globale Maximum für  $f$ , denn für alle  $(x, y)$  im Definitionsbereich ist  $0 < (1 - x^2) \leq 1$  und  $0 < \cos(y) \leq 1$ , somit

$$0 < (1 - x^2) \cos(y) \leq 1.$$

Sei wieder  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge. Im folgenden Kapitel werden wir zeigen:

**9.12** *Jede stetig differenzierbare Funktion  $f: U \rightarrow \mathbb{R}$  ist stetig.*

Mit ähnlichen Argumenten werden wir zudem beweisen:

**9.13** Ist  $U \subseteq \mathbb{R}^n$  offen und  $f: U \rightarrow \mathbb{R}$  stetig differenzierbar, so existiert für jedes  $x \in U$  und jedes  $v \in \mathbb{R}^n$  der Grenzwert

$$(D_v f)(x) := \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}$$

und es gilt

$$(D_v f)(x) = \nabla f(x) \cdot v$$

(Skalarprodukt).

Dann ist also  $(D_v f)(x) = \left. \frac{d}{dt} \right|_{t=0} f(x + tv)$ . Man nennt  $(D_v f)(x)$  die *Richtungsableitung* von  $f$  an der Stelle  $x$  in der Richtung  $v$ .

Bevor wir 9.12 und 9.13 beweisen, schauen wir uns eine Anwendung an. Man kann den Gradienten als die *Richtung des steilsten Anstiegs* interpretieren. Befindet man sich an der Stelle  $x \in U$  und bewegt sich in  $U$  in Richtung des Gradienten, so wächst die Funktion  $f$  am schnellsten.

Ist etwa  $f(x, y)$  die Höhe der Bergoberfläche über dem Punkt  $(x, y)$  in der Ebene, so kommen Sie am schnellsten bergauf, wenn Sie sich in Richtung  $\nabla f(x, y)$  bewegen.

**9.14** Sei  $f: U \rightarrow \mathbb{R}$  stetig differenzierbar und  $x \in U$ . Ist  $\nabla f(x) \neq 0$ , so wird das Maximum

$$\max\{(D_v f)(x) : v \in \mathbb{R}^n \text{ mit } |v| = 1\}$$

genau für

$$v = \frac{\nabla f(x)}{|\nabla f(x)|}$$

angenommen.

Analog ist  $-\nabla f(x)$  die Richtung des steilsten Abstiegs.

Zum Nachweis brauchen wir Information darüber, wann in der Cauchy-Schwarzschen Ungleichung Gleichheit auftritt:

**9.15** Sind  $u, v \in \mathbb{R}^n$  mit  $|v| = 1$  und ist

$$|u \cdot v| = |u| |v| = |u|,$$

so ist  $u$  ein Vielfaches von  $v$ .

Nachweis: Setzen wir  $u_{\parallel} := (u \cdot v)v$  und  $u_{\perp} := u - u_{\parallel}$ , so ist  $u = u_{\parallel} + u_{\perp}$ .  
Wegen

$$u_{\parallel} \cdot u_{\perp} = (u \cdot v)v \cdot u_{\perp} = (u \cdot v)(v \cdot u) - (u \cdot v)v \cdot (u \cdot v)v = |u \cdot v|^2 - |u \cdot v|^2 |v|^2 = 0$$

sind die Vektoren  $u_{\parallel}$  und  $u_{\perp}$  orthogonal. Ist  $u$  kein Vielfaches von  $v$  und somit  $u_{\perp} \neq 0$ , so gilt nach dem Satz des Pythagoras (siehe 7.53)

$$|v|^2 = |v_{\parallel}|^2 + |v_{\perp}|^2 > |v_{\parallel}|^2.$$

Wegen

$$v \cdot u = v \cdot u_{\parallel} + \underbrace{v \cdot u_{\perp}}_{=0} = v \cdot u_{\parallel}$$

folgt

$$|v \cdot u| = |v \cdot u_{\parallel}| \leq |v| |u_{\parallel}| = |u_{\parallel}| < |u|.$$

Beweis für 9.14. Nach der Cauchy-Schwarzschen Ungleichung ist für jedes  $v \in \mathbb{R}^n$  mit  $|v| = 1$

$$v \cdot \nabla f(x) \leq |v \cdot \nabla f(x)| \leq |v| |\nabla f(x)| = |\nabla f(x)|.$$

Die obere Schranke wird auch angenommen, denn mit  $u := \frac{\nabla f(x)}{|\nabla f(x)|}$  ist

$$u \cdot \nabla f(x) = \frac{1}{|\nabla f(x)|} \nabla f(x) \cdot \nabla f(x) = \frac{1}{|\nabla f(x)|} |\nabla f(x)|^2 = |\nabla f(x)|.$$

Analog wird für  $-u$  das Minimum

$$(-u) \cdot \nabla f(x) = -|\nabla f(x)|$$

angenommen. Ist  $v$  weder  $+u$  noch  $-u$ , so ist der Einheitsvektor  $v$  kein Vielfaches des Einheitsvektors  $u$  und somit nach 9.15

$$|u \cdot v| < 1,$$

also  $u \cdot \nabla f(x) \leq |u \cdot \nabla f(x)| < |\nabla f(x)|$ .

**Beispiel 9.16** Ein Käfer befindet sich an der Stelle  $(1, 0)$  auf der Herdplatte

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 16\}$$

vom Radius 4 cm. Die Temperatur an der Stelle  $(x, y)$  sei

$$T(x, y) = 40 - x^2 - y^2.$$

In welcher Richtung sollte er davonkrabbeln, damit es auf die Dauer nicht zu heiß wird?

Antwort: In die Richtung  $-\nabla T(1, 0)$  des steilsten Abstiegs. Es ist  $\nabla T(x, y) = (-2x, -2y)$ , somit  $-\nabla T(1, 0) = (2, 0)$ . Der Käfer sollte sich also radial nach außen bewegen, längs der  $x$ -Achse.

Analog dazu, wie die Tangente eine differenzierbare Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  um eine Stelle  $x \in \mathbb{R}$  approximiert, können wir partiell differenzierbare Funktionen durch um eine Stelle  $x$  approximieren durch eine affin-lineare Funktion.

**9.17** Ist  $U \subseteq \mathbb{R}^n$  eine offene Menge,  $f: U \rightarrow \mathbb{R}$  stetig differenzierbar und  $x \in U$ , so setzen wir

$$R(y) := f(y) - f(x) - (y - x) \cdot \nabla f(x)$$

für  $y \in U$ , so dass also

$$\begin{aligned} f(y) &= f(x) + \nabla f(x) \cdot (y - x) + R(y) \\ &= f(x) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)(y_j - x_j) + R(y). \end{aligned}$$

Dann gilt

$$\lim_{y \rightarrow x} \frac{R(y)}{|y - x|} = 0.$$

Den Nachweis führen wir im folgenden Kapitel.

**Beispiel 9.18** (Fehlerabschätzung bei Messfehlern). Es sei  $U \subseteq \mathbb{R}^n$  offen und  $f: U \rightarrow \mathbb{R}$  stetig differenzierbar. Im Labor haben wir gewisse Messwerte  $x_1, \dots, x_n$  erhalten. Diese können wegen der Messungenauigkeit abweichen von der exakten Werten  $y_1, \dots, y_n$ . Meist kennt man obere Schranken  $\Delta_j$  für den Messfehler:

$$|y_j - x_j| \leq \Delta_j$$

für  $j \in \{1, \dots, n\}$ . Mit der affin-linearen Approximation

$$f(y) - f(x) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)(y_j - x_j) + R(y)$$

erhält man als exakten Messfehler

$$|f(y) - f(x)| = \left| \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)(y_j - x_j) + R(y) \right|.$$

Man vernachlässigt nun das Restglied  $R(y)$  und erhält

$$\begin{aligned} |f(y) - f(x)| &\approx \left| \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)(y_j - x_j) \right| \\ &\leq \sum_{j=1}^n \left| \frac{\partial f}{\partial x_j}(x) \right| |y_j - x_j| \\ &\leq \sum_{j=1}^n \left| \frac{\partial f}{\partial x_j}(x) \right| \Delta_j. \end{aligned}$$

Näherungsweise gilt für den Fehler der aus den Messwerten berechneten Größe  $f(x)$  im Vergleich zu  $f(y)$  also

$$|f(y) - f(x)| \lesssim \sum_{j=1}^n \left| \frac{\partial f}{\partial x_j}(x) \right| \Delta_j.$$

## 10 Beweise für 9.12, 9.13 und 9.17

Die Beweise für die drei genannten Fakten sind sehr ähnlich und daher hier gemeinsam dargestellt. In der Vorlesung behandeln wir an der Tafel nur den Fall  $n = 2$ ,  $U = \mathbb{R}^2$ , was die Notation einfacher macht. Der Vollständigkeit halber finden Sie hier dennoch die (etwas technischen) allgemeinen Beweise.

Wir beginnen mit einer Vorüberlegung. Sei  $x = (x_1, \dots, x_n) \in U$  und  $\varepsilon > 0$  derart, dass  $x + t e_1 \in U$  für alle  $t \in ]-\varepsilon, \varepsilon[$ . Für jedes  $t \in ]-\varepsilon, \varepsilon[$  ist dann

$$f(x_1 + t, x_2, \dots, x_n) - f(x_1, \dots, x_n) = \frac{\partial f}{\partial x_1}(\xi, x_2, \dots, x_n) t \quad (51)$$

für ein  $\xi$  zwischen  $x_1$  und  $x_1 + t$ . Für  $t = 0$  sind nämlich beide Seiten 0 mit  $\xi := x_1$ . Ist  $t \neq 0$ , betrachten wir  $x_2, \dots, x_n$  als Konstanten, so dass also  $y \mapsto f(y, x_2, \dots, x_n)$  eine differenzierbare Funktion der reellen Variablen  $y \in [x_1, x_1 + t]$  (bzw.  $y \in [x_1 - t, x_1]$ ) ist. Nach dem Mittelwertsatz der

Differentialrechnung (siehe Mathematik 1) ist die Sekantensteigung gleich der Tangentensteigung an einer Zwischenstelle  $\xi$  zwischen  $x_1$  und  $x_1 + t$ , also

$$\frac{f(x_1 + t, x_2, \dots, x_n) - f(x_1, \dots, x_n)}{t} = \frac{\partial f}{\partial x_1}(\xi, x_2, \dots, x_n).$$

Multiplizieren mit  $t$  liefert (51).

Setzen wir  $\tau := \xi - x_1$ , so ist  $\xi = x_1 + \tau$  und

$$f(x + te_1) - f(x) = \frac{\partial f}{\partial x_1}(x + \tau e_1) t$$

für ein  $\tau$  zwischen 0 und  $t$ . Analog sehen wir, dass für jedes  $j \in \{1, \dots, n\}$  und  $t \in ]-\varepsilon, \varepsilon[$

$$f(x + te_j) - f(x) = \frac{\partial f}{\partial x_j}(x + \tau e_j) t$$

für ein  $\tau$  zwischen 0 und  $t$ , wenn  $x + se_j \in U$  für alle  $s \in ]-\varepsilon, \varepsilon[$ .

**Beweis für 9.12.** In der Situation von 9.12 zeigen wir nun Stetigkeit von  $f$  an der Stelle  $x = (x_1, \dots, x_n) \in U$ . Da  $U$  offen ist, gibt es ein  $\varepsilon > 0$  mit  $B_{\varepsilon\sqrt{n}}(x) \subseteq U$ . Dann ist

$$]x_1 - \varepsilon, x_1 + \varepsilon[ \times \dots \times ]x_n - \varepsilon, x_n + \varepsilon[ \subseteq B_{\varepsilon\sqrt{n}}(x) \subseteq U,$$

denn für jedes  $y = (y_1, \dots, y_n)$  im linken offenen Quader gilt  $|y - x|^2 = \sum_{j=1}^n (y_j - x_j)^2 < \sum_{j=1}^n \varepsilon^2 \leq n\varepsilon^2$ . Weiter ist laut Vorüberlegung

$$\begin{aligned} f(y) - f(x) &= \sum_{k=1}^n \left( f \left( x + \sum_{j=1}^k (y_j - x_j) e_j \right) - f \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j \right) \right) \\ &= \sum_{k=1}^n \left( f \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j + (y_k - x_k) e_k \right) - f \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j \right) \right) \\ &= \sum_{k=1}^n \frac{\partial f}{\partial x_k} \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j + \tau_k e_k \right) (y_k - x_k) \end{aligned} \quad (52)$$

mit Zahlen  $\tau_k$  zwischen 0 und  $y_k - x_k$ , so dass also  $\tau_k \rightarrow 0$  für  $y \rightarrow x$ . Aufgrund der Stetigkeit der beteiligten Funktionen geht die rechte Seite von (52) gegen

$$\sum_{k=1}^n \frac{\partial f}{\partial x_k} \left( x + \sum_{j=1}^{k-1} (x_j - x_j) e_j + 0e_k \right) (x_k - x_k) = 0$$

für  $y \rightarrow x$ , also  $f(y) \rightarrow f(x)$ .

**Beweis von 9.13.** Es sei  $x \in U$  und  $\varepsilon$  wie zuvor. Für  $v = (v_1, \dots, v_n) \in \mathbb{R}^n$  wählen wir  $\delta > 0$  so klein, dass  $\delta|v_j| < \varepsilon$  für alle  $j \in \{1, \dots, n\}$ . Für alle  $t \in ]-\delta, \delta[$  ist dann also

$$y := x + tv \in ]x_1 - \varepsilon, x_1 + \varepsilon[ \times \dots \times ]x_n - \varepsilon, x_n + \varepsilon[.$$

Zu  $y$  finden wir  $\tau_1, \dots, \tau_n$  wie oben mit  $\tau_j$  zwischen 0 und  $y_j - x_j = tv$ , so dass also

$$|\tau_j| < |t| |v|$$

und insbesondere  $\tau_j \rightarrow 0$  für  $t \rightarrow 0$ . Für  $t \in ]-\delta, \delta[ \setminus \{0\}$  liefert Teilen von (52) durch  $t$

$$\begin{aligned} \frac{f(x + tv) - f(x)}{t} &= \sum_{k=1}^n \frac{\partial f}{\partial x_k} \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j + \tau_k e_k \right) v_k \\ &= \sum_{k=1}^n \frac{\partial f}{\partial x_k} \left( x + t \sum_{j=1}^{k-1} v_j e_j + \tau_k e_k \right) v_k; \end{aligned}$$

hierbei haben wir benutzt, dass  $y_k - x_k = v_k$ . Wegen der Stetigkeit der beteiligten Funktion folgt für  $t \rightarrow 0$

$$\frac{f(x + tv) - f(x)}{t} \rightarrow \sum_{k=1}^n \frac{\partial f}{\partial x_k}(x) v_k = v \cdot \nabla f(x).$$

Es ist also  $D_v f(x) = v \cdot \nabla f(x)$ .

**Beweis für 9.17.** Gegeben  $x \in U$  sei  $\varepsilon > 0$  wie im Beweis von 9.12. Für  $y \in ]x_1 - \varepsilon, x_1 + \varepsilon[ \times \dots \times ]x_n - \varepsilon, x_n + \varepsilon[$  ist nach (52) dann

$$f(y) - f(x) - (y - x) \cdot \nabla f(x) = \sum_{k=1}^n \left( \frac{\partial f}{\partial x_k} \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j + \tau_k e_k \right) - \frac{\partial f}{\partial x_k}(x) \right) (y_k - x_k).$$

Somit gilt

$$\begin{aligned} & \left| \frac{f(y) - f(x) - (y - x) \cdot \nabla f(x)}{|y - x|} \right| \\ & \leq \sum_{k=1}^n \underbrace{\left| \frac{\partial f}{\partial x_k} \left( x + \sum_{j=1}^{k-1} (y_j - x_j) e_j + \tau_k e_k \right) - \frac{\partial f}{\partial x_k}(x) \right|}_{\rightarrow 0} \underbrace{\frac{|y_k - x_k|}{|y - x|}}_{\leq 1} \rightarrow 0 \end{aligned}$$

für  $y \rightarrow x$  und folglich  $\frac{R(x)}{|y-x|} \rightarrow 0$ .

## 11 Höhere Ableitungen

In diesem Kapitel untersuchen wir höhere partielle Ableitungen einer Funktion  $f: U \rightarrow \mathbb{R}$  auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ .

**11.1** Ist  $f$  partiell differenzierbar, so sind die partiellen Ableitungen von  $f$  ebenfalls Funktionen

$$\frac{\partial f}{\partial x_j}: U \rightarrow \mathbb{R}$$

auf der offenen Menge  $U$ . Sind all diese partiell differenzierbar, nennen wir  $f$  *zweimal* partiell differenzierbar. Sind  $f: U \rightarrow \mathbb{R}$  stetig, die ersten partiellen Ableitungen  $\frac{\partial f}{\partial x_j}: U \rightarrow \mathbb{R}$  stetig und die zweiten partiellen Ableitungen

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f$$

alle stetig, so nennt man  $f$  *zweimal stetig differenzierbar* (oder kurz: eine  $C^2$ -Funktion).

Statt  $\frac{\partial^2 f}{\partial x_i \partial x_i}$  schreibt man meist  $\frac{\partial^2 f}{\partial x_i^2}$ .

**Beispiel 11.2** Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $x = (x_1, x_2) \mapsto x_1 \sin(x_2)$  ist partiell differenzierbar mit

$$\frac{\partial f}{\partial x_1}(x) = \sin(x_2) \quad \text{und} \quad \frac{\partial f}{\partial x_2}(x) = x_1 \cos(x_2).$$

Diese partiellen Ableitungen sind ebenfalls wieder partiell differenzierbar. Ableiten nach  $x_1$  liefert

$$\frac{\partial^2 f}{\partial x_1^2}(x) = 0 \quad \text{und} \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) = \cos(x_2).$$

Ableiten der ersten partiellen Ableitungen nach  $x_2$  liefert

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x) = \cos(x_2) \quad \text{und} \quad \frac{\partial^2 f}{\partial x_2^2}(x) = -x_1 \sin(x_2).$$

Da  $f$ , alle ersten partiellen Ableitungen und alle zweiten partiellen Ableitungen stetig sind, ist  $f$  eine  $C^2$ -Funktion.

Beachten Sie, dass im vorigen Beispiel die gemischten zweiten Ableitungen

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x)$$

gleich sind: beide sind  $\cos(x_2)$ . Das ist kein Zufall; allgemein gilt:

**11.3** (Satz von Schwarz). *Ist  $f: U \rightarrow \mathbb{R}$  eine zweimal stetig differenzierbare Funktion auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ , so gilt für alle  $i, j \in \{1, \dots, n\}$*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x).$$

**11.4** Für jede  $C^2$ -Funktion  $f: U \rightarrow \mathbb{R}$  auf  $U \subseteq \mathbb{R}^n$  ist für jedes  $x \in U$  die  $n \times n$ -Matrix

$$H_f(x) := \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

der zweiten partiellen Ableitungen also eine *symmetrische* Matrix. Man nennt  $H_f(x) \in \mathbb{R}^{n \times n}$  die *Hesse-Matrix* von  $f$ .

**Beispiel 11.5** Im vorigen Beispiel ist

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) \end{pmatrix} = \begin{pmatrix} 0 & \cos(x_2) \\ \cos(x_2) & -x_1 \sin(x_2) \end{pmatrix}.$$

Bevor wir den Satz von Schwarz beweisen, geben wir eine Anwendung der Hesse-Matrix: Diese liefert eine hinreichende Bedingung dafür, dass ein kritischer Punkt von  $f$  eine lokale Minimalstelle ist (und entsprechend für Maximalstellen).

**11.6** Eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *positiv definit*, wenn all ihre Eigenwerte  $> 0$  sind. Sind alle Eigenwerte negativ, wird  $A$  *negativ definit* genannt. Hat  $A$  sowohl einen positiven Eigenwert als auch einen negativen Eigenwert, wird  $A$  *indefinit* genannt.

**11.7** (Hinreichende Bedingung für lokale Extremstellen). *Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge,  $f: U \rightarrow \mathbb{R}$  eine  $C^2$ -Funktion und  $x \in U$  ein kritischer Punkt von  $f$ , d.h. es ist  $\nabla f(x) = 0$ . Dann gilt:*

- (a) *Ist die Hesse-Matrix  $H_f(x)$  positiv definit, so ist  $x$  eine lokale Minimalstelle von  $f$ .*
- (b) *Ist  $H_f(x)$  negativ definit, so ist  $x$  eine lokale Maximalstelle von  $x$ .*
- (c) *Ist  $H_f(x)$  indefinit, so ist  $x$  keine lokale Extremstelle von  $f$ .*

Positive Definitheit erkennt man am besten mit dem Hurwitz-Kriterium aus der linearen Algebra, das wir als Black Box benutzen.<sup>15</sup>

**11.8** (Hurwitz-Kriterium). *Eine quadratische Matrix  $A = (a_{ij}) \in \mathbb{R}^n$  ist genau dann positiv definit, wenn*

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0$$

für alle  $k \in \{1, 2, \dots, n\}$ .

Also ist eine symmetrische  $2 \times 2$ -Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

genau dann positiv definit, wenn  $a_{11} = \det(a_{11}) > 0$  und  $\det(A) > 0$ .

**Bemerkung 11.9** Eine  $n \times n$ -Matrix  $A$  ist genau dann negativ definit, wenn  $-A$  positiv definit ist. Letzteres können Sie mit dem Hurwitz-Kriterium nachrechnen.

**Beispiel 11.10** Die Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

ist positiv definit, denn es ist  $2 > 0$  und  $\det(A) = 4 - 1 = 3 > 0$ .

<sup>15</sup>Einen Beweis finden Sie z.B. im Analysis 2-Skript auf der Homepage des Dozenten.

**Beispiel 11.11** Die Matrix Die Matrix

$$A = \begin{pmatrix} -2 & 3 \\ 3 & -5 \end{pmatrix}$$

ist negativ definit, denn

$$-A = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}$$

ist positiv definit nach dem Hurwitz-Kriterium, weil  $2 > 0$  und  $\det(-A) = 10 - 9 = 1 > 0$ .

**Beispiel 11.12** Wir wissen aus Beispiel 9.11 dass die stetig differenzierbare Funktion

$$f: ]-1, 1[ \times ]-\frac{\pi}{2}, \frac{\pi}{2}[ \rightarrow \mathbb{R}, \quad (x, y) \mapsto (1 - x^2) \cos(y)$$

am kritischen Punkt  $(0, 0)$  ein lokales (und sogar globales) Maximum besitzt. Dies können wir auch mittels der Hessematrix nachrechnen. Es ist

$$\frac{\partial f}{\partial x}(x, y) = -2x \cos y \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = -(1 - x^2) \sin y.$$

Diese Funktionen stetig differenzierbar (also  $f$  eine  $C^2$ -Funktion). Berechnung der zweiten partiellen Ableitung liefert die Hesse-Matrix

$$H_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix} = \begin{pmatrix} -2 \cos y & 2x \sin(y) \\ 2x \sin(y) & -(1 - x^2) \cos(y) \end{pmatrix}.$$

Am kritischen Punkt  $(0, 0)$  erhalten wir

$$H_f(0, 0) = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$$

Dies ist eine Diagonalmatrix, die Diagonaleinträge  $-2$  und  $-1$  sind also die Eigenwerte der Matrix. Da beide negativ sind, ist die Matrix negativ definit und somit liegt am kritischen Punkt  $(0, 0)$  von  $f$  eine lokale Maximalstelle vor.

Unsere hinreichende Bedingung für lokale Extremstellen beruht auf Taylor-Entwicklung 2. Ordnung.

**11.13** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $f: U \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Sei  $x \in U$ . Für das Restglied  $R(y)$  in

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle y - x, H_f(x)(y - x) \rangle + R(y) \quad (53)$$

gilt dann

$$\lim_{y \rightarrow x} \frac{R(y)}{|y - x|^2} = 0.$$

Für ein  $r > 0$  enthält  $U$  die Kugel  $B_r(x)$ . Für  $y \in B_r(x)$  betrachten wir die stetige Funktion

$$h: [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x + t(y - x)).$$

Diese ist differenzierbar: Für jedes  $t \in [0, 1]$  ist

$$h'(t) = D_{y-x}f(x + t(y - x))$$

die Richtungsableitung von  $f$  an der Stelle  $x + t(y - x)$  in der Richtung  $y - x$ . Nach 9.13 ist also

$$h'(t) = \langle \nabla f(x + t(y - x)), y - x \rangle = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x + t(y - x))(y_j - x_j). \quad (54)$$

Diese Funktion kann erneut nach  $t$  abgeleitet werden; wir erhalten wieder eine Richtungsableitung, nämlich

$$\begin{aligned} h''(t) &= \sum_{j=1}^n \left\langle \nabla \left( \frac{\partial f}{\partial x_j} \right) (x + t(y - x)), y - x \right\rangle (y_j - x_j) \\ &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x + t(y - x))(y_i - x_i)(y_j - x_j) \end{aligned} \quad (55)$$

$$= \langle y - x, H_f(x + t(y - x))(y - x) \rangle; \quad (56)$$

der zweiten Zeile sehen wir an, dass  $h''(t)$  eine stetige Funktion von  $t$  ist. Taylorapproximation zweiter Ordnung für  $h$  liefert

$$h(1) = h(0) + h'(0)(1 - 0) + \frac{1}{2}h''(0)(1 - 0)^2 + R_2(1) \quad (57)$$

mit dem Restglied

$$R_2(1) = \int_0^1 (1-t)(h''(t) - h''(0)) dt, \quad (58)$$

siehe (22) in Kapitel 1. Nun ist  $h(0) = f(x)$  und  $h(1) = f(y)$ . Setzen wir dies in (57) ein sowie  $h'(0)$  aus (54) und  $h''(0)$  aus (56), so erhalten wir (53) mit

$$R(y) = R_2(1) = \int_0^1 (1-t) \sum_{i,j=1}^n \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(x + t(y-x)) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right) (y_i - x_i)(y_j - x_j) dt$$

wobei (55) in (58) eingesetzt wurde mit dem gegebenen  $t$  sowie mit  $t = 0$  (so dass  $x + t(y-x) = x$ ). Wir benutzen nun, dass die zweiten partiellen Ableitungen stetig sind. Ist  $\varepsilon > 0$  gegeben, finden wir also ein  $\delta \in ]0, r]$  derart, dass für alle  $z \in B_\delta(x) \subseteq \mathbb{R}^n$  und  $i, j \in \{1, \dots, n\}$

$$\left| \frac{\partial^2 f}{\partial x_i \partial x_j}(z) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right| < \frac{\varepsilon}{n^2}.$$

Für alle  $y \in B_\delta(x)$  ist  $z = x + t(y-x) \in B_\delta(x)$  für alle  $t \in [0, 1]$ . Wir schließen, dass

$$\begin{aligned} |R(y)| &\leq \int_0^1 \underbrace{(1-t)}_{\leq 1} \sum_{i,j=1}^n \underbrace{\left| \frac{\partial^2 f}{\partial x_i \partial x_j}(x + t(y-x)) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right|}_{\leq \varepsilon/n^2} |y_i - x_i| |y_j - x_j| dt \\ &\leq \sum_{i,j=1}^n \frac{\varepsilon}{n^2} |y_i - x_i| |y_j - x_j| \leq \sum_{i,j=1}^n \frac{\varepsilon}{n^2} |y - x|^2 = \varepsilon |y - x|^2, \end{aligned}$$

wobei  $|y_i - x_i| \leq |y - x|$  benutzt wurde. Also ist

$$\frac{|R(y)|}{|y - x|^2} \leq \varepsilon \quad \text{für alle } y \neq x \text{ mit } |y - x| < \delta$$

und somit  $\lim_{y \rightarrow x} \frac{R(y)}{|y-x|^2} = 0$  gezeigt.

Bevor wir 11.7 beweisen, schauen wir noch ein prototypisches Beispiel ein.

**Beispiel 11.14** Wir betrachten die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto y^2 - x^2.$$

Diese ist  $C^2$  mit

$$\nabla f(x, y) = \begin{pmatrix} -2x \\ 2y \end{pmatrix}$$

und

$$H_f(x, y) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Es ist also  $(0, 0)$  der einzige kritische Punkt von  $f$ . Die Eigenwerte von  $H_f(0, 0)$  sind 2 und  $-2$ . Also ist  $H_f(0, 0)$  indefinit. Folglich ist  $(0, 0)$  weder eine lokale Minimalstelle noch eine lokale Maximalstelle. Man kann den Graphen von  $f$  (bzw.  $f|_{[-1,1] \times [-1,1]}$ ) leicht skizzieren; dieser ist eine sogenannte *Sattelfläche*.

Dass kein lokales Extremum vorliegt, könnte man auch von Hand sehen: Für alle  $t > 0$  (die wir beliebig klein machen können) ist

$$f(t, 0) = -t^2 < 0 = f(0, 0),$$

also  $(0, 0)$  keine lokale Minimalstelle. Für alle  $t > 0$  ist weiter

$$f(0, t) = t^2 > 0 = f(0, 0),$$

also  $(0, 0)$  keine lokale Maximalstelle.

**Beweis für 11.7.** (a) Da die Hesse-Matrix  $H_f(x)$  positiv definit ist, sind all ihre Eigenwerte positiv. Als symmetrische reelle Matrix ist  $H_f(x)$  diagonalisierbar; es existiert eine Orthonormalbasis  $v_1, \dots, v_n$  für  $\mathbb{R}^n$ , bestehend aus Eigenvektoren für  $H_f(x)$  (siehe 6.33). Sei

$$H_f(x)v_j = \lambda_j v_j$$

mit  $\lambda_j > 0$ . Schreiben wir  $w \in \mathbb{R}^n$  als Linearkombination der Basisvektoren,

$$w = \sum_{j=1}^n t_j v_j,$$

so ist

$$|w|^2 = \sum_{j=1}^n t_j^2 |v_j|^2 = \sum_{j=1}^n t_j^2$$

nach dem Satz des Pythagoras (siehe 7.53). Weiter ist  $H_f(x)w = \sum_{j=1}^n t_j \lambda_j v_j$  und somit

$$\langle w, H_f(x)w \rangle = \left\langle \sum_{k=1}^n t_k w_k, \sum_{j=1}^n t_j \lambda_j w_j \right\rangle = \sum_{j,k=1}^n \lambda_j t_j t_k \underbrace{\langle w_k, w_j \rangle}_{=\delta_{jk}} = \sum_{j=1}^n \lambda_j t_j^2 \geq \lambda \sum_{j=1}^n t_j^2 = \lambda |w|^2$$

mit  $\lambda := \min\{\lambda_1, \dots, \lambda_n\}$ . Für jedes  $w \in \mathbb{R}^n$  gilt also

$$\langle w, H_f(x)w \rangle \geq \lambda |w|^2. \quad (59)$$

Da  $x$  ein kritischer Punkt von  $f$  ist, lautet die Taylorentwicklung zweiter Ordnung

$$f(y) = f(x) + \frac{1}{2} \langle y - x, H_f(x)(y - x) \rangle + R(y). \quad (60)$$

Da  $R(y)/|y - x|^2 \rightarrow 0$  für  $y \rightarrow x$ , gibt es ein  $\delta > 0$  derart, dass

$$\frac{R(y)}{|y - x|^2} < \frac{1}{2} \lambda$$

für alle  $y \in U$  mit  $y \neq x$  und  $|y - x| < \delta$  (vgl. 7.33). Dann ist insbesondere

$$R(y) \geq -|R(y)| > \frac{1}{2} \lambda |y - x|^2. \quad (61)$$

Schätzen wir das Skalarprodukt in (60) via (59) nach unten ab mit  $w = y - x$  und  $R(y)$  durch (61), so erhalten wir

$$f(y) > f(x) + \frac{1}{2} \lambda |y - x|^2 - \frac{1}{2} \lambda |y - x|^2 = f(x).$$

(b) Nach (a) ist  $x$  lokale Minimalstelle von  $-f$ , also lokale Maximalstelle von  $f$ .

(c) Ist ähnlich zu (a) und wird daher in der Vorlesung übersprungen. Der Vollständigkeit halber: Wir zeigen, dass  $x$  keine lokale Minimalstelle ist (analog ist  $x$  keine lokale Maximalstelle). Es sei  $\lambda$  ein negativer Eigenwert von  $H_f(x)$  und  $v$  ein zugehöriger Eigenvektor mit  $|v| = 1$ . Es gibt ein  $\delta > 0$  derart, dass  $B_\delta(x) \subseteq U$  und

$$\frac{|R(y)|}{|y - x|^2} < \frac{1}{2} |\lambda|$$

für alle  $y \in B_\delta(x)$  mit  $y \neq x$  und somit

$$R(y) < \frac{1}{2}|\lambda| |y - x|^2 = -\frac{1}{2}\lambda |y - x|^2.$$

Für jedes  $r \in ]0, \delta]$  ist

$$y := x + \frac{r}{2}v \in B_r(x)$$

und

$$\begin{aligned} f(y) &= f(x) + \frac{1}{2} \left(\frac{r}{2}\right)^2 \langle v, H_f(x)v \rangle + R(y) = f(x) + \lambda \frac{1}{2} \left(\frac{r}{2}\right)^2 + R(y) \\ &< f(x) + \lambda \frac{1}{2} \left(\frac{r}{2}\right)^2 - \frac{1}{2}\lambda |y - x|^2 = f(x). \end{aligned}$$

Also ist  $x$  keine lokale Minimalstelle.

Wir beweisen nun eine erste Fassung der Kettenregel.

**11.15** (Kettenregel). *Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $f: U \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion. Weiter sei  $I \subseteq \mathbb{R}$  ein nicht-entartetes Intervall und  $\gamma = (\gamma_1, \dots, \gamma_n): I \rightarrow \mathbb{R}^n$  eine  $C^1$ -Kurve mit  $\gamma(I) \subseteq U$ . Dann ist  $f \circ \gamma: I \rightarrow \mathbb{R}$  stetig differenzierbar und*

$$(f \circ \gamma)'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t)$$

für alle  $t \in I$ , also

$$(f \circ \gamma)'(t) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\gamma(t)) \gamma'_j(t).$$

**Beispiel 11.16** Wir betrachten die  $C^1$ -Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto (x - 1)^2 + 5y^2$$

und die  $C^1$ -Kurve  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $t \mapsto (\cos(t), \sin(t))$ . Finden Sie  $(f \circ \gamma)'(t)$  allein mit der Kettenregel (also, ohne vorher  $f(\gamma(t))$  explizit auszurechnen).

Lösung: Es ist

$$\nabla f(x, y) = \begin{pmatrix} 2(x - 1) \\ 10y \end{pmatrix},$$

also

$$\nabla f(\gamma(t)) = \begin{pmatrix} 2(\cos t - 1) \\ 10 \sin t \end{pmatrix}.$$

Weiter ist

$$\gamma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}.$$

Die Kettenregel liefert nun

$$\begin{aligned} (f \circ \gamma)'(t) &= \nabla f(\gamma(t)) \cdot \gamma'(t) = \begin{pmatrix} 2(\cos t - 1) \\ 10 \sin t \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \\ &= 10 \sin t \cos t - 2(\cos t - 1) \sin t. \end{aligned}$$

**Beweis der Kettenregel.** Sei  $t \in I$  und  $(t_k)_{k \in \mathbb{N}}$  eine Folge in  $I \setminus \{t\}$ , die gegen  $t$  konvergiert. Wir benutzen die affin-lineare Approximation

$$f(y) = f(x) + \nabla f(x) \cdot (y - x) + R(y)$$

von  $f$  an der Stelle  $x = \gamma(t)$ . Mit  $y = \gamma(t_k)$  erhalten wir

$$f(\gamma(t_k)) - f(\gamma(t)) = \nabla f(x) \cdot (\gamma(t_k) - \gamma(t)) + R(\gamma(t_k)).$$

Teilen durch  $t_k - t$  liefert

$$\frac{f(\gamma(t_k)) - f(\gamma(t))}{t_k - t} = \nabla f(x) \cdot \frac{\gamma(t_k) - \gamma(t)}{t_k - t} + \frac{R(\gamma(t_k))}{t_k - t}.$$

Der Differenzenquotient hinter dem Gradienten konvergiert gegen  $\gamma'(t)$  für  $k \rightarrow \infty$ . Es bleibt nur noch zu zeigen, dass  $R(\gamma(t_k))/(t_k - t) \rightarrow 0$ . Die gegen  $|\gamma'(t)|$  konvergente Zahlenfolge

$$\left| \frac{\gamma(t_k) - \gamma(t)}{t_k - t} \right|$$

ist beschränkt; es gibt also ein  $M \geq 0$  derart, dass

$$\left| \frac{\gamma(t_k) - \gamma(t)}{t_k - t} \right| \leq M \quad \text{für alle } k \in \mathbb{N}.$$

Gegeben  $\varepsilon > 0$  gibt es ein  $\delta > 0$  derart, dass

$$\frac{|R(z)|}{|z - x|} \leq \frac{\varepsilon}{M + 1}$$

für alle  $z \in U$  mit  $z \neq x$  und  $|z - x| < \delta$ . Es existiert ein  $N \in \mathbb{N}$  derart, dass

$$|\gamma(t_k) - \gamma(t)| < \delta$$

für alle  $k \geq N$ . Für diese  $k$  ist dann also

$$\frac{|R(\gamma(t_k))|}{|t_k - t|} \leq \frac{\varepsilon}{M+1} \frac{|\gamma(t_k) - \gamma(t)|}{|t_k - t|} \leq \varepsilon \frac{M}{M+1} < \varepsilon.$$

Also gilt  $\lim_{k \rightarrow \infty} \frac{R(t_k)}{t_k - t} = 0$ .

**11.17** Ist  $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ ,  $(s, t) \mapsto f(s, t)$  eine stetige Funktion, so ist auch

$$g: [a, b] \rightarrow \mathbb{R}, \quad s \mapsto \int_c^d f(s, t) dt$$

stetig.

Sei  $\varepsilon > 0$ . Da der Quader  $[a, b] \times [c, d]$  kompakt und somit  $f$  nach 7.52 gleichmäßig stetig ist, gibt es ein  $\delta > 0$  derart, dass für alle  $(s_1, t_1), (s_2, t_2)$  in  $[a, b] \times [c, d]$  gilt: Ist  $|(s_2, t_2) - (s_1, t_1)| < \delta$ , so folgt

$$|f(s_2, t_2) - f(s_1, t_1)| < \varepsilon.$$

Insbesondere gilt für jedes  $t \in [c, d]$  und alle  $s_1, s_2 \in [a, b]$  mit  $|s_2 - s_1| < \delta$ , dass

$$|f(s_2, t) - f(s_1, t)| < \varepsilon,$$

denn es ist  $|(s_2, t) - (s_1, t)| = |(s_2 - s_1, 0)| = |s_2 - s_1| < \delta$ . Somit ist für alle  $s_1, s_2 \in [a, b]$  mit  $|s_2 - s_1| < \delta$

$$\begin{aligned} |g(s_2) - g(s_1)| &= \left| \int_c^d (f(s_2, t) - f(s_1, t)) dt \right| \\ &\leq \int_c^d \underbrace{|f(s_2, t) - f(s_1, t)|}_{< \varepsilon} dt \\ &\leq \varepsilon(d - c). \end{aligned}$$

Da  $\varepsilon(d - c)$  beliebig klein gemacht werden kann, ist  $g$  an der Stelle  $s_1$  stetig.

**Bemerkung 11.18** Analog ist

$$[a_1, b_1] \times \cdots \times [a_n, b_n] \rightarrow \mathbb{R}, \quad s = (s_1, \dots, s_n) \mapsto \int_c^d f(s, t) dt$$

stetig für jede stetige Funktion  $f: [a_1, b_1] \times \cdots \times [a_n, b_n] \times [c, d] \rightarrow \mathbb{R}$ . Sie brauchen nur  $[a, b]$  durch  $[a_1, b_1] \times \cdots \times [a_n, b_n]$  zu ersetzen im vorigen Beweis.

**11.19** (Differenzieren unter dem Integral) *Die Funktion  $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ ,  $(s, t) \mapsto f(s, t)$  sei stetig. Weiter existiere die partielle Ableitung*

$$\frac{\partial f}{\partial s}: [a, b] \times [c, d] \rightarrow \mathbb{R}$$

und sei stetig. Dann ist die Funktion

$$g: [a, b] \rightarrow \mathbb{R}, \quad s \mapsto \int_c^d f(s, t) dt$$

stetig differenzierbar und für jedes  $s \in [a, b]$  ist

$$g'(s) = \int_c^d \frac{\partial f}{\partial s}(s, t) dt.$$

Es gilt also

$$\frac{d}{ds} \int_c^d f(s, t) dt = \int_c^d \frac{\partial f}{\partial s}(s, t) dt.$$

Zum Nachweis sei  $s \in [a, b]$ . Für  $r \in [a, b]$  mit  $r \neq s$  ist

$$\begin{aligned} \frac{g(r) - g(s)}{r - s} &= \int_c^d \frac{f(r, t) - f(s, t)}{r - s} dt \\ &= \int_c^d \int_s^r \frac{\frac{\partial f}{\partial s}(\sigma, t)}{r - s} d\sigma dt \\ &= \int_c^d \int_0^1 \frac{\partial f}{\partial s}(s + u(r - s), t) du dt; \end{aligned} \quad (62)$$

hierbei wurde für festes  $t$  der Hauptsatz der Differential- und Integralrechnung auf die Funktion  $s \mapsto f(s, t)$  angewandt und anschließend die Substitution  $\sigma = s + u(r - s)$  mit  $u \in [0, 1]$  und  $d\sigma = (r - s) du$  durchgeführt. Da die Funktion

$$[a, b] \times [c, d] \times [0, 1] \rightarrow \mathbb{R}, \quad (r, t, u) \mapsto \frac{\partial f}{\partial s}(s + u(r - s), t)$$

stetig ist, ist nach Bemerkung 11.18 die Funktion

$$h: [a, b] \times [c, d] \rightarrow \mathbb{R}, \quad (r, t) \mapsto \int_0^1 \frac{\partial f}{\partial s}(s + u(r - s), t) du$$

stetig. Wegen 11.17 ist dann auch

$$k: [a, b] \rightarrow \mathbb{R}, \quad r \mapsto \int_c^d h(r, t) dt$$

stetig. Einsetzen in (62) zeigt, dass für  $r \in [a, b] \setminus \{s\}$

$$\frac{g(r) - g(s)}{r - s} = \int_c^d h(r, t) dt = k(r).$$

Für  $r \rightarrow s$  folgt wegen der Stetigkeit von  $k$

$$\frac{g(r) - g(s)}{r - s} \rightarrow k(s) = \int_c^d h(s, t) dt, \quad (63)$$

wobei

$$h(s, t) = \int_0^1 \frac{\partial f}{\partial s}(s + u(s - s), t) du = \int_0^1 \frac{\partial f}{\partial s}(s, t) du = \frac{\partial f}{\partial s}(s, t).$$

Einsetzen in (63) liefert  $g'(s) = \int_c^d \frac{\partial f}{\partial s}(s, t) dt$ .

**Beispiel 11.20** Man berechne das Integral

$$I := \int_0^1 \frac{\ln(x + 1)}{x^2 + 1} dx.$$

Lösung. Wir betrachten

$$g: [0, 1] \rightarrow \mathbb{R}, \quad g(t) = \int_0^1 \frac{\ln(tx + 1)}{x^2 + 1} dx.$$

Wegen  $\ln(1) = 0$  ist dann  $g(0) = 0$ . Weiter ist  $g(1) = I$ . Nach dem Hauptsatz der Integral- und Differentialrechnung ist

$$g(t) = g(0) + \int_0^t g'(s) ds = \int_0^t g'(s) ds.$$

Daher interessiert uns

$$g'(t) = \int_0^1 \frac{\partial}{\partial t} \frac{\ln(tx+1)}{x^2+1} dx = \int_0^1 \frac{x}{(tx+1)(x^2+1)} dx;$$

die Ableitung durfte wegen 11.19 unter das Integral gezogen werden. Wir benutzen nun die Partialbruchzerlegung

$$\frac{x}{(tx+1)(x^2+1)} = \frac{1}{t^2+1} \left( \frac{t}{x^2+1} + \frac{x}{x^2+1} - \frac{t}{tx+1} \right)$$

mit Stammfunktion

$$h_t(x) = \frac{1}{t^2+1} \left( t \arctan(x) + \frac{1}{2} \ln(x^2+1) - \ln(tx+1) \right)$$

bzgl.  $x$  und folgern, dass

$$\begin{aligned} g'(t) &= [h_t(x)]_{x=0}^{x=1} = \frac{1}{t^2+1} \left( t \arctan(1) + \frac{1}{2} \ln(2) - \ln(t+1) \right) \\ &= \frac{1}{t^2+1} \left( t \frac{\pi}{4} + \frac{1}{2} \ln(2) - \ln(t+1) \right). \end{aligned}$$

Also ist

$$I = g(1) = \int_0^1 g'(s) ds = \int_0^1 \frac{t\pi + 2 \ln(2)}{4(t^2+1)} dt - \underbrace{\int_0^1 \frac{\ln(t+1)}{t^2+1} dt}_{=I}.$$

Auflösen nach  $I$  liefert

$$\begin{aligned} I &= \frac{1}{8} \int_0^1 \frac{t\pi + 2 \ln(2)}{t^2+1} dt = \frac{1}{8} \left[ \frac{\pi}{2} \ln(t^2+1) + 2 \ln(2) \arctan(t) \right]_{t=0}^{t=1} \\ &= \frac{\pi}{16} \ln(2) + \frac{\ln(2)}{4} \underbrace{\arctan(1)}_{=\pi/4} = \frac{\pi \ln(2)}{8}. \end{aligned}$$

**Beweis des Satzes von Schwarz.** Gegeben  $x = (x_1, \dots, x_n) \in U$  existiert ein  $r > 0$  mit  $B_{r\sqrt{n}}(x) \subseteq U$ . Dann ist

$$]x_1 - r, x_1 + r[ \times \cdots \times ]x_n - r, x_n + r[ \subseteq B_{r\sqrt{n}}(x) \subseteq U$$

(vgl. 7.53). Seien  $i, j \in \{1, \dots, n\}$  mit  $i \neq j$ . Da beim partiellen Ableiten nach  $x_i$  und  $x_j$  all die anderen Variablen festgehalten werden, genügt es, den Fall  $n = 2$  zu diskutieren und eine  $C^2$ -Funktion

$$f: ]a_1, b_1[ \times ]a_2, b_2[ \rightarrow \mathbb{R}$$

zu betrachten mit reellen  $a_1 < b_1$  und  $a_2 < b_2$ . Sei  $x_2 \in ]a_2, b_2[$  und  $y_2 \in ]a_2, b_2[$  mit  $y_2 \neq x_2$ . Für jedes  $x_1 \in ]a_1, b_1[$  ist dann

$$\frac{f(x_1, y_2) - f(x_1, x_2)}{y_2 - x_2} = \int_{x_2}^{y_2} \frac{\frac{\partial f}{\partial x_2}(x_1, s)}{y_2 - x_2} ds = \int_0^1 \frac{\partial f}{\partial x_2}(x_1, x_2 + t(y_2 - x_2)) dt. \quad (64)$$

Hierbei wurde bei festgehaltener erster Variable der Hauptsatz benutzt und dann  $s = x_2 + t(y_2 - x_2)$ ,  $ds = (y_2 - x_2) dt$  substituiert. Ableiten von (64) nach  $x_1$  liefert

$$\Delta_{y_2} := \frac{\frac{\partial f}{\partial x_1}(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)}{y_2 - x_2} = \int_0^1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2 + t(y_2 - x_2)) dt. \quad (65)$$

Der Differenzenquotient in (65) ergibt sich dabei direkt aus dem in (64) durch partielles Ableiten nach  $x_1$ . Partielles Ableiten des rechten Integrals in (64) liefert nach 11.19 das Integral in (65); die Ableitung darf ins Integral gezogen werden. Betrachtung des Differenzenquotienten in (65) zeigt, dass

$$\lim_{y_2 \rightarrow x_2} \Delta_{y_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2). \quad (66)$$

Nun ist aber

$$h: ]a_2, b_2[ \times ]0, 1[ \rightarrow \mathbb{R}, \quad (y_2, t) \mapsto \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2 + t(y_2 - x_2))$$

eine stetige Funktion. Nach 11.17 ist somit<sup>16</sup>

$$g: ]a_2, b_2[ \rightarrow \mathbb{R}, \quad y_2 \mapsto \int_0^1 h(y_2, t) dt$$

---

<sup>16</sup>Genauer hat man Stetigkeit von  $g$  zunächst auf kleineren kompakten Intervallen der Form  $]a_2 + \varepsilon, b_2 - \varepsilon[$  mit einem  $\varepsilon > 0$  (aber dann auch auf deren Vereinigung  $]a_2, b_2[$ ).

stetig. Man beachte, dass  $h(x_2, t) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2)$  für alle  $t \in [0, 1]$  und somit  $g(x_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2)$ . Die Integraldarstellung aus (65) liefert

$$\Delta_{y_2} = \int_0^1 h(y_2, t) dt = g(y_2)$$

für alle  $y_2 \in ]a_2, b_2[$  mit  $y_2 \neq x_2$ . Die Stetigkeit von  $g$  liefert

$$\lim_{y_2 \rightarrow x_2} \Delta_{y_2} = g(x_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2). \quad (67)$$

Vergleich von (66) und (67) liefert die Gleichheit der gemischten zweiten partiellen Ableitungen.

Bisher haben wir stetige Funktionen betrachtet, reellwertige einmal stetig differenzierbare Funktionen und reellwertige zweimal stetig differenzierbare Funktionen.

**11.21** (Höhere Differenzierbarkeit). Ist  $k \in \mathbb{N} \cup \{\infty\}$  und existieren die partiellen Ableitungen

$$\frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}} := \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_j}} f$$

für alle  $j \in \mathbb{N}$  mit  $j \leq k$  und alle  $i_1, \dots, i_j \in \{1, \dots, n\}$ , so sagen wir,  $f$  sei  $k$  mal partiell differenzierbar. Sind zudem  $f: U \rightarrow \mathbb{R}$  und die vorigen partiellen Ableitungen

$$\frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}: U \rightarrow \mathbb{R}$$

stetige Funktionen, wird  $f$  eine  $k$  mal stetig differenzierbare Funktion (oder kurz: eine  $C^k$ -Funktion) genannt.

**Bemerkung 11.22** Für  $k \in \mathbb{N}$  mit  $k \geq 2$  ist eine Funktion  $f: U \rightarrow \mathbb{R}$  genau dann  $C^k$ , wenn  $f$  eine  $C^1$ -Funktion ist und jede der partiellen Ableitungen  $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$  eine  $C^{k-1}$ -Funktion.

**11.23** Statt  $\frac{\partial}{\partial x_i} \cdots \frac{\partial}{\partial x_i}$  (mit  $m$  Faktoren) schreiben wir auch  $\frac{\partial^m}{\partial x_i^m}$ .

**Beispiel 11.24** Jede konstante Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \mapsto c$  ist  $C^\infty$ , also  $C^k$  für alle  $k \in \mathbb{N}_0$ . Dann  $f$  ist stetig, also  $C^0$ . Sei  $k \in \mathbb{N}$ . Induktionsvoraussetzung: Jede konstante Funktion ist  $C^{k-1}$ . Nun gilt

$$\frac{\partial f}{\partial x_i} = 0$$

für alle  $i \in \{1, \dots, n\}$ , d.h. alle partiellen Ableitungen erster Ordnung von  $f$  sind wieder konstant (also stetig, so dass  $f$  eine  $C^1$ -Funktion ist) und sogar  $C^{k-1}$ , per Induktionsvoraussetzung. Somit ist  $f$  eine  $C^k$ -Funktion, nach 11.22.

**Beispiel 11.25** Jede lineare Abbildung

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \sum_{j=1}^n a_j x_j$$

mit  $a_1, \dots, a_j \in \mathbb{R}$  ist  $C^\infty$ . Denn es ist

$$\frac{\partial f}{\partial x_j} = a_j$$

eine konstante (und somit stetige) Funktion für alle  $j \in \{1, \dots, n\}$ , somit  $f$  eine  $C^1$ -Funktion. Für jedes  $k \in \mathbb{N}$  ist  $\frac{\partial f}{\partial x_i}$  als konstante Funktion eine  $C^{k-1}$ -Funktion (siehe Beispiel 11.24), also  $f$  eine  $C^k$ -Funktion nach 11.22.

**11.26** (Partielle Ableitungen in Standard-Reihenfolge). Ist  $f: U \rightarrow \mathbb{R}$  eine  $C^k$ -Funktion auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^n$ , so schreiben wir für jeden Multiindex  $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$  mit  $|\alpha| \leq k$

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f.$$

**11.27** Es seien  $U \subseteq \mathbb{R}^n$  eine offene Menge,  $f: U \rightarrow \mathbb{R}$  eine  $C^k$ -Funktion mit  $k \in \mathbb{N}$  und  $i_1, \dots, i_k \in \{1, \dots, n\}$ . Für  $j \in \{1, \dots, n\}$  sei  $\alpha_j$  die Anzahl der  $\ell \in \{1, \dots, k\}$  mit

$$i_\ell = j.$$

Mit  $\alpha := (\alpha_1, \dots, \alpha_n)$  gilt dann

$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f = \frac{\partial^{|\alpha|} f}{\partial x^\alpha},$$

wobei  $|\alpha| = k$ .

**Beispiel 11.28** Ist  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  eine  $C^3$ -Funktion, so ist

$$\frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_1}(x_1, x_2) = \frac{\partial^3 f}{\partial x_1^2 \partial x_2}(x_1, x_2) = \frac{\partial^{|\alpha|}}{\partial x^\alpha}(x_1, x_2)$$

mit  $\alpha = (2, 1)$ .

Nachweis für 11.27 (in der Vorlesung übersprungen): Der Beweis ist per Induktion nach  $k$ . Der Fall  $k = 0$  ist trivial,  $f$  null mal abgeleitet ist per Definition  $f = \frac{\partial^{|\alpha|}}{\partial x^\alpha} f$  mit  $\alpha = (0, \dots, 0)$ .

Sei nun  $k \geq 1$ . Wird nur nach der letzten Variable abgeleitet,  $i_1, \dots, i_k = n$ , sind wir fertig. Andernfalls ist

$$j := \min\{i_1, \dots, i_k\} < n.$$

Dann ist also  $\alpha_j \geq 1$  aber  $\alpha_1 = \dots = \alpha_{j-1} = 0$ . Sei  $\ell \in \{1, \dots, n\}$  minimal gewählt mit  $i_\ell = j$ . Ist  $\ell > 1$ , so können wir in

$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_{\ell-2}}} \frac{\partial}{\partial x_{i_{\ell-1}}} \frac{\partial}{\partial x_{i_\ell}} \frac{\partial}{\partial x_{i_{\ell+1}}} \cdots \frac{\partial}{\partial x_{i_k}} f(x)$$

$\leftrightarrow$

mit dem Satz von Schwarz nacheinander

$$\frac{\partial}{\partial x_{\ell-1}}, \dots, \frac{\partial}{\partial x_{i_1}}$$

mit  $\frac{\partial}{\partial x_{i_\ell}}$  vertauschen; wir erhalten

$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f = \frac{\partial}{\partial x_{i_\ell}} \left( \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_{\ell-1}}} \frac{\partial}{\partial x_{i_{\ell+1}}} \cdots \frac{\partial}{\partial x_{i_k}} f \right),$$

wobei  $x_{i_\ell} = x_j$ . Per Induktionsvoraussetzung ist die Funktion in Klammern gleich

$$\frac{\partial^{k-1}}{\partial x^{\alpha - e_j}} f = \frac{\partial^{\alpha_j - 1}}{\partial x_j^{\alpha_j - 1}} \frac{\partial^{\alpha_{j+1}}}{\partial x_{j+1}^{\alpha_{j+1}}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f;$$

insgesamt erhalten wir also

$$\frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}} \frac{\partial^{\alpha_{j+1}}}{\partial x_{j+1}^{\alpha_{j+1}}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f = \frac{\partial^{|\alpha|}}{\partial x^\alpha} f.$$

Aus der Mathematik 1 wissen wir für Funktionen  $f, g: I \rightarrow \mathbb{R}$  einer reellen Variablen, dass

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

und  $(rf)' = rf'$  für  $r \in \mathbb{R}$ . Da wir beim partiellen Ableiten alle bis auf eine Variable festhalten, folgt:

**11.29** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $k \in \mathbb{N}_0$ . Dann ist für alle  $C^k$ -Funktionen  $f, g: U \rightarrow \mathbb{R}$  auch  $f + g: U \rightarrow \mathbb{R}$ ,  $x \mapsto f(x) + g(x)$  eine  $C^k$ -Funktion und auch  $rf: U \rightarrow \mathbb{R}$ ,  $x \mapsto rf(x)$  für jedes  $r \in \mathbb{R}$ . Es gilt

$$\frac{\partial^{|\alpha|}(f + g)}{\partial x^\alpha} = \frac{\partial^{|\alpha|}f}{\partial x^\alpha} + \frac{\partial^{|\alpha|}g}{\partial x^\alpha}$$

und

$$\frac{\partial^\alpha(rf)}{\partial x^\alpha} = r \frac{\partial^{|\alpha|}f}{\partial x^\alpha}$$

für alle  $\alpha \in (\mathbb{N}_0)^n$  mit  $|\alpha| \leq k$ .

Beweis für  $f + g$  per Induktion nach  $k$  (der Falle  $rf$  wird analog behandelt). Induktionsanfang  $k = 0$ : Summen stetiger Funktionen sind stetig.

Induktionsschritt: Sei  $k \geq 1$  und gelte die Aussage für  $k - 1$  an Stelle von  $k$ . Nach der Vorüberlegung existieren die ersten partiellen Ableitungen von  $f + g$  und sind gleich

$$\frac{\partial(f + g)}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}.$$

Da die rechte Seite stetig ist, ist  $f$  eine  $C^1$ -Funktion. Da die rechte Seite per Induktionsvoraussetzung sogar  $C^{k-1}$  ist, ist  $f$  eine  $C^k$ -Funktion (siehe 11.22). Jedes  $\frac{\partial}{\partial x_i}$  macht aus einer Summe von Funktionen die Summe der partielle Ableitungen, also auch die Komposition  $\frac{\partial^{|\alpha|}}{\partial x^\alpha}$ .

**11.30** (Leibniz-Regel) Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $k \in \mathbb{N}_0$ . Dann ist für alle  $C^k$ -Funktionen  $f, g: U \rightarrow \mathbb{R}$  auch  $fg: U \rightarrow \mathbb{R}$ ,  $x \mapsto f(x)g(x)$  eine  $C^k$ -Funktion. Es gilt

$$\frac{\partial^{|\alpha|}(fg)}{\partial x^\alpha} = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \frac{\partial^{|\beta|}f}{\partial x^\beta} \frac{\partial^{|\alpha-\beta|}g}{\partial x^{\alpha-\beta}}.$$

**Beispiel 11.31** Jede Polynomfunktion  $p: \mathbb{R}^n \rightarrow \mathbb{R}$  ist  $C^k$  für jedes  $k \in \mathbb{N}$  und somit eine  $C^\infty$ -Funktion.

Dies ist klar, wenn  $p$  die Nullfunktion ist. Andernfalls hat  $p$  einen Grad  $m > 0$  und es ist

$$p(x) = \sum_{|\alpha| \leq m} a_\alpha x^\alpha$$

mit gewissen Koeffizienten  $a_\alpha \in \mathbb{R}$ ; summiert wird über alle Multiindizes  $\alpha \in (\mathbb{N}_0)^n$  mit  $|\alpha| \leq m$ . Nach 11.29 wird  $p$  eine  $C^\infty$ -Funktion sein, wenn wir dies für jedes Monom

$$\mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto x^\alpha$$

zeigen können. Dies erfolgt per Induktion nach  $|\alpha|$ . Die Behauptung ist klar, wenn  $|\alpha| = 0$  oder  $|\alpha| = 1$ ; dann ist das Monom konstant bzw. linear und somit  $C^\infty$  nach Beispiel 11.24 bzw. Beispiel 11.25. Ist  $|\alpha| \geq 2$ , so wähle  $j \in \{1, \dots, n\}$  mit  $\alpha_j \geq 1$ . Da  $x^{\alpha - e_j}$  per Induktionsvoraussetzung schon  $C^\infty$  ist, ist

$$x^\alpha = x_j x^{\alpha - e_j}$$

$C^\infty$  nach 11.30.

**11.32** Wir zeigen zunächst, dass  $fg$  eine  $C^k$ -Funktion ist, per Induktion nach  $k$ . Ist  $k = 0$ , so ist  $fg$  stetig, also eine  $C^0$ -Funktion. Sei nun  $k \geq 1$  und gelte die Aussage bereits für  $k - 1$  an Stelle von  $k$ . da partielle Ableitungen nach  $x_j$  gebildet werden, indem alle anderen Variablen festgehalten werden, existiert nach der Produktregel der Mathematik 1 die partielle Ableitung  $\frac{\partial(fg)}{\partial x_j}$  und ist gegeben durch

$$\frac{\partial(fg)}{\partial x_j} = \left( \frac{\partial f}{\partial x_j} \right) g + f \left( \frac{\partial g}{\partial x_j} \right). \quad (68)$$

Da diese Funktion stetig ist, ist  $fg$  eine  $C^1$ -Funktion. Weiter sind  $\frac{\partial f}{\partial x_j}$ ,  $g$ ,  $f$  und  $\frac{\partial g}{\partial x_j}$  alles  $C^{k-1}$ -Funktionen; per Induktionsvoraussetzung sind daher beide Summanden in 68  $C^{k-1}$ -Funktionen und somit auch die Summe,  $\frac{\partial(fg)}{\partial x_j}$ , nach 11.29. Also ist  $fg$  eine  $C^k$ -Funktion, nach 11.22. Gegeben  $x \in U$  ist

$$x \in ]a_1, b_1[ \times \dots \times ]a_n, b_n[ \subseteq U$$

mit geeigneten  $a_j < b_j$  (siehe oben). Halten wir  $(x_1, \dots, x_{n-1})$  fest, liefert die Leibnizregel der Mathematik 1 für  $x_n \in ]a_n, b_n[$

$$\frac{\partial^{\alpha_n}(fg)}{\partial x_n^{\alpha_n}} = \sum_{\beta_n=0}^{\alpha_n} \binom{\alpha_n}{\beta_n} \frac{\partial^{\beta_n} f}{\partial x_n^{\beta_n}} \frac{\partial^{\alpha_n - \beta_n} g}{\partial x_n^{\alpha_n - \beta_n}}.$$

Fahren wir nun ebenso mit  $x_{n-1}, \dots, x_1$  fort, erhalten wir schließlich

$$\begin{aligned} \frac{\partial^{|\alpha|}(fg)}{\partial x^\alpha} &= \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}(fg) \\ &= \sum_{\beta_1=0}^{\alpha_1} \cdots \sum_{\beta_n=0}^{\alpha_n} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}} f \frac{\partial^{\alpha_1-\beta_1}}{\partial x_1^{\alpha_1-\beta_1}} \frac{\partial^{\alpha_n-\beta_n}}{\partial x_n^{\alpha_n-\beta_n}} g \\ &= \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \frac{\partial^{|\beta|} f}{\partial x^\beta} \frac{\partial^{|\alpha-\beta|} g}{\partial x^{\alpha-\beta}}. \end{aligned}$$

Wir beenden das Kapitel mit einer weiteren Fassung der Kettenregel.

**11.33** *Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge,  $I \subseteq \mathbb{R}$  ein nicht entartetes Intervall,  $k \in \mathbb{N}_0 \cup \{\infty\}$ ,  $f: U \rightarrow \mathbb{R}$  eine  $C^k$ -Funktion und  $\gamma = (\gamma_1, \dots, \gamma_n): I \rightarrow \mathbb{R}^n$  eine  $C^k$ -Kurve mit  $\gamma(I) \subseteq U$ . Dann ist*

$$f \circ \gamma: I \rightarrow \mathbb{R}, \quad t \mapsto f(\gamma(t))$$

eine  $C^k$ -Funktion.

Es genügt,  $k \in \mathbb{N}_0$  zu betrachten. Der Beweis ist per Induktion nach  $k$ . Ist  $k = 0$ , so ist  $f \circ \gamma$  stetig und somit  $C^0$ .

Sein nun  $k \geq 1$  und gelte die Aussage bereits für  $k - 1$  an Stelle von  $k$ . Nach 11.15 ist dann  $f \circ \gamma$  eine  $C^1$ -Funktion und

$$(f \circ \gamma)'(t) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\gamma(t)) \gamma_j'(t)$$

für alle  $t \in I$ , also

$$(f \circ \gamma)' = \sum_{j=1}^n \left( \frac{\partial f}{\partial x_j} \circ \gamma \right) \gamma_j'.$$

Für jedes  $j \in \{1, \dots, n\}$  sind  $\frac{\partial f}{\partial x_j}$  und  $\gamma$  beide  $C^{k-1}$ -Funktionen. Per Induktionsvoraussetzung ist  $\frac{\partial f}{\partial x_j} \circ \gamma$  also eine  $C^{k-1}$ -Funktion. Da auch  $\gamma_j'$  eine  $C^{k-1}$ -Funktion ist, ist  $\left( \frac{\partial f}{\partial x_j} \circ \gamma \right) \gamma_j'$  eine  $C^{k-1}$ -Funktion nach 11.30 und somit

$$(f \circ \gamma)' = \sum_{j=1}^n \left( \frac{\partial f}{\partial x_j} \circ \gamma \right) \gamma_j'$$

eine  $C^{k-1}$ -Funktion, nach 11.29. Nach 11.22 ist also  $f \circ \gamma$  eine  $C^k$ -Funktion.

## 12 Potentialfunktionen

Ist  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge und  $\phi: U \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion, so ist der Gradient

$$\nabla\phi: U \rightarrow \mathbb{R}^n, \quad x \mapsto \left( \frac{\partial\phi}{\partial x_1}, \dots, \frac{\partial\phi}{\partial x_n} \right)$$

ein stetiges Vektorfeld auf  $U$ . Schreiben wir kürzer  $D_j\phi := \frac{\partial\phi}{\partial x_j}$  für  $j \in \{1, \dots, n\}$ , so ist also

$$\nabla\phi = (D_1\phi, \dots, D_n\phi).$$

**Beispiel 12.1** Für  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x_1, x_2) \mapsto \frac{1}{4}(x_1^2 + x_2^2)$  ist

$$\nabla\phi(x_1, x_2) = \left( \frac{1}{2}x_1, \frac{1}{2}x_2 \right).$$

**Definition 12.2** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge und  $F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$  ein stetiges Vektorfeld. Eine  $C^1$ -Funktion  $\phi: U \rightarrow \mathbb{R}$  wird eine *Potentialfunktion* für  $F$  genannt, wenn

$$F = \nabla\phi,$$

also  $F_j = \frac{\partial\phi}{\partial x_j}$  für alle  $j \in \{1, \dots, n\}$ .

Wozu sind Potentialfunktionen gut? Zum Beispiel ermöglichen sie uns, Wegintegrale 2. Art sehr einfach auszurechnen:

**12.3** Ist  $U \subseteq \mathbb{R}^n$  offen und  $\phi: U \rightarrow \mathbb{R}$  eine Potentialfunktion für das stetige Vektorfeld  $F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$ , so gilt

$$\int_{\gamma} F \cdot ds = \phi(\gamma(\mathbf{b})) - \phi(\gamma(\mathbf{a})) \tag{69}$$

für jeden  $C^1$ -Weg  $\gamma: [a, b] \rightarrow U$ .

In der Tat ist

$$\begin{aligned} \int_{\gamma} F \cdot ds &= \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt = \sum_{j=1}^n F_j(\gamma(t)) \gamma'_j(t) dt \\ &= \int_a^b \sum_{j=1}^n \frac{\partial\phi}{\partial x_j}(\gamma(t)) \gamma'_j(t) dt \\ &= \int_a^b (\phi \circ \gamma)'(t) dt = [\phi(\gamma(t))]_a^b = \phi(\gamma(b)) - \phi(\gamma(a)); \end{aligned}$$

hierbei wurde beim Übergang zur letzten Zeile die Kettenregel in der Form 11.15 benutzt und dann der Hauptsatz der Integral- und Differentialrechnung.

**Beispiel 12.4** Für das Vektorfeld  $F = \nabla\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  aus Beispiel 12.1 und

$$\gamma: [0, \pi] \rightarrow \mathbb{R}^2, \quad (\cos(t), \sin(t))$$

ist

$$\int_{\gamma} F \cdot ds = \phi(1, 0) - \phi(-1, 0) = \frac{1}{4} - \frac{1}{4} = 0.$$

**Bemerkung 12.5** Formel (69) gilt ebenso, wenn  $\gamma: [a, b] \rightarrow U$  lediglich *stückweise* stetig differenzierbar ist.

Sei nämlich  $a = t_0 < t_1 < \dots < t_m = b$  eine Unterteilung des Intervalls  $[a, b]$  derart, dass der Teilweg  $\gamma_j := \gamma|_{[t_{j-1}, t_j]}$  stetig differenzierbar ist für all  $j \in \{1, \dots, m\}$ . Dann ist

$$\begin{aligned} \int_{\gamma} F \cdot ds &= \sum_{j=1}^m \int_{\gamma_j} F \cdot ds = \sum_{j=1}^m (\phi(\gamma(t_j)) - \phi(\gamma(t_{j-1}))) \\ &= \phi(\gamma(t_m)) - \phi(\gamma(t_0)) = \phi(\gamma(b)) - \phi(\gamma(a)); \end{aligned}$$

beim Übergang zur letzten Zeile wurde benutzt, dass eine “Teleskopsumme” vorliegt, in der Summanden mit positivem und negativen Vorzeichen vorkommen und sich auslöschen.

**Bemerkung 12.6** Hat ein Vektorfeld  $F$  eine Potentialfunktion  $\phi$ , so hängen Wegintegrale  $\int_{\gamma} F \cdot ds$  nach (69) nur vom Anfangs- und Endpunkt des Weges ab, nicht vom Verlauf des Weges dazwischen. Insbesondere ist für jeden (stückweise) stetig differenzierbaren *geschlossenen* Weg<sup>17</sup>  $\gamma: [a, b] \rightarrow U$  also

$$\int_{\gamma} F \cdot ds = 0.$$

Anwender schreiben für Wegintegrale über geschlossene Integrationswege gern  $\oint_{\gamma} F \cdot ds$ . In dieser Notation haben wir also stets  $\oint_{\gamma} F \cdot ds = 0$ , wenn  $F$  eine Potentialfunktion besitzt.

---

<sup>17</sup>Ein Weg  $\gamma: [a, b] \rightarrow U$  heißt geschlossen, wenn sein Anfangs- und Endpunkt übereinstimmen, also  $\gamma(a) = \gamma(b)$ .

**Beispiel 12.7** Ist  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein Kraftfeld und  $F(x_1, x_2, x_3)$  die Kraft, die auf einen Körper an der Stelle  $(x_1, x_2, x_3)$  wirkt, so ist die beim Bewegen des Körpers von  $\gamma(a) \in \mathbb{R}^3$  nach  $\gamma(b) \in \mathbb{R}^3$  längs eines stückweise stetig differenzierbaren Weges  $\gamma: [a, b] \rightarrow \mathbb{R}^3$

$$W = - \int_{\gamma} F \cdot ds.$$

Besitzt  $F$  eine Potentialfunktion  $\phi$ , so hängt die geleistete Arbeit also nur vom Anfangspunkt und Endpunkt, nicht vom Weg zwischen den Endpositionen ab.

Wann hat ein Vektorfeld eine Potentialfunktion? Eine einfache notwendige Bedingung gibt es für stetig differenzierbare Vektorfelder.

**Definition 12.8** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge. Eine Funktion  $F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$  heißt *stetig differenzierbares Vektorfeld*, wenn  $F_1, \dots, F_n: U \rightarrow \mathbb{R}$  stetig differenzierbare Funktionen sind.

**12.9** (Integrabilitätsbedingung). Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge und  $F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Hat  $F$  eine Potentialfunktion, so gilt

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial F_k}{\partial x_j} \quad \text{für alle } j, k \in \{1, \dots, n\}.$$

In der Tat: Ist  $F = \nabla\phi$  mit einer  $C^1$ -Funktion  $\phi$ , so ist  $\frac{\partial\phi}{\partial x_j} = F_j$  stetig diff'bar für alle  $j \in \{1, \dots, n\}$  und somit  $\phi$  eine  $C^2$ -Funktion. Nach dem Satz von Schwarz gilt somit

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} \phi = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} \phi = \frac{\partial F_k}{\partial x_j}.$$

**Beispiel 12.10** Hat das Vektorfeld  $F = (F_1, F_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $(x_1, x_2) \mapsto (-x_2, x_1)$  eine Potentialfunktion?

Das Vektorfeld ist stetig differenzierbar. Da

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial}{\partial x_2}(-x_2) = -1$$

nicht mit

$$\frac{\partial F_2}{\partial x_1} = \frac{\partial}{\partial x_1}(x_1) = 1$$

übereinstimmt, ist die Integrabilitätsbedingung verletzt. Also kann  $F$  keine Potentialfunktion haben.

**Beispiel 12.11** Hat das Vektorfeld

$$F = (F_1, F_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (\sin y, 1 + x \cos y)$$

eine Potentialfunktion?

Das Vektorfeld ist stetig differenzierbar. Da

$$\frac{\partial F_1}{\partial y} = \cos(y) = \frac{\partial F_2}{\partial x},$$

ist die Integrabilitätsbedingung erfüllt. Es spricht also nichts gegen die Existenz einer Potentialfunktion und in der Tat können wir eine solche konstruieren mit der *Methode des sukzessiven Integrierens*.

Wir starten hierzu mit einer  $C^1$ -Funktion  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Genau dann gilt

$$\frac{\partial \phi}{\partial x}(x, y) = F_1(x, y) = \sin y$$

für alle  $(x, y) \in \mathbb{R}^2$ , wenn für jedes feste  $y \in \mathbb{R}$

$$\frac{\partial \phi}{\partial x}(x, y) = \sin y$$

erfüllt ist, also

$$\phi(x, y) = x \sin y + C(y) \tag{70}$$

ist mit einem  $C(y) \in \mathbb{R}$ . Hierbei wurde  $y$  festgehalten und auf die so erhaltene Funktion von  $x$  der Hauptsatz der Differential- und Integralrechnung angewandt. Dann ist

$$C(y) = \phi(x, y) - x \sin y$$

eine  $C^1$ -Funktion. Aus (70) folgt

$$\frac{\partial \phi}{\partial y}(x, y) = x \cos y + C'(y).$$

Also gilt

$$\frac{\partial \phi}{\partial y}(x, y) = F_2(x, y) = 1 + x \cos y$$

genau dann, wenn

$$C'(y) = 1,$$

also  $C(y) = y + K$  mit einer Konstanten  $K \in \mathbb{R}$ . Also liefert

$$\phi(x, y) = x \sin y + y + K$$

eine Potentialfunktion für  $F$ .

Das beschriebene Vorgehen funktioniert stets für  $C^1$ -Vektorfelder auf  $\mathbb{R}^n$  oder auf einem offenen Quader  $]a_1, b_1[ \times \cdots \times ]a_n, b_n[$ , welche die Integrabilitätsbedingung erfüllen. Somit:

**12.12** *Ein stetig differenzierbares Vektorfeld*

$$F = (F_1, \dots, F_n): ]a_1, b_1[ \times \cdots \times ]a_n, b_n[ \rightarrow \mathbb{R}^n$$

hat genau dann eine Potentialfunktion, wenn  $F$  die Integrabilitätsbedingung erfüllt.

Im Falle von Vektorfeldern auf  $U \subseteq \mathbb{R}^3$  gibt es weitere nützliche Notationen.

**Definition 12.13** Es sei  $U \subseteq \mathbb{R}^3$  eine offene Teilmenge und

$$F = (F_1, F_2, F_3): U \rightarrow \mathbb{R}^3$$

ein stetig differenzierbares Vektorfeld. Wir definieren die *Rotation* von  $F$  als das stetige Vektorfeld

$$\operatorname{rot} F := \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}.$$

Eine andere Notation für  $\operatorname{rot}(F)$  ist  $\nabla \times F$ , denn formal gerechnet liefert das Vektorprodukt

$$\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$$

die Rotation (wenn man den Operator immer vor die Funktion stellt). Im Englischen nennt man die Rotation "curl" und schreibt auch  $\operatorname{curl}(F)$ .

**Beispiel 12.14** Für das Vektorfeld  $F = (F_1, F_2, F_3): \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,

$$(x_1, x_2, x_3) \mapsto (x_2, x_1 x_3, \cos(x_1))$$

ist mit  $x := (x_1, x_2, x_3)$

$$\operatorname{rot} F(x) = \begin{pmatrix} 0 - x_1 \\ 0 + \sin(x_1) \\ x_3 - 1 \end{pmatrix}.$$

**12.15** Ein  $C^1$ -Vektorfeld  $F = (F_1, F_2, F_3): U \rightarrow \mathbb{R}^3$  erfüllt offenbar genau dann die Integrabilitätsbedingung aus 12.9, wenn  $\operatorname{rot} F = 0$ .

**12.16** Ein  $C^1$ -Vektorfeld  $F = (F_1, F_2, F_3): U \rightarrow \mathbb{R}^3$  auf einer offenen Teilmenge  $U \subseteq \mathbb{R}^3$  heißt *wirbelfrei*, wenn  $\operatorname{rot} F = 0$ .

**Beispiel 12.17** Das Gravitationsfeld  $F = (F_1, F_2, F_3): \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^3$ ,

$$x = (x_1, x_2, x_3) \mapsto -\frac{x}{|x|^3} = -\frac{1}{(x_1^2 + x_2^2 + x_3^2)^{\frac{3}{2}}}(x_1, x_2, x_3)$$

ist wirbelfrei (Übung).

**Definition 12.18** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge und

$$F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$$

ein stetig differenzierbares Vektorfeld. Die Funktion

$$\operatorname{div}(F) := \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_n}{\partial x_n}$$

wird die *Divergenz von F* genannt; diese ist also eine stetige Funktion  $U \rightarrow \mathbb{R}$ . Eine andere Notation ist  $\nabla \cdot F := \operatorname{div}(F)$ . Ein stetig differenzierbares Vektorfeld  $F: U \rightarrow \mathbb{R}^n$  heißt *quellenfrei*, wenn  $\operatorname{div}(F) = 0$ .

**Beispiel 12.19** (a) Für das stetig differenzierbare Vektorfeld

$$F = (F_1, F_2, F_3): \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2, x_3)$$

berechnen wir

$$\operatorname{div}(F) = \frac{\partial}{\partial x_1} x_1 + \frac{\partial}{\partial x_2} x_2 + \frac{\partial}{\partial x_3} x_3 = 1 + 1 + 1 = 3.$$

Insbesondere ist  $F$  nicht quellenfrei (anschaulich sieht man dies auch, wenn man das Vektorfeld skizziert, siehe Vorlesung).

(b) Das Vektorfeld

$$F = (F_1, F_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x_1, x_2) \mapsto (-x_2, x_1)$$

aus Beispiel 12.10 ist quellenfrei, denn

$$\operatorname{div}(F) = \frac{\partial}{\partial x_1}(-x_2) + \frac{\partial}{\partial x_2}(x_1) = 0 + 0 = 0.$$

Die Thematik wird in der Mathematik 3 fortgesetzt. Insbesondere lernen wir dann den Gaußschen Integralsatz kennen und den Stokesschen Integralsatz. Es wird dann klar, was Wirbelfreiheit mit Wirbeln zu tun hat und Quellenfreiheit mit Quellen.

## 13 Differenzierbare vektorwertige Funktionen

**Definition 13.1** Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge. Eine Funktion

$$f = (f_1, \dots, f_m): U \rightarrow \mathbb{R}^m, \quad x = (x_1, \dots, x_n) \mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

wird *stetig differenzierbar* genannt (oder kurz: eine  $C^1$ -Funktion), wenn alle Komponenten  $f_1, \dots, f_m: U \rightarrow \mathbb{R}$  stetig differenzierbar sind. Wir definieren dann die Jacobi-Matrix  $J_f(x) \in \mathbb{R}^{m \times n}$  von  $f$  an der Stelle  $x \in U$  als

$$J_f(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Ist  $k \in \mathbb{N}_0 \cup \{\infty\}$  und jeder der Komponenten  $f_1, \dots, f_m$  eine  $C^k$ -Funktion, so wird  $f$  eine  $C^k$ -Funktion genannt.

**Beispiel 13.2** Die Funktion

$$f = (f_1, f_2): \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad x = (x_1, x_2, x_3) \mapsto \begin{pmatrix} x_1 x_3 \\ x_1 e^{x_2 x_3} \end{pmatrix}$$

von drei reellen Variablen ist stetig differenzierbar, da ihre zwei Komponenten es sind. Die Jacobi-Matrix an einer Stelle  $x$  lautet

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \frac{\partial f_1}{\partial x_3}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \frac{\partial f_2}{\partial x_3}(x) \end{pmatrix} = \begin{pmatrix} x_3 & 0 & x_1 \\ e^{x_2 x_3} & x_1 x_3 e^{x_2 x_3} & x_1 x_2 e^{x_2 x_3} \end{pmatrix}.$$

Schauen wir uns einige Spezialfälle systematisch an.

**Beispiel 13.3** (Der Fall  $n = 1$ ). Ist  $I \subseteq \mathbb{R}$  ein offenes Intervall und  $\gamma = (\gamma_1, \dots, \gamma_m): U \rightarrow \mathbb{R}^m$  eine  $C^1$ -Kurve, so ist die Jacobi-Matrix  $J_\gamma(x) \in \mathbb{R}^{m \times 1} = \mathbb{R}^m$  der Spaltenvektor

$$J_\gamma(x) = \gamma'(x)$$

mit  $\gamma'(x) = (\gamma'_1(x), \dots, \gamma'_m(x))^T$ .

**Beispiel 13.4** (Der Fall  $m = 1$ ). Ist  $U \subseteq \mathbb{R}^n$  offen und  $f: U \rightarrow \mathbb{R}$  eine reellwertige  $C^1$ -Funktion, so ist die Jacobi-Matrix

$$J_f(x) = \left( \frac{\partial f}{\partial x_1}(x) \cdots \frac{\partial f}{\partial x_n}(x) \right) = (\nabla f(x))^T$$

die zum Gradienten transponierte Zeilenmatrix.

**Beispiel 13.5** (Der Fall  $n = m$ ). Stetig differenzierbare Vektorfelder

$$F = (F_1, \dots, F_n): U \rightarrow \mathbb{R}^n$$

auf einer offenen Menge  $U \subseteq \mathbb{R}^n$  haben wir bereits im vorigen Kapitel kennengelernt.

**Beispiel 13.6** Jede konstante Funktion  $f = (f_1, \dots, f_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $x \mapsto c$  mit  $c = (c_1, \dots, c_m) \in \mathbb{R}^m$  ist stetig differenzierbar, mit Jacobi-Matrix  $J_f(x) = 0 \in \mathbb{R}^{m \times n}$  für alle  $x \in \mathbb{R}^n$ .

Denn wir wissen, dass  $\frac{\partial f_i}{\partial x_j} = \frac{\partial}{\partial x_j} c_i = 0$  für alle  $i \in \{1, \dots, m\}$  und  $j \in \{1, \dots, n\}$ .

**Beispiel 13.7** Es sei  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  eine  $m \times n$ -Matrix und

$$\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto Ax$$

die zugehörige lineare Abbildung. Dann ist  $\phi_A$  stetig differenzierbar und für jedes  $x \in \mathbb{R}^n$  ist die Jacobi-Matrix

$$J_{\phi_A}(x) = A$$

gleich der gegebenen Matrix.

Schreiben wir  $f = (f_1, \dots, f_m) := \phi_A$ , so ist nämlich

$$f_i(x) = \sum_{k=1}^n a_{ik} x_k$$

für  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , also  $f_i$  ein Polynom vom Grad  $\leq 1$  und somit  $f_i$  stetig differenzierbar, für jedes  $i \in \{1, \dots, m\}$ . Weiter ist

$$\frac{\partial f_i}{\partial x_j}(x) = \sum_{k=1}^n a_{ik} \underbrace{\frac{\partial}{\partial x_j}(x_k)}_{=\delta_{jk}} = a_{ij}$$

für alle  $j \in \{1, \dots, n\}$ . Somit ist  $J_f(x) = \left( \frac{\partial f_i}{\partial x_j}(x) \right) = (a_{ij}) = A$ .

**13.8** (Kettenregel) *Es seien  $U \subseteq \mathbb{R}^n$  und  $V \subseteq \mathbb{R}^m$  offene Teilmengen,  $g = (g_1, \dots, g_m): U \rightarrow \mathbb{R}^m$ ,  $x = (x_1, \dots, x_n) \mapsto g(x)$  eine  $C^1$ -Funktion mit  $g(U) \subseteq V$  und  $f = (f_1, \dots, f_\ell): V \rightarrow \mathbb{R}^\ell$ ,  $y = (y_1, \dots, y_m) \mapsto f(y)$  eine  $C^1$ -Funktion. Dann ist  $f \circ g: U \rightarrow \mathbb{R}^\ell$ ,  $x \mapsto f(g(x))$  stetig differenzierbar mit Jacobi-Matrix*

$$J_{f \circ g}(x) = J_f(g(x))J_g(x)$$

*an der Stelle  $x \in U$ . Sind  $f$  und  $g$  beides  $C^k$ -Abbildungen, so ist auch  $f \circ g$  eine  $C^k$ -Abbildung.*

Zum Nachweis betrachten wir für  $i \in \{1, \dots, \ell\}$  die  $i$ te Komponente  $f_i \circ g$  von  $f \circ g$ . Um die partielle Ableitung von  $f_i \circ g$  nach der Variablen  $x_k$  an der Stelle  $x \in U$  zu berechnen, betrachten wir für ein  $\varepsilon > 0$  die Hilfsfunktion

$$\gamma = (\gamma_1, \dots, \gamma_m): ]-\varepsilon, \varepsilon[ \rightarrow U, \quad t \mapsto g(x + te_k)$$

mit  $\gamma_j(t) = g_j(x + te_k)$  und somit  $\gamma'_j(0) = \frac{\partial g_j}{\partial x_k}(x)$ . Nach 11.15 ist  $f_i \circ \gamma$  stetig differenzierbar mit Ableitung

$$(f_i \circ \gamma)'(0) = \sum_{j=1}^m \frac{\partial f_i}{\partial y_j}(\gamma(0))\gamma'_j(0) = \sum_{j=1}^m \frac{\partial f_i}{\partial y_j}(g(x)) \frac{\partial g_j}{\partial x_k}(x).$$

also existiert  $\frac{\partial(f_i \circ g)}{\partial x_k}(x)$  und ist gleich

$$\sum_{j=1}^m \frac{\partial f_i}{\partial y_j}(g(x)) \frac{\partial g_j}{\partial x_k}(x).$$

Diese Zahl stimmt mit der  $i$ ten Komponente von  $J_f(g(x))J_g(x)$  überein.

Sind  $f$  und  $g$  beides  $C^k$ -Funktionen, so ist nach dem Vorigen  $f_i \circ g$  stetig differenzierbar für alle  $i \in \{1, \dots, \ell\}$ . Da

$$\frac{\partial(f_i \circ g)}{\partial x_k} = \sum_{j=1}^m \left( \frac{\partial f_i}{\partial y_j} \circ g \right) \frac{\partial g_j}{\partial x_k}$$

per Induktionsvoraussetzung und 11.30 eine  $C^{k-1}$ -Funktion ist, ist  $f_i \circ g$  eine  $C^k$ -Funktion und somit  $f \circ g$  eine  $C^k$ -Funktion.

**13.9** (Mittelwertsatz in Integralform) *Es sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $f = (f_1, \dots, f_m): U \rightarrow \mathbb{R}^m$  eine  $C^1$ -Funktion. Sind  $x, y \in U$  mit  $x+t(y-x) \in U$  für alle  $t \in [0, 1]$ , so ist*

$$f(y) - f(x) = \int_0^1 J_f(x + t(y-x))(y-x) dt.$$

Es genügt zu zeigen, dass auf beiden Seiten die  $i$ ten Komponenten gleich sind für alle  $i \in \{1, \dots, m\}$ , also  $f_i(y) - f_i(x) = \int_0^1 J_{f_i}(x + t(y-x))(y-x) dt$ . Wir dürfen daher annehmen, dass  $m = 1$  ist. Wir betrachten also eine  $C^1$ -Funktion

$$f: U \rightarrow \mathbb{R}.$$

Die Hilfsfunktion  $h: [0, 1] \rightarrow \mathbb{R}$ ,  $t \mapsto f(x+t(y-x))$  ist an jeder Stelle  $t \in [0, 1]$  differenzierbar mit

$$h'(t) = D_{y-x}f(x+t(y-x)) = \nabla f(x+t(y-x)) \cdot (y-x) = J_f(x+t(y-x))(y-x).$$

Nach dem Hauptsatz der Integral- und Differentialrechnung ist

$$f(y) - f(x) = h(1) - h(0) = \int_0^1 h'(t) dt = \int_0^1 J_f(x + t(y-x))(y-x) dt.$$

## 14 Lösen nicht-linearer Gleichungen

Zur Motivation betrachten wir für  $c \in \mathbb{R}$  die "Höhenlinie"

$$H_c := \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$$

der Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2$ . Wir stellen fest:

- Für  $c > 0$  ist  $H_c$  der Kreis  $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = c\}$  vom Radius  $\sqrt{c}$  um den Ursprung. Dieser ist lokal um jede Stelle  $(x_0, y_0) \in H_c$  ein Graph einer  $C^1$ -Funktion  $\phi(x)$  von  $x$  (nämlich  $\phi(x) = \sqrt{c - x^2}$  oder  $\phi(x) = -\sqrt{c - x^2}$ ) oder einer Funktion von  $y$ . Skizzen wurden in der Vorlesung gegeben.
- Für  $c = 0$  ist  $H_0 = \{(0, 0)\}$  eine einpunktige Menge
- Für  $c < 0$  ist  $H_c = \emptyset$ .

An der entarteten Stelle  $(0, 0)$  ist  $\nabla f(0, 0) = (0, 0)$ , also sowohl  $\frac{\partial f}{\partial x}(0, 0) = 0$  als auch  $\frac{\partial f}{\partial y}(0, 0) = 0$ . Dies ist kein Zufall.

**14.1** (Satz über implizite Funktionen). *Es sei  $U \subseteq \mathbb{R}^2$  eine offene Menge und  $f: U \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion. Es sei  $(x_0, y_0) \in U$  ein Punkt derart, dass*

$$f(x_0, y_0) = 0$$

und

$$\frac{\partial f}{\partial y}(x_0, y_0) \neq 0.$$

*Dann gibt es eine offene Menge  $V \subseteq \mathbb{R}$  mit  $x_0 \in V$  und eine offene Menge  $W \subseteq \mathbb{R}$  mit  $y_0 \in W$  derart, dass  $V \times W \subseteq U$  und*

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \text{graph}(\phi) = \{(x, \phi(x)) : x \in V\}$$

*für eine  $C^1$ -Funktion  $\phi: V \rightarrow W$ .*

**Bemerkung 14.2** Man beachte, dass für  $c \in \mathbb{R}$

$$\{(x, y) \in U : f(x, y) = c\} = \{(x, y) \in U : f(x, y) - c = 0\}.$$

Indem man  $f - c$  statt  $f$  betrachtet, braucht man statt Höhenlinien also nur Nullstellenmengen diskutieren.

**Beispiel 14.3** Im motivierenden Beispiel sei  $c > 0$  und  $(x_0, y_0) \in H_c$ , also  $(x_0, y_0)$  eine Nullstelle von  $g: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto f(x, y) - c$ . Ist  $y_0 \neq 0$ , so ist

$$\frac{\partial g}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 2y_0 \neq 0$$

und somit nach dem Satz über implizite Funktionen nahe  $(x_0, y_0)$  die Menge  $H_c$  der Graph einer Funktion von  $x$  (nämlich  $\sqrt{c - x^2}$  wenn  $y_0 > 0$ , sonst  $-\sqrt{c - x^2}$ ). Ist  $x_0 \neq 0$ , können wir den Satz mit vertauschten Rollen von  $x$  und  $y$  anwenden und  $H_c$  nahe  $(x_0, y_0)$  als Graph  $\{(\phi(y), y) : y \in V\}$  einer Funktion von  $y$  schreiben.

**Beispiel 14.4** *Die Gleichung*

$$x e^y + y e^x = 0$$

ist erfüllt für  $(x, y) = (0, 0)$ . Zeigen Sie, dass für alle  $x \in \mathbb{R}$  nahe 0 eine Lösung der Form  $(x, \phi(x))$  existiert.

Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x e^y + y e^x$  ist stetig differenzierbar. Es ist  $f(x_0, y_0) = 0$  mit  $(x_0, y_0) := (0, 0)$  und

$$\frac{\partial f}{\partial y}(x, y) = x e^y + e^x$$

für alle  $(x, y) \in \mathbb{R}^2$ , also

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial f}{\partial y}(0, 0) = 1 \neq 0.$$

Nach dem Satz über implizite Funktionen gibt es offene Mengen  $V, W \subseteq \mathbb{R}$  mit  $0 \in V$  und  $0 \in W$  und eine  $C^1$ -Funktion  $\phi: V \rightarrow W$  derart, dass

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \{(x, \phi(x)) : x \in V\}.$$

**Bemerkung 14.5** *In der Situation von 14.1 gilt*

$$\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))}$$

für alle  $x \in V$  mit  $\frac{\partial f}{\partial y}(x, \phi(x)) \neq 0$

(was aus Stetigkeitsgründen für alle  $x \in V$  nahe  $x_0$  erfüllt ist).

Anwendung der Kettenregel auf

$$f(x, \phi(x)) = 0$$

liefert nämlich

$$\frac{\partial f}{\partial x}(x, \phi(x)) + \frac{\partial f}{\partial y}(x, \phi(x))\phi'(x) = 0.$$

Ist  $\frac{\partial f}{\partial y}(x, \phi(x)) \neq 0$ , können wir nach  $\phi'(x)$  auflösen.

**Beispiel 14.6** In Beispiel 14.4 haben wir  $\frac{\partial f}{\partial y}(0, 0) = 1$  (siehe oben) und

$$\frac{\partial f}{\partial x}(x, y) = e^y + y e^x,$$

also  $\frac{\partial f}{\partial x}(0, 0) = 1$ . Mit Bemerkung 14.5 folgt

$$\phi'(0) = -\frac{\frac{\partial f}{\partial x}(0, 0)}{\frac{\partial f}{\partial y}(0, 0)} = -1.$$

Wir können also  $\phi'(0)$  berechnen, obwohl wir  $\phi$  nicht in geschlossener Form hinschreiben können!

Wir erwähnen noch eine allgemeinere Fassung des Satzes über implizite Funktionen.

**14.7** (Satz über implizite Funktionen). *Es sei  $U \subseteq \mathbb{R}^n$  mit  $n \geq 2$  eine offene Menge und  $f: U \rightarrow \mathbb{R}$ ,  $x = (x_1, \dots, x_n) \mapsto f(x)$  eine  $C^1$ -Funktion. Es sei  $(x_0, y_0) \in U$  mit  $x_0 \in \mathbb{R}^{n-1}$  und  $y_0 \in \mathbb{R}$  ein Punkt derart, dass*

$$f(x_0, y_0) = 0$$

und

$$\frac{\partial f}{\partial x_n}(x_0, y_0) \neq 0.$$

*Dann gibt es eine offene Menge  $V \subseteq \mathbb{R}^{n-1}$  mit  $x_0 \in V$  und eine offene Menge  $W \subseteq \mathbb{R}$  mit  $y_0 \in W$  derart, dass  $V \times W \subseteq U$  und*

$$\{(x, y) \in V \times W : f(x, y) = 0\} = \text{graph}(\phi) = \{(x, \phi(x)) : x \in V\}$$

*für eine  $C^1$ -Funktion  $\phi: V \rightarrow W$ .*

**Bemerkung 14.8** Es gibt auch Versionen des Satzes für vektorwertige Funktionen (wir verweisen auf die Literatur).

**Definition 14.9** Eine Abbildung  $\psi: U \rightarrow V$  zwischen offenen Teilmengen  $U \subseteq \mathbb{R}^n$  und  $V \subseteq \mathbb{R}^n$  heißt  *$C^1$ -Diffeomorphismus*, wenn  $\psi$  bijektiv ist und sowohl  $\psi$  als auch  $\psi^{-1}: V \rightarrow U$  stetig differenzierbar sind.

**14.10** Ist  $\psi: U \rightarrow V$  ein  $C^1$ -Diffeomorphismus zwischen offenen Teilmengen  $U$  und  $V$  in  $\mathbb{R}^n$ , so ist für jedes  $x \in U$  die Jacobi-Matrix  $J_\psi(x) \in \mathbb{R}^{n \times n}$  invertierbar. Für jedes  $x \in U$  gilt weiter

$$J_{\psi^{-1}}(\psi(x)) = (J_\psi(x))^{-1}.$$

Die zur Einheitsmatrix  $\mathbf{1} \in \mathbb{R}^{n \times n}$  gehörige lineare Abbildung ist  $\phi_{\mathbf{1}}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x \mapsto \mathbf{1}x = x$ , also  $\phi_{\mathbf{1}} = \text{id}_{\mathbb{R}^n}$ . Wir schließen, dass

$$J_{\text{id}_U}(x) = J_{\phi_{\mathbf{1}}}(x) = \mathbf{1}$$

für alle  $x \in U$ . Anwendung der Kettenregel auf

$$\psi^{-1} \circ \psi = \text{id}_U$$

liefert also

$$J_{\psi^{-1}}(\psi(x))J_\psi(x) = J_{\text{id}_U}(x) = \mathbf{1}.$$

Analog ist  $J_\psi(x) \circ J_{\psi^{-1}}(\psi(x)) = J_{\text{id}_V}(\psi(x)) = \mathbf{1}$ , somit die Matrix  $J_\psi(x)$  invertierbar mit  $(J_\psi(x))^{-1} = J_{\psi^{-1}}(\psi(x))$ .

**14.11** (Satz über die Umkehrfunktion/Umkehrabbildung) *Es sei  $U \subseteq \mathbb{R}^n$  eine offene Teilmenge,  $f: U \rightarrow \mathbb{R}^n$  eine stetig differenzierbare Abbildung und  $x_0 \in U$  derart, dass die Jacobi-Matrix  $J_f(x_0) \in \mathbb{R}^{n \times n}$  invertierbar ist. Dann existiert eine offene Teilmenge  $V \subseteq U$  mit  $x_0 \in V$  derart, dass  $f(V) \subseteq \mathbb{R}^n$  offen und  $f|_V: V \rightarrow f(V)$  ein  $C^1$ -Diffeomorphismus ist.*

**Bemerkung 14.12** Dies ist ein Satz über das Auflösen von Gleichungen: Für alle  $y \in f(V)$  existiert genau ein  $x \in V$  derart, dass  $f(x) = y$  (nämlich  $x = (f|_V)^{-1}(y)$ ).

**Beispiel 14.13** Wir betrachten die Funktion

$$f = (f_1, f_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad x = (x_1, x_2) \mapsto (x_1 + \sin(x_1x_2), x_2 + e^{x_1x_2})$$

um die Stelle  $x_0 = (0, 0)$ . Es ist  $f(x_0) = (0, 1)$ . *Besitzt die Gleichung  $f(x_1, x_2) = (y_1, y_2)$  für  $(y_1, y_2) \in \mathbb{R}^2$  nahe  $(0, 1)$  immer eine Lösung  $(x_1, x_2)$ ?*

Antwort: Ja, dies folgt mit dem Satz über die Umkehrfunktion. Die Funktion  $f$  ist stetig differenzierbar und hat an der Stelle  $x = (x_1, x_2) \in \mathbb{R}^2$  die Jacobi-Matrix

$$J_f(x) = \begin{pmatrix} 1 + x_2 \cos(x_1 x_2) & x_1 \cos(x_1 x_2) \\ x_2 e^{x_1 x_2} & 1 + x_1 e^{x_1 x_2} \end{pmatrix}.$$

Es ist also

$$J_f(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

die Einheitsmatrix und somit eine invertierbare Matrix. Nach dem Satz über die Umkehrfunktion gibt es eine offene Teilmenge  $V \subseteq \mathbb{R}^2$  mit  $(0, 0) \in V$  derart, dass  $f(V)$  offen in  $\mathbb{R}^2$  ist und  $f|_V: V \rightarrow f(V)$  ein  $C^1$ -Diffeomorphismus. Dann ist  $(0, 1) = f(0, 0) \in f(V)$ . Für  $y \in \mathbb{R}^2$  nahe zu  $(0, 1)$  in dem Sinne, dass  $y \in f(V)$ , ist also  $y = f(x)$  für ein  $x \in V$  (nämlich für  $x := (f|_V)^{-1}(y)$ ).

**Bemerkung 14.14** Sowohl der Satz über implizite Funktionen als auch der Satz über die Umkehrfunktion sind Sätze über das Lösen nicht-linearer Gleichungen bzw. Gleichungssysteme.

- Im Falle des Satzes über implizite Funktionen haben wir eine Gleichung,  $f(x_1, \dots, x_{n-1}, y) = 0$ , für  $y$  als Funktion von  $(x_1, \dots, x_{n-1})$ .
- Im Falle des Satzes über die Umkehrfunktion haben wir ein System von  $n$  nicht-linearen Gleichungen für  $n$  Unbekannte  $(x_1, \dots, x_n)$ , nämlich

$$\begin{aligned} f_1(x_1, \dots, x_n) &= y_1 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= y_n \end{aligned}$$

zu gegebenen  $(y_1, \dots, y_n) \in \mathbb{R}^n$  nahe  $f(x_0)$ .

**14.15** Ist  $x \in \mathbb{R}^n$ , so nennt man eine offene Teilmenge  $W \subseteq \mathbb{R}^n$  mit  $x \in W$  übrigens auch eine *offene Umgebung* von  $x$ . Eine Menge  $V \subseteq \mathbb{R}^n$ , die eine offene Umgebung  $W$  von  $x$  enthält (also  $W \subseteq V$ ), nennt man eine *Umgebung* von  $x$ .

**Beispiel 14.16** Im Satz über die Umkehrfunktion ist  $V$  eine offene  $x_0$ -Umgebung und  $f(V)$  eine offene Umgebung von  $f(x_0)$ .

## 15 Extrema unter Nebenbedingungen

Sei  $U \subseteq \mathbb{R}^n$  eine offene Menge und  $f: U \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion. In der Praxis hat man oft  $f(x_1, \dots, x_n)$  zu maximieren oder minimieren, allerdings nicht für beliebige  $x = (x_1, \dots, x_n) \in U$ , sondern nur für diejenigen  $x \in U$ , die eine *Nebenbedingung*

$$g(x_1, \dots, x_n) = 0$$

erfüllen mit einer stetig differenzierbaren Funktion  $g: U \rightarrow \mathbb{R}$  (oder mehrere Nebenbedingungen,  $g_1(x_1, \dots, x_n) = \dots = g_m(x_1, \dots, x_n) = 0$ ).

**Beispiel 15.1** Maximiere  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto xy$  auf dem Einheitskreis  $\mathbb{S} \subseteq \mathbb{R}^2$ , also unter der Nebenbedingung  $x^2 + y^2 - 1 = 0$ .

In diesem Kapitel lernen wir Techniken kennen, mit denen sich solche Extremwertprobleme unter Nebenbedingungen behandeln lassen. Außerhalb der Optimierung kommen die Methoden zum Beispiel auch in der Theoretischen Mechanik zum Einsatz.

### Parametrisierungsmethode

Zum Beispiel im Falle von zwei Variablen: Wenn man eine Parameterdarstellung  $x = x(t)$ ,  $y = y(t)$ ,  $t \in I \subseteq \mathbb{R}$  der Kurve  $g(x, y) = 0$  finden kann, so bestimmt man für die Funktion

$$F: I \rightarrow \mathbb{R}, \quad t \mapsto f(x(t), y(t))$$

mit den (in Kapitel 1 wiederholten) Methoden der Mathematik 1 die Extremstellen.

Extremalität von  $F(t)$  ist notwendig dafür, dass  $f$  an der Stelle  $(x(t), y(t))$  extremal ist unter der Nebenbedingung  $g(x, y) = 0$ .

**Beispiel 15.2** Im obigen Beispiel 15.1 nimmt die stetige Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto xy$  auf der kompakten Menge  $\mathbb{S}$  ein Maximum an. Wählen wir die Parametrisierung  $\mathbb{R} \rightarrow \mathbb{S}$ ,  $t \mapsto (\cos(t), \sin(t))$  und definieren

$$F(t) := f(\cos(t), \sin(t)) = \cos(t) \sin(t),$$

so ist  $F(t) = \frac{1}{2} \sin(2t)$ . Wir könnten nun über Untersuchung von  $0 = F'(t) = \cos(2t)$  und  $F''$  die Maximalstellen von  $F$  finden. Aber wir wissen ja, wo der Sinus seine Maximalstellen hat;  $F(t)$  wird maximal für

$$2t = \frac{\pi}{2} + 2\pi k$$

mit  $k \in \mathbb{Z}$ , also  $t = \frac{\pi}{4} + \pi k$ . Dann ist  $F(\frac{\pi}{4} + \pi k) = \frac{1}{2} \sin(\pi/2 + 2\pi k) = \frac{1}{2}$ . Das Maximum der Funktion  $f(x, y) = xy$  unter der Nebenbedingung  $g(x, y) = x^2 + y^2 - 1 = 0$  ist also  $\frac{1}{2}$ . Das Maximum wird genau an den Stellen

$$(\cos(\pi/4), \sin(\pi/4)) = \left( \frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2} \right)$$

(mit  $k = 0$ ) und

$$(\cos(\pi/4 + \pi), \sin(\pi/4 + \pi)) = \left( -\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2} \right)$$

(mit  $k = 1$ ) angenommen.<sup>18</sup>

## Explizite Methode

Man kann  $f(x_1, \dots, x_n)$  leicht minimieren oder maximieren unter der Nebenbedingung  $g(x_1, \dots, x_n) = 0$ , wenn man die Bedingung  $g(x_1, \dots, x_n) = 0$  nach einer der Variablen auflösen kann, zum Beispiel nach  $x_n$ . Sei also

$$x_n = h(x_1, \dots, x_{n-1})$$

mit einer  $C^1$ -Funktion  $h$ . Wir haben dann die Funktion

$$f(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1}))$$

von  $n-1$  Variablen zu minimieren bzw. maximieren. Die Extremstellen kann man finden mit der Methode aus Kapitel 11.

**Beispiel 15.3** Sei  $T(x, y, z) = xz - zy$  für  $(x, y, z) \in \mathbb{R}^3$ . Frage: *Wo nimmt  $T$  auf der Sphäre*

$$K = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$$

---

<sup>18</sup>Andere Werte von  $k$  liefern die gleichen Punkte wegen der  $2\pi$ -Periodizität von Sinus und Cosinus.

ein globales Maximum an und wie groß ist dieses?

Lösung: Weil  $T(-x, -y, -z) = T(x, y, z)$ , genügt es, das globale Maximum auf der oberen Halbsphäre  $\{(x, y, z) \in K : z \geq 0\}$  zu finden. Für  $(x, y, z) = (x, y, 0)$  im Äquator wird das Maximum nicht angenommen, denn es ist  $T(x, y, 0) = 0$  aber zum Beispiel  $T(\frac{1}{2}\sqrt{2}, 0, \frac{1}{2}\sqrt{2}) = \frac{1}{2} > 0$ . Das Maximum wird also auf der Menge

$$M = \{(x, y, z) \in K : z > 0\}$$

angenommen, also auf

$$M = \left\{ \left( x, y, \sqrt{1 - x^2 - y^2} \right) : (x, y) \in \mathbb{R}^2 \text{ mit } x^2 + y^2 < 1 \right\}.$$

Für  $(x, y)$  in der offenen Kreisscheibe  $D := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$  können wir nach  $z$  auflösen und haben

$$z = h(x, y) := \sqrt{1 - x^2 - y^2},$$

wobei  $h: D \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion ist. Einsetzen in  $T$  liefert

$$F: D \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y, h(x, y)) = (x - y)\sqrt{1 - x^2 - y^2}$$

mit

$$\begin{aligned} \nabla F(x, y) &= \left( \sqrt{1 - x^2 - y^2} - \frac{(x - y)x}{\sqrt{1 - x^2 - y^2}}, -\sqrt{1 - x^2 - y^2} - \frac{(x - y)y}{\sqrt{1 - x^2 - y^2}} \right) \\ &= \frac{1}{\sqrt{1 - x^2 - y^2}} \left( (1 - x^2 - y^2) - (x - y)x, -(1 - x^2 - y^2) - (x - y)y \right) \\ &= \frac{1}{\sqrt{1 - x^2 - y^2}} (1 - 2x^2 - y^2 + xy, -1 + x^2 + 2y^2 - xy). \end{aligned}$$

Somit ist  $(x, y) \in D$  genau dann ein kritischer Punkt, wenn

$$1 - 2x^2 - y^2 + xy = 0 \tag{71}$$

und

$$1 - x^2 - 2y^2 + xy = 0. \tag{72}$$

Ziehen wir (72) von (71) ab, erhalten wir die notwendige Bedingung

$$-x^2 + y^2 = 0,$$

also  $x^2 = y^2$ . Somit ist  $x = y$  oder  $x = -y$ . Im Falle  $x = y$  ist jedoch  $F(x, y) = F(y, y) = 0$  und wir haben gesehen, dass 0 nicht das Maximum von  $T$  ist. Also muss  $x = -y$  sein. Mit  $x = -y$  wird aus den Gleichungen (71) und (72) jeweils

$$1 - 4y^2 = 0.$$

Es ist also  $y = \pm \frac{1}{2}$  und wir erhalten als Kandidaten für lokale Maximalstellen

$$\left(-\frac{1}{2}, \frac{1}{2}\right) \quad \text{und} \quad \left(\frac{1}{2}, -\frac{1}{2}\right).$$

Nun ist

$$F\left(\frac{1}{2}, -\frac{1}{2}\right) = \frac{1}{2}\sqrt{2} \quad \text{und} \quad F\left(-\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{2}\sqrt{2}.$$

Also ist  $\frac{1}{2}\sqrt{2}$  das globale Maximum von  $F$  und somit auch von  $T$  unter der Nebenbedingung  $x^2 + y^2 + z^2 - 1 = 0$ . Es wird von  $T$  angenommen an den Stellen

$$p := \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\sqrt{2}\right)$$

und  $-p$ .

## Methode der Lagrange-Multiplikatoren

**15.4** *Es sei  $U \subseteq \mathbb{R}^n$  offen;  $f: U \rightarrow \mathbb{R}$  und  $g: U \rightarrow \mathbb{R}$  seien stetig differenzierbar. Hat  $f$  an einer Stelle  $a \in U$  ein Extremum unter der Nebenbedingung  $g(x) = 0$  und ist  $\nabla g(a) \neq 0$ , so gilt*

$$\nabla f(a) = \lambda \nabla g(a)$$

für ein  $\lambda \in \mathbb{R}$ .

Man nennt  $\lambda$  einen Lagrange-Multiplikator.

**Bemerkung 15.5** Um Extrema von  $f$  unter der Nebenbedingung  $g(x) = 0$  zu finden, berechne man alle  $a \in U$  mit  $g(a) = 0$ , für welche

$$\nabla f(a) = \lambda \nabla g(a)$$

für ein  $\lambda \in \mathbb{R}$ . Wenn  $f$  ein Extremum unter der Nebenbedingung  $g(x) = 0$  besitzt und  $\nabla g(x) \neq 0$  für alle  $x \in U$  mit  $g(x) = 0$ , so muss unter den gefundenen Punkten  $a$  die Extremstelle sein.

**Bemerkung 15.6** Für ein *praktisches Verfahren* betrachtet man die Hilfsfunktion

$$L: U \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, \lambda) \mapsto f(x) - \lambda g(x)$$

mit  $x = (x_1, \dots, x_n)$ . Extremstellen von  $f$  unter der Nebenbedingung  $g(x) = 0$  können Punkte  $a \in U$  mit  $g(a) = 0$  und  $\nabla g(a) = 0$  sein; oder sie erfüllen  $\nabla L(a, \lambda) = 0$  für ein  $\lambda \in \mathbb{R}$ , sind also kritische Punkte von  $L$ . Um letzteres einzusehen, schauen wir uns die partiellen Ableitungen von  $L$  an: Wir haben

$$\begin{aligned} \frac{\partial L}{\partial x_1}(a, \lambda) &= \frac{\partial f}{\partial x_1}(a) - \lambda \frac{\partial g}{\partial x_1}(a) \\ &\vdots \\ \frac{\partial L}{\partial x_n}(a, \lambda) &= \frac{\partial f}{\partial x_n}(a) - \lambda \frac{\partial g}{\partial x_n}(a) \\ \frac{\partial L}{\partial \lambda}(a, \lambda) &= -g(a). \end{aligned}$$

Die Bedingung  $\nabla L(a, \lambda) = 0$  bedeutet also, dass die Nebenbedingung  $g(a) = 0$  erfüllt ist und  $\nabla f(a) = \lambda \nabla g(a)$ .

**Beispiel 15.7** Finde die Scheitelpunkte (Punkte mit maximalem bzw. minimalem Abstand vom Ursprung) der Ellipse

$$E := \{(x, y) \in \mathbb{R}^2 : x^2 + xy + y^2 = 5\}.$$

Wir suchen also Extrema der Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2$$

unter der Nebenbedingung  $g(x, y) := x^2 + xy + y^2 - 5 = 0$ . Da  $E$  kompakt ist, nimmt  $f$  auf  $E$  ein globales Maximum an und auch ein globales Minimum. Da  $(0, 0) \notin E$ , können wir  $f$  und  $g$  als Funktionen  $\mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$  betrachten. Dies hat den Vorteil, dass

$$\nabla g(x, y) = (2x + y, 2y + x) = x(2, 1) + y(1, 2) \neq (0, 0)$$

für alle  $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$  (da die Vektoren  $(2, 1)$  und  $(1, 2)$  linear unabhängig sind). Wir betrachten nun die Hilfsfunktion

$$L: (\mathbb{R}^2 \setminus \{(0,0)\}) \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, y, \lambda) \mapsto f(x, y) - \lambda g(x, y) \\ = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 5).$$

Dem gerade diskutierten Vorgehen folgend haben wir kritische Punkte von  $L$  zu finden, berechnen also den Gradienten

$$\nabla L(x, y, \lambda) = (2x - \lambda 2x - \lambda y, 2y - \lambda x - \lambda 2y, -(x^2 + xy + y^2 - 5)).$$

Für einen kritischen Punkt muss der Gradient  $(0, 0, 0)$  sein, also die folgenden drei Bedingungen erfüllt:

$$2x - \lambda 2x - \lambda y = 0 \tag{73}$$

$$2y - \lambda x - \lambda 2y = 0 \tag{74}$$

$$x^2 + xy + y^2 - 5 = 0. \tag{75}$$

Ziehen wir das  $x$ -fache von (74) vom  $y$ -fachen von (73) ab, erhalten wir die notwendige Bedingung

$$\lambda(x^2 - y^2) = 0,$$

so dass also  $\lambda = 0$  sein muss oder  $x = \pm y$ . Im Falle  $\lambda = 0$  folgt aus (73), dass  $x = 0$  und aus (74), dass  $y = 0$  sein muss. Den Punkt  $(0, 0)$ , der nicht auf der Ellipse liegt, haben wir aber schon ausgeschlossen. Im Fall  $x = y$  wird aus (75) die Bedingung

$$y^2 + y^2 + y^2 = 5,$$

so dass also  $y^2 = \frac{5}{3}$  und  $y = \pm\sqrt{5/3}$ . Wir erhalten also zwei Kandidaten für Extremstellen,

$$(x_1, y_1) = (\sqrt{5/3}, \sqrt{5/3}) \quad \text{und} \quad (x_2, y_2) = (-\sqrt{5/3}, -\sqrt{5/3}).$$

Im Fall  $x = -y$  wird aus (75) die Bedingung

$$y^2 - y^2 + y^2 = 5,$$

so dass also  $y = \pm\sqrt{5}$ . Wir erhalten somit zwei weitere Kandidaten für Extremstellen,

$$(x_3, y_3) = (-\sqrt{5}, \sqrt{5}) \quad \text{und} \quad (x_4, y_4) = (\sqrt{5}, -\sqrt{5}).$$

Einsetzen in  $f$  liefert

$$f(x_1, y_1) = f(x_2, y_2) = \frac{10}{3}, \quad f(x_3, y_3) = f(x_4, y_4) = 10.$$

Das Maximum von  $f$  unter der Nebenbedingung  $g(x, y) = 0$  ist also 10 und es wird an den Stellen  $(x_3, y_3)$  und  $(x_4, y_4)$  angenommen. Das Minimum von  $f$  unter der Nebenbedingung  $g(x, y) = 0$  ist  $10/3$  und dieses wird an den Stellen  $(x_1, y_1)$  und  $(x_2, y_2)$  angenommen. Die Scheitelpunkte sind  $(x_j, y_j)$  mit  $j \in \{1, 2, 3, 4\}$ .

Wir erwähnen, dass die Methode der Lagrange-Multiplikatoren in geeigneter Form auch angewandt werden kann, wenn mehrere Nebenbedingungen vorliegen.

**15.8** Es sei  $U \subseteq \mathbb{R}^n$  offen und es seien  $f: U \rightarrow \mathbb{R}$  sowie  $g_1, \dots, g_m: U \rightarrow \mathbb{R}$  stetig differenzierbare Funktionen. Wir suchen Extrema von  $f$  unter den Nebenbedingungen

$$g_1(x) = \dots = g_m(x) = 0, \tag{76}$$

mit  $x = (x_1, \dots, x_n) \in U$ . Praktisches Vorgehen: Man betrachtet die Hilfsfunktion

$$L: U \times \mathbb{R}^m \rightarrow \mathbb{R}, \quad (x, \lambda_1, \dots, \lambda_m) \mapsto f(x) - \sum_{j=1}^m \lambda_j g_j(x).$$

Dann gilt: *Die Extremstellen von  $f$  unter den Nebenbedingungen (76) sind unter denjenigen  $a \in U$  zu finden, für welche*

$$\nabla L(a, \lambda_1, \dots, \lambda_m) = (0, \dots, 0)$$

*für gewisse  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  gilt; oder aber  $g_1(a) = \dots = g_m(a) = 0$  ist und die Vektoren  $\nabla g_1(a), \dots, \nabla g_m(a)$  nicht linear unabhängig sind.*

Ein Beispiel lernen wir in der Übung kennen.

## 16 Iterationsverfahren

Ist  $f: A \rightarrow \mathbb{R}^d$  eine Funktion auf einer Teilmenge  $A \subseteq \mathbb{R}^d$ , so nennen wir ein  $a \in A$  einen *Fixpunkt* von  $f$ , wenn

$$f(a) = a.$$

In diesem Kapitel soll zunächst erläutert werden, wie Fixpunkte näherungsweise (etwa mit dem Computer) berechnet werden können. Anschließend wird eine Methode zur approximativen Berechnung von Nullstellen vorgestellt (ein vereinfachtes Newtonverfahren).

### Approximative Berechnung von Fixpunkten

**Definition 16.1** Es sei  $f: A \rightarrow \mathbb{R}^d$  eine Abbildung derart, dass  $f(A) \subseteq A$  (eine sogenannte *Selbstabbildung* von  $A$ ). Wir nennen  $f$  eine *Kontraktion*, wenn ein  $L \in [0, 1[$  existiert mit

$$|f(y) - f(x)| \leq L |y - x| \quad \text{für alle } x, y \in A.$$

Ein solches  $L$  wird eine *Kontraktionskonstante* für  $f$  genannt.

Ist  $f: A \rightarrow A$  eine Selbstabbildung, so können wir  $f$  beliebig oft iterieren: Für  $x \in A$  definieren wir

$$f^0(x) := x, \quad f^1(x) := f(x), \quad f^2(x) := f(f(x)), \quad f^3(x) := f(f(f(x)))$$

und rekursiv  $f^n(x) := f(f^{n-1}(x))$  für alle  $n \in \mathbb{N}$ .

**16.2** (Banachscher Fixpunktsatz). *Es sei  $A \neq \emptyset$  eine abgeschlossene Teilmenge von  $\mathbb{R}^d$  und  $f: A \rightarrow A$  eine Selbstabbildung von  $A$ , welche eine Kontraktion ist. Sei  $L \in [0, 1[$  eine Kontraktionskonstante für  $f$ . Dann gilt:*

- (a)  $f$  hat genau einen Fixpunkt  $a \in A$ .
- (b) Für jedes  $x_0 \in A$  gilt  $\lim_{n \rightarrow \infty} f^n(x_0) = a$ .
- (c) Es gilt folgende **a priori-Abschätzung**: Für alle  $n \in \mathbb{N}_0$  ist

$$|f^n(x_0) - a| \leq \frac{L^n}{1 - L} |f(x_0) - x_0|.$$

Bevor wir den Satz beweisen, schauen wir ein Beispiel an.

**Beispiel 16.3** Zeigen Sie, dass die Abbildung

$$f: \left[0, \frac{1}{2}\right] \rightarrow \left[0, \frac{1}{2}\right], \quad x \mapsto \frac{1}{2} \cos(x)$$

genau einen Fixpunkt  $a$  besitzt. Finden Sie einen Näherungswert für  $a$ , der um weniger als  $\frac{1}{100}$  abweicht. Finden Sie auch eine Dezimalzahl mit zwei Stellen hinter dem Komma, deren genannte Dezimalstellen mit denen von  $a$  übereinstimmen (so dass wir dann also  $a$  auf zwei Stellen hinter dem Komma genau kennen).

Lösung. Für  $x \in [0, \frac{1}{2}] \subseteq [0, \pi/2]$  ist  $0 \leq \cos(x)$ , also  $\frac{1}{2} \cos(x) \in [0, \frac{1}{2}]$ . Somit ist  $f$  eine Selbstabbildung von  $[0, \frac{1}{2}]$ . Für alle  $x, y \in [0, \frac{1}{2}]$  mit  $x \leq y$  gilt wegen  $f'(t) = -\frac{1}{2} \sin(t)$  mit  $|f'(t)| = \frac{1}{2} \sin(t)$

$$|f(y) - f(x)| = \left| \int_x^y f'(t) dt \right| \leq \int_x^y |f'(t)| dt = \int_x^y \frac{1}{2} \sin(t) dt \leq (y - x) \frac{\sin \frac{1}{2}}{2},$$

da  $\sin|_{[0, \frac{1}{2}]}$  monoton wachsend ist. Mit  $L = \frac{1}{2} \sin \frac{1}{2} = 0,2397127693 < 1$  ist also

$$|f(y) - f(x)| \leq L |y - x| \quad \text{für alle } x, y \in [0, \frac{1}{2}].$$

Wir haben gezeigt, dass  $f$  eine Kontraktion ist mit der Kontraktionskonstanten  $L$ . Nach dem Banachschen Fixpunktsatz hat  $f$  also genau einen Fixpunkt  $a$  und es gilt

$$a = \lim_{n \rightarrow \infty} f^n(x_0)$$

für jedes  $x_0 \in [0, \frac{1}{2}]$ . Um  $a$  näherungsweise zu berechnen, starten wir mit  $x_0 := \frac{1}{2}$  und berechnen  $x_n := f^n(x_0) = f(x_{n-1})$  für kleine  $n \in \mathbb{N}$ . Im Rahmen der Rechengenauigkeit des Taschenrechners ist

$$x_1 = \frac{1}{2} \cos \frac{1}{2} = 0,43879128094,$$

also

$$|x_1 - x_0| = 0,06120871906$$

in der a priori-Abschätzung

$$|a - x_n| \leq \frac{L^n}{1 - L} |x_1 - x_0|.$$

Um die erste Aufgabe zu lösen, wollen wir die rechte Seite  $< \frac{1}{100}$  machen, was für alle  $n \geq 2$  der Fall ist, denn es ist

$$\frac{L^0}{1-L}|x_1 - x_0| = 0,08050736166, \quad \frac{L^1}{1-L}|x_1 - x_0| = 0,01929864261,$$

$$\Delta_2 := \frac{L^2}{1-L}|x_1 - x_0| = 0,00462613106 < \frac{1}{100}.$$

Die Zahl

$$x_2 = \frac{1}{2} \cos(x_1) = 0,45263292166$$

weicht vom Fixpunkt  $a$  als bereits um weniger als  $\frac{1}{100}$  ab. Genauer ist  $a$  im Intervall  $[x_2 - \Delta_2, x_2 + \Delta_2]$ , also in

$$[0,4480067906, 0,45725905272].$$

Die zweite Nachkommastelle von  $a$  könnte also statt 5 vielleicht auch 4 sein, dies können wir noch nicht ausschließen. Und die nächste Approximation ist

$$x_3 = \frac{1}{2} \cos(x_2) = 0,44964937621.$$

Jedoch ist

$$x_4 = \frac{1}{2} \cos(x_3) = 0,45029977813$$

und  $\Delta_4 := \frac{L^4}{1-L}|x_1 - x_0| = 0,00026582772$ , also  $a$  im Intervall

$$[x_4 - \Delta_4, x_4 + \Delta_4] = [0,45003395041, 0,45056560585].$$

Die zweite Nachkommastelle von  $a$  muss also 5 sein, es ist

$$a = 0,45\dots$$

und sogar  $a = 0,450\dots$  mit weiteren uns noch nicht bekannten Nachkommastellen.

Für ein etwas vollständigeres Bild zur Konvergenzgeschwindigkeit zeigt Ihnen die folgende Tabelle die Zahlenwerte von  $x_n$  und  $\Delta_n := \frac{L^n}{1-L}|x_1 - x_0|$  für  $n \in \{0, \dots, 10\}$ , so dass also  $|a - x_n| \leq \Delta_n$  aufgrund der a priori-Abschätzung.

$n$	$x_n$	$\Delta_n$
0	0,5	0,08050736166
1	0,43879128094	0,01929864261
2	0,45263292166	0,00462613106
3	0,44964937621	0,00110894268
4	0,45029977813	0,00026582772
5	0,45015833437	0,00006372229
6	0,45018911053	0,00001527504
7	0,45018241484	0,00000366162
8	0,4501838716	0,000000877738085
9	0,45018355466	0,000000210405027
10	0,45018362361	0,0000000504367717

Wir entnehmen der Tabelle, dass  $a = 0,450183\dots$ , denn es ist

$$a \in [x_{10} - \Delta_{10}, x_{10} + \Delta_{10}] = [0,45018357317, 0,45018367404].$$

**16.4** Beweis des Fixpunktsatzes. Es sei  $x_0 \in A$ . Für  $n \in \mathbb{N}$  definieren wir rekursiv

$$x_n := f(x_{n-1}),$$

so dass also  $x_n = f^n(x_0)$ . Wir zeigen zunächst, dass

$$|x_{n+1} - x_n| \leq L^n |x_1 - x_0| \quad \text{für alle } n \in \mathbb{N}_0. \quad (77)$$

Induktionsanfang: Es gilt  $|x_1 - x_0| = 1|x_1 - x_0| = L^0 |x_1 - x_0| \leq L^0 |x_1 - x_0|$ .

Gelte nun für ein  $n \in \mathbb{N}_0$  bereits  $|x_{n+1} - x_n| \leq L^n |x_1 - x_0|$ . Dann folgt

$$|x_{n+2} - x_{n+1}| = |f(x_{n+1}) - f(x_n)| \leq L |x_{n+1} - x_n| \leq L L^n |x_1 - x_0| = L^{n+1} |x_1 - x_0|,$$

wie benötigt.

Nun zeigen wir, dass

$$|x_{n+m} - x_n| \leq \frac{L^n}{1-L} |x_1 - x_0| \quad \text{für alle } n, m \in \mathbb{N}_0. \quad (78)$$

Für  $m = 0$  ist die linke Seite 0, die Abschätzung somit trivial. Sei nun

$m \geq n$ . Dann gilt

$$\begin{aligned} |x_{n+m} - x_n| &= \left| \sum_{j=0}^{m-1} (x_{n+j+1} - x_{n+j}) \right| \leq \sum_{j=0}^{m-1} |x_{n+j+1} - x_{n+j}| \\ &\leq \sum_{j=0}^{m-1} L^{n+j} |x_1 - x_0| = L^n \sum_{j=0}^{m-1} L^j |x_1 - x_0| \\ &= L^n \frac{1 - L^m}{1 - L} |x_1 - x_0| \leq \frac{L^n}{1 - L} |x_1 - x_0|, \end{aligned}$$

wie benötigt. Hierbei wurde zunächst  $x_{n+m} - x_n$  als Teleskopsumme geschrieben und diese mit der Dreiecksungleichung abgeschätzt. Dann wurde (77) benutzt zum Abschätzen der Summanden und anschließend die geometrische Summenformel für Partialsummen der geometrischen Reihe benutzt.

Da  $\lim_{n \rightarrow \infty} L^n = 0$ , gibt es für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  derart, dass

$$\frac{L^n}{1 - L} |x_1 - x_0| < \varepsilon \quad \text{für alle } n \geq N.$$

Für alle  $n \geq N$  und  $k \geq n$  gilt mit  $m := k - n$  also

$$|x_k - x_n| = |x_{n+m} - x_n| \leq \frac{L^n}{1 - L} |x_1 - x_0| < \varepsilon.$$

Also ist  $(x_n)_{n \in \mathbb{N}}$  eine Cauchyfolge in  $\mathbb{R}^d$  bzw. jede der Komponenten eine Cauchyfolge in  $\mathbb{R}$ , weswegen  $x_n$  gegen ein  $a \in \mathbb{R}^d$  konvergiert. Da  $x_n \in A$  für alle  $n \in \mathbb{N}_0$  und  $A$  in  $\mathbb{R}^d$  abgeschlossen ist, folgt  $a = \lim_{n \rightarrow \infty} x_n \in A$ . Da  $f$  stetig ist, folgt

$$f(a) = f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = a,$$

da  $(x_{n+1})_{n \in \mathbb{N}}$  als Teilfolge von  $(x_n)_{n \in \mathbb{N}}$  ebenfalls gegen  $a$  konvergiert. Also ist  $a$  ein Fixpunkt von  $f$ . Mit  $m \rightarrow \infty$  folgt aus (78) die Ungleichung

$$|a - x_n| \leq \frac{L^n}{1 - L} |x_1 - x_0|,$$

so dass also die a-priori-Abschätzung gilt. Ist  $b \in A$  ebenfalls ein Fixpunkt von  $f$ , so gilt

$$|a - b| = |f(a) - f(b)| \leq L |a - b|,$$

also

$$(1 - L)|a - b| \leq 0.$$

Wäre  $a \neq b$ , so wäre  $|a - b| > 0$ , wegen  $1 - L > 0$  also auch  $(1 - L)|a - b| > 0$ , Widerspruch. Also muss  $a = b$  sein, der Fixpunkt von  $f$  ist also eindeutig.

## Vorbereitende Überlegungen

Ist  $F: U \rightarrow \mathbb{R}^m$  eine  $C^1$ -Funktion auf einer offenen Menge  $U \subseteq \mathbb{R}^n$ , liefert und der Mittelwertsatz eine Integralformel für  $f(y) - f(x)$  für alle  $x, y \in U$ , deren Verbindungsstrecke in  $U$  enthalten ist. Es ist nützlich, eine Bezeichnung für Mengen zu haben, in denen diese Bedingung immer erfüllt ist.

**16.5** Eine Teilmenge  $U \subseteq \mathbb{R}^n$  heißt *konvex*, wenn für alle  $x, y \in U$  ihre Verbindungsstrecke in  $U$  enthalten ist, also

$$x + t(y - x) \in U \quad \text{für alle } t \in [0, 1].$$

**Beispiel 16.6** Für alle  $x_0 \in \mathbb{R}^n$  und  $r > 0$  ist die offene Kugel  $B_r(x_0) = \{x \in \mathbb{R}^n : |x - x_0| < r\}$  konvex.

Sind nämlich  $x, y \in B_r(x_0)$ , so ist  $|x - x_0| < r$  und  $|y - x_0| < r$ . Für alle  $t \in [0, 1]$  ist  $x_0 = tx_0 + (1 - t)x_0$ , also

$$\begin{aligned} |x + t(y - x) - x_0| &= |ty + (1 - t)x - tx_0 - (1 - t)x_0| \\ &\leq |t(y - x_0)| + |(1 - t)(x - x_0)| \\ &= t|y - x_0| + (1 - t)|x - x_0| \\ &< tr + (1 - t)r = r. \end{aligned}$$

Somit ist  $x + t(x - y) \in B_r(x_0)$ .

Analog sieht man, dass abgeschlossene Kugeln  $\overline{B}_r(x_0)$  stets konvex sind.

Die Operatornorm einer Matrix  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  ist sehr praktisch, da sie uns für jeden Vektor  $x \in \mathbb{R}^n$  die Abschätzung

$$|Ax| \leq \|A\|_{\text{op}} |x|$$

ermöglicht. Für praktische Rechnungen würden wir aber gern  $\|A\|_{\text{op}}$  mit Hilfe der Beträge  $|a_{ij}|$  der Matrixeinträge nach oben abschätzen können. Indem

wir den Vektorraum  $\mathbb{R}^{m \times n}$  der  $m \times n$ -Matrizen mit  $\mathbb{R}^{mn}$  identifizieren, nämlich  $A = (a_{ij})$  mit

$$(a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn})^T \in \mathbb{R}^{mn},$$

können wir einer Matrix ihre euklidische Länge zuordnen,

$$|A| := \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}.$$

Nützlich ist auch das Maximum der Beträge der Matrixeinträge,

$$\|A\|_\infty := \max\{|a_{ij}| : i = 1, \dots, m; j = 1, \dots, n\}.$$

Für einen Spaltenvektor  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  benutzt man entsprechend manchmal seine sogenannte *Maximum-Norm*

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}.$$

Es gilt

$$\|x\|_\infty \leq |x| \leq \sqrt{n} \|x\|_\infty \quad \text{für alle } x \in \mathbb{R}^n, \quad (79)$$

denn für jedes  $j \in \{1, \dots, n\}$  ist

$$|x_j| = \sqrt{x_j^2} \leq \sqrt{x_1^2 + \dots + x_n^2} = |x|;$$

Bildung des Maximums über alle  $j$  liefert  $\|x\|_\infty \leq |x|$ . Da  $x_j^2 \leq (\|x\|_\infty)^2$ , ist umgekehrt

$$|x| = \sqrt{x_1^2 + \dots + x_n^2} \leq \sqrt{\|x\|_\infty^2 + \dots + \|x\|_\infty^2} \leq \sqrt{n} \|x\|_\infty,$$

mit  $n$  Summanden in der Wurzel.

**16.7** Für jede  $m \times n$ -Matrix  $A = (a_{ij})$  gilt

$$\|A\|_\infty \leq |A| \quad \text{und} \quad |A| \leq n \|A\|_\infty. \quad (80)$$

Weiter gilt

$$\|A\|_\infty \leq \|A\|_{\text{op}} \leq nm \|A\|_\infty.$$

Es ist (80) ein Spezialfall von (79). Für alle  $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$  ist wegen  $|e_j| = 1 \leq 1$

$$\|A\|_{\text{op}} \geq |Ae_j| = \left| \sum_{k=1}^m a_{kj} e_k \right| = \sqrt{\sum_{k=1}^m a_{kj}^2} \geq \sqrt{a_{ij}^2} = |a_{ij}|.$$

Übergang zum Supremum in  $(i, j)$  liefert

$$\|A\|_{\infty} \leq \|A\|_{\text{op}}.$$

Ist  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  mit  $|x| \leq 1$ , so ist  $|x_i| \leq 1$  für alle  $i \in \{1, \dots, n\}$ , somit

$$\begin{aligned} |Ax| &= \left| \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e_i \right| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij} x_j e_i| \\ &= \sum_{i=1}^m \sum_{j=1}^n \underbrace{|a_{ij}|}_{\leq \|A\|_{\infty}} \underbrace{|x_j|}_{\leq 1} \\ &\leq \sum_{i=1}^m \sum_{j=1}^n \|A\|_{\infty} = nm \|A\|_{\infty}. \end{aligned}$$

Übergang zum Supremum in  $x$  liefert  $\|A\|_{\text{op}} \leq nm \|A\|_{\infty}$ .

## Approximative Berechnung von Nullstellen

Wir erläutern nun, wie sich unter geeigneten Voraussetzungen Nullstellen approximativ berechnen lassen mit einem vereinfachten Newtonverfahren. Zunächst geben wir eine grobe Idee des eigentlichen Newton-Verfahrens.

**16.8** Im Falle differenzierbarer Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  einer Variablen funktioniert das Newton-Verfahren (das manche von Ihnen aus der Schule kennen) wie folgt: Kennen wir bereits ein  $x_0 \in \mathbb{R}$  nahe einer Nullstelle  $a$  von  $f$  und ist  $f'(x_0) \neq 0$ , betrachten wir die Tangente

$$y = f(x_0) + f'(x_0)(x - x_0)$$

der Funktion  $f$  um die Stelle  $x_0$  und suchen deren Nullstelle  $x_1$  (in der Hoffnung, damit der Nullstelle  $a$  näher zu kommen als mit  $x_0$ ). Wir haben also

$$0 = f(x_0) + f'(x_0)(x_1 - x_0)$$

aufzulösen nach  $x_1$  und erhalten

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

Man macht nun so weiter, definiert also

$$x_2 := x_1 - \frac{f(x_1)}{f'(x_1)}$$

und rekursiv  $x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$  für  $n \in \mathbb{N}$ . Unter geeigneten Voraussetzungen konvergieren dann die Punkte  $x_n$  gegen die Nullstelle  $a$  (und sogar sehr schnell, im Falle einer  $C^2$ -Funktion  $f$  hat man sogenannte “quadratische Konvergenz”). Bei einem *vereinfachten Newtonverfahren* arbeitet man durchgehend mit  $f'(x_0)$  statt mit  $f'(x_n)$ , betrachtet stattdessen also

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_0)}.$$

Im allgemeinen ist die Konvergenz gegen  $a$  dann nicht ganz so schnell, das Verfahren ist mathematisch aber viel einfacher zu behandeln (eine Konvergenzuntersuchung des Newton-Verfahrens wäre für uns recht anstrengend). Im Falle eine Funktion  $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$  nimmt man entsprechend

$$x_{n+1} := x_n - J_f(x_n)^{-1}(f(x_n))$$

im Falle des Newtonverfahrens und

$$x_{n+1} := x_n - J_f(x_0)^{-1}(f(x_n))$$

im Falle des vereinfachten Newtonverfahrens. Für das vereinfachte Verfahren finden Sie im folgenden sowohl Beispiele als auch einen Konvergenzbeweis unter geeigneten Voraussetzungen.

**16.9** *Es sei  $x_0 \in \mathbb{R}^d$ ,  $r > 0$  und  $f: \overline{B}_r(x_0) \rightarrow \mathbb{R}^d$  eine stetige Abbildung, die auf  $B_r(x_0)$  stetig differenzierbar ist. Wir nehmen an, dass die  $n \times n$ -Matrix  $J_f(x_0)$  invertierbar ist und*

$$\sigma := \sup \{ \|J_f(x) - J_f(x_0)\|_{\text{op}} : x \in B_r(x_0) \} < \frac{1}{\|J_f(x_0)^{-1}\|_{\text{op}}}; \quad (81)$$

also ist  $\theta := \frac{1}{\|J_f(x_0)^{-1}\|_{\text{op}}} - \sigma > 0$ . Dann gilt

$$\overline{B}_{\theta r}(f(x_0)) \subseteq f(\overline{B}_r(x_0)).$$

Ist  $|f(x_0)| \leq \theta r$ ,

so ist  $0 \in \overline{B}_{\theta r}(f(x_0))$  und somit hat  $f$  eine Nullstelle  $a$ . Diese lässt sich berechnen als

$$a = \lim_{n \rightarrow \infty} g^n(x_0)$$

mit  $g: \overline{B}_r(x_0) \rightarrow \overline{B}_r(x_0)$ ,  $x \mapsto x - J_f(x_0)^{-1}(f(x))$ .

**Beispiel 16.10** Zeigen Sie, dass die Funktion

$$f: \left[-\frac{1}{2}, \frac{1}{2}\right] \rightarrow \mathbb{R}, \quad x \mapsto x + \frac{1}{5} \cos x - \frac{1}{4} e^{x/2}$$

eine Nullstelle hat und geben Sie eine Iterationsvorschrift an, mit der diese approximativ berechnet werden kann.

Für  $x$  im Definitionsbereich ist

$$f'(x) = 1 - \frac{1}{5} \sin(x) - \frac{1}{8} e^{x/2},$$

also

$$f'(0) = 1 - \frac{1}{8} = \frac{7}{8}$$

und somit

$$\frac{1}{|f'(0)^{-1}|} = |f'(0)| = \frac{7}{8}.$$

Weiter gilt für alle  $x \in [-\frac{1}{2}, \frac{1}{2}]$

$$|f'(x) - f'(0)| = \left| \frac{1}{5} \sin(x) - \frac{1}{8} (e^{x/2} - 1) \right| \leq \frac{1}{5} + \frac{1}{8} (e^{x/4} + 1) \leq \frac{1}{5} + \frac{1}{8} \cdot \frac{7}{3} = \frac{24 + 35}{120} = \frac{59}{120} \leq \frac{1}{2}$$

unter Benutzung der Dreiecksungleichung und der Abschätzung  $e^{1/4} = 1,2840254\dots \leq \frac{4}{3}$ . Also ist

$$\sigma := \sup\{|f'(x) - f'(0)| : x \in [-\frac{1}{2}, \frac{1}{2}]\} \leq \frac{1}{2}$$

und somit

$$\theta := \frac{1}{|f'(0)^{-1}|} - \sigma = \frac{7}{8} - \sigma \geq \frac{7}{8} - \frac{1}{2} = \frac{3}{8}.$$

Mit  $r := \frac{1}{2}$  gilt nun

$$|f(0)| = \left| \frac{1}{5} - \frac{1}{4} \right| = \frac{1}{20} \leq \frac{3}{16} \leq \theta r.$$

Nach 16.9 hat  $f$  also eine Nullstelle  $a$  in  $[-\frac{1}{2}, \frac{1}{2}]$  und es ist

$$a = \lim_{n \rightarrow \infty} g^n(0)$$

mit  $g: [-\frac{1}{2}, \frac{1}{2}] \rightarrow [-\frac{1}{2}, \frac{1}{2}]$ ,  $x \mapsto x - \frac{f(x)}{f'(0)} = x - \frac{8}{7}f(x)$ .

**Beispiel 16.11** Wir betrachten die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f(x, y) = \left( x + \frac{1}{20} \cos(y), y + \frac{1}{24} \cos(x) \right).$$

Wir wollen klären, ob  $f$  in der abgeschlossenen Einheitskugel  $\overline{B}_1(0, 0) \subseteq \mathbb{R}^2$  eine Nullstelle besitzt. Es ist

$$J_f(x, y) = \begin{pmatrix} 1 & -\frac{1}{20} \sin(y) \\ -\frac{1}{24} \sin(x) & 1 \end{pmatrix},$$

also  $J_f(0, 0) = \mathbf{1}$  die Einheitsmatrix. Weiter gilt für alle  $(x, y) \in \mathbb{R}^2$

$$\begin{aligned} \|J_f(x, y) - J_f(0, 0)\|_{\text{op}} &= \left\| \begin{pmatrix} 0 & \frac{1}{20} \sin(y) \\ \frac{1}{24} \sin(x) & 0 \end{pmatrix} \right\|_{\text{op}} \\ &\leq 4 \max\{|\sin(y)|/20, |\sin(x)|/24\} \leq \frac{1}{5}, \end{aligned}$$

nach 16.7. Insbesondere ist

$$\sigma := \sup\{\|J_f(x, y) - J_f(0, 0)\|_{\text{op}} : (x, y) \in B_1(0, 0)\} \leq \frac{1}{5} < 1 = \frac{1}{\|J_f(0, 0)^{-1}\|_{\text{op}}}$$

und  $\theta := \frac{1}{\|J_f(0, 0)^{-1}\|_{\text{op}}} - \sigma \geq 4/5$ . Nun ist  $f(0, 0) = (1/20, 1/24)$ , somit

$$|f(0, 0)| = \sqrt{(1/20)^2 + (1/24)^2} \leq \sqrt{(1/20)^2 + (1/20)^2} = \sqrt{2} \cdot \frac{1}{20} \leq 2 \cdot \frac{1}{20} = \frac{1}{10} \leq \frac{4}{5} \leq \theta r$$

mit  $r := 1$ . Nach 16.9 hat  $f|_{\overline{B}_1(0, 0)}$  eine Nullstelle  $a$  und diese ist

$$a = \lim_{n \rightarrow \infty} x_n$$

mit  $x_0 := (0, 0)$  und  $x_{n+1} := x_n - f(x_n)$  für  $n \in \mathbb{N}$ .

**16.12** Beweis für 16.9 (wird in der Vorlesung übersprungen). Wir kürzen  $A := J_f(x_0)$  ab. Für  $y \in \overline{B}_{\theta r}(f(x_0))$  haben wir

$$y = f(x_0) + \delta$$

mit  $\delta := y - f(x_0)$ , wobei  $\delta$  die Länge  $|\delta| \leq \theta r$  hat. Wir betrachten die Abbildung

$$g: \overline{B}_r(x_0) \rightarrow \mathbb{R}^d, \quad x \mapsto x - A^{-1}(f(x) - y).$$

Wir zeigen nun, dass  $g$  eine Selbstabbildung von  $\overline{B}_r(x_0)$  ist und zudem eine Kontraktion.

Zunächst betrachten wir das Restglied

$$R(x) = f(x) - f(x_0) - A(x - x_0) \tag{82}$$

in der affin-linearen Approximation

$$f(x) = f(x_0) + A(x - x_0) + R(x).$$

Wir entnehmen der Formel (82), dass  $R$  stetig ist,  $R(x_0) = 0$  und zudem  $R$  auf  $B_r(x_0)$  eine  $C^1$ -Funktion ist mit Jacobi-Matrix

$$J_R(x) = J_f(x) - A$$

für alle  $x \in B_r(x_0)$ , so dass also

$$\|J_R(x)\|_{\text{op}} \leq \sigma.$$

Der Mittelwertsatz in Integralform zeigt nun, dass für alle  $x, z \in B_r(x_0)$

$$\begin{aligned} |R(x) - R(z)| &= \left| \int_0^1 J_R(z + t(x - z))(x - z) dt \right| \\ &\leq \int_0^1 \|J_R(z + t(x - z))\|_{\text{op}} |x - z| dt \leq \sigma |x - z|. \end{aligned}$$

Dann gilt

$$|R(x) - R(z)| \leq \sigma |x - z| \quad \text{für alle } x, z \in \overline{B}_r(x_0),$$

denn wählen wir eine Folge  $t_n \in [0, 1[$  mit  $t_n \rightarrow 1$  für  $n \rightarrow \infty$ , so gilt so ist  $x_n := x_0 + t_n(x - x_0) \in B_r(x_0)$  und  $z_n := x_0 + t_n(z - x_0) \in B_r(x_0)$  für alle  $n \in \mathbb{N}$  und es gilt

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{und} \quad \lim_{n \rightarrow \infty} z_n = z.$$

Da  $R$  stetig ist, ist also

$$R(x) = \lim_{n \rightarrow \infty} R(x_n) \quad \text{und} \quad R(z) = \lim_{n \rightarrow \infty} R(z_n).$$

Nun gilt für jedes  $n \in \mathbb{N}$

$$|R(x_n) - R(z_n)| \leq \sigma |x_n - z_n| = \sigma |x_0 + t_n(x - x_0) - x_0 - t_n(z - x_0)| = \sigma t_n |x - z|.$$

Mit  $n \rightarrow \infty$  folgt, dass  $|R(x) - R(z)| \leq \sigma |x - z|$ .

Da  $R(x_0) = 0$ , gilt insbesondere für alle  $x \in \overline{B}_r(x_0)$

$$|R(x)| = |R(x) - R(x_0)| \leq \sigma |x - x_0|.$$

Es ist  $g(\overline{B}_r(x_0)) \subseteq \overline{B}_r(x_0)$ : Für  $x \in \overline{B}_r(x_0)$  ist nämlich

$$\begin{aligned} |g(x) - x_0| &= |x - A^{-1}f(x) + A^{-1}y - x_0| \\ &= |A^{-1}Ax - A^{-1}f(x) + A^{-1}f(x_0) + A^{-1}\delta - A^{-1}Ax_0| \\ &\leq \|A^{-1}\|_{\text{op}} |Ax - f(x) + f(x_0) + \delta - Ax_0| \\ &\leq \|A^{-1}\|_{\text{op}} (|f(x) - f(x_0) - A(x - x_0)| + |\delta|) \\ &\leq \|A^{-1}\|_{\text{op}} (|R(x)| + |\delta|) \\ &\leq \|A^{-1}\|_{\text{op}} (\sigma r + \theta r) = r \|A^{-1}\|_{\text{op}} \underbrace{(\sigma + \theta)}_{=1/\|A^{-1}\|_{\text{op}}} = r \end{aligned}$$

unter Benutzung von  $|R(x)| \leq \sigma |x - x_0| \leq \sigma r$ .

Man beachte, dass  $\sigma < \frac{1}{\|A^{-1}\|_{\text{op}}}$  nach (81). Multiplikation beider Seiten mit  $\|A^{-1}\|_{\text{op}}$  liefert

$$L := \|A^{-1}\|_{\text{op}} \sigma < 1.$$

Wir zeigen nun, dass  $g$  eine Kontraktion ist mit Kontraktionskonstante  $L$ . In der Tat gilt für alle  $x, z \in \overline{B}_r(x_0)$

$$\begin{aligned} |g(z) - g(x)| &= |z - A^{-1}(f(z) - y) - x + A^{-1}(f(x) - y)| \\ &= |z - x - A^{-1}(f(z) - f(x))| \\ &= |A^{-1}(f(x) - Ax - (f(z) - Az))| \\ &\leq \|A^{-1}\|_{\text{op}} |f(x) - Ax - (f(z) - Az)| \\ &= \|A^{-1}\|_{\text{op}} |f(x) - f(x_0) - A(x - x_0) - (f(z) - f(x_0) - A(x - x_0))| \\ &= \|A^{-1}\|_{\text{op}} |R(x) - R(z)| \\ &\leq \|A^{-1}\|_{\text{op}} \sigma |x - z| = L |x - z|. \end{aligned}$$

Nach dem Banachschen Fixpunktsatz hat  $g$  genau einen Fixpunkt  $a$ . Man beachte, dass für  $a \in \overline{B}_r(x_0)$  äquivalent sind:

$$a = g(a) \quad \Leftrightarrow \quad a = a - A^{-1}(f(a) - y) \quad \Leftrightarrow \quad f(a) = y.$$

Also gibt es genau ein  $a \in \overline{B}_r(x_0)$  (nämlich den Fixpunkt von  $g$ ) mit  $f(a) = y$ .

**Bemerkung 16.13** Ist  $f = (f_1, \dots, f_d): \overline{B}_r(x_0) \rightarrow \mathbb{R}^d$  stetig auf  $\overline{B}_r(x_0) \subseteq \mathbb{R}^d$  mit  $r > 0$  und  $x_0 \in \mathbb{R}^d$  und zudem  $f$  stetig differenzierbar auf  $B_r(x_0)$ , so kann man für jedes  $\varepsilon > 0$  ein  $s \in ]0, r]$  finden mit

$$\sup\{\|f'(x) - f'(x_0)\|_{\text{op}} : x \in B_s(x_0)\} \leq \varepsilon.$$

Ist  $f'(x_0)$  invertierbar, können wir durch Verkleinern von  $r$  also immer erreichen, dass

$$\sup\{\|f'(x) - f'(x_0)\|_{\text{op}} : x \in B - r(x_0)\} < \frac{1}{\|f'(x_0)^{-1}\|_{\text{op}}},$$

wie in 16.9 benötigt. Wegen der Stetigkeit der partiellen Ableitungen gibt es nämlich ein  $s \in ]0, r]$  derart, dass

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(x_0) \right| \leq \frac{\varepsilon}{d^2}$$

für alle  $i, j \in \{1, \dots, d\}$  und alle  $x \in B_s(x_0)$ . Nach 16.7 gilt dann

$$\|J_f(x) - J_f(x_0)\|_{\text{op}} \leq d^2 \frac{\varepsilon}{d^2} = \varepsilon$$

für alle  $x \in B_s(x_0)$ .

## 17 Gewöhnliche Differentialgleichungen

Zeitabhängige Größen lassen sich oft durch Differentialgleichungen beschreiben. Diese werden in Natur- und Ingenieurwissenschaften ständig benutzt.

### Grundbegriffe und erste Beispiele

**Beispiel 17.1** (Ein Populationsmodell/Exponentielles Wachstum).

Wir betrachten eine Population von Bakterien in einer Petrischale. Zu einer festen Zeit  $t_0$  (zu der wir die Beobachtung beginnen) gebe es  $N(t_0) = N_0 > 0$  Bakterien. Wie groß ist die Zahl  $N(t)$  der Bakterien zu einer beliebigen Zeit  $t$ ? Wir modellieren das Problem wie folgt: Da  $N(t)$  sehr groß ist und gezeichnet von einer Kurve nicht zu unterscheiden wäre, machen wir die vereinfachende Annahme, dass  $N(t)$  eine stetig differenzierbare Funktion von  $t$  ist. Wir machen weiter die Annahme, dass die Änderungsrate/Ableitung  $N'(t)$  zur Zeit  $t$  proportional zu  $N(t)$  ist (liegen z.B. doppelt so viele Bakterien vor, erwarten wir die doppelte Zuwachsrage).<sup>19</sup> Mit einer reellen Konstanten  $\alpha > 0$  ist also

$$N'(t) = \alpha N(t). \quad (83)$$

Um die Gleichung zu lösen, teilen wir durch  $N(t) > 0$  und erhalten

$$\alpha = \frac{N'(t)}{N(t)} = \frac{d}{dt}(\ln(N(t))).$$

Integrieren beider Seiten von  $t_0$  bis  $t$  liefert

$$\alpha(t - t_0) = \ln(N(t)) - \ln(N(t_0)) = \ln \frac{N(t)}{N_0}.$$

Wir wenden die Exponentialfunktion auf beide Seiten an und lösen nach  $N(t)$  auf, mit dem Ergebnis

$$N(t) = N_0 e^{\alpha(t-t_0)}. \quad (84)$$

Dann ist  $N(t_0) = N_0 e^0 = N_0$  und in der Tat (unter Benutzung der Kettenregel)

$$N'(t) = \alpha N_0 e^{\alpha(t-t_0)} = \alpha N(t).$$

---

<sup>19</sup>Dies ist vernünftig, so lange die Population nicht zu groß ist, so dass ausreichend Nährstoffe und Platz zur Verfügung stehen.

Die durch (84) gegebene Funktion  $N: \mathbb{R} \rightarrow \mathbb{R}$  löst also die Gleichung (83) unter der Anfangsbedingung  $N(t_0) = N_0$  und ist die einzige solche (stets positive) Funktion (da wir in jedem Schritt des Rechenwegs nur Schlussfolgerungen gezogen haben). Nach (84) zeigt die Bakterienpopulation exponentielles Wachstum. Da  $N(t) \rightarrow \infty$  für  $t \rightarrow \infty$ , ist die Lösung für große  $t$  übrigens nicht realistisch und unsere Modellierung war zu schlecht (in eine Petrischale passen nicht beliebig viele Bakterien).

Im vorigen Beispiel haben wir eine Differentialgleichung gelöst (genauer: ein Anfangswertproblem). Die Begriffe sind wie folgt:

**Definition 17.2** Eine *Differentialgleichung* (erster Ordnung) in  $\mathbb{R}^n$  ist eine Gleichung der Form

$$y'(t) = f(t, y(t)) \quad (85)$$

(oder kurz:  $y' = f(t, y)$ ) mit einer gegebenen Funktion  $f: U \rightarrow \mathbb{R}^n$  auf einer Teilmenge  $U \subseteq \mathbb{R} \times \mathbb{R}^n$ . Eine *Lösung der Differentialgleichung* (85) ist eine  $C^1$ -Funktion  $\phi: I \rightarrow \mathbb{R}^n$  auf einem nicht entarteten Intervall  $I \subseteq \mathbb{R}$  derart, dass

$$(t, \phi(t)) \in U \quad \text{für alle } t \in I$$

und

$$\phi'(t) = f(t, \phi(t)) \quad \text{für alle } t \in I.$$

**Definition 17.3** Gegeben  $f$  wie zuvor und  $(t_0, y_0) \in U \subseteq \mathbb{R} \times \mathbb{R}^n$  nennt man eine Lösung  $\phi: I \rightarrow \mathbb{R}^n$  der Differentialgleichung (85) eine *Lösung des Anfangswertproblems*

$$\begin{cases} y' &= f(t, y) \\ y(t_0) &= y_0, \end{cases}$$

wenn zudem  $t_0 \in I$  und  $\phi(t_0) = y_0$  gilt.

**Bemerkung 17.4** (a) Wir betrachten ausschließlich *explizite* Differentialgleichungen der Form (85), nicht implizite Differentialgleichungen der Form

$$0 = f(t, y(t), y'(t)).$$

(b) In der Literatur spricht man oft lediglich im Falle  $n = 1$  von Differentialgleichungen, im Falle  $n \geq 2$  von *Systemen* von Differentialgleichungen.

(c) Später betrachten wir auch Differentialgleichungen höherer Ordnung, der Form

$$y^{(m)}(t) = f(t, y(t), y'(t), y''(t), \dots, y^{(m-1)}(t)).$$

(d) Die von uns betrachteten Differentialgleichungen sind *gewöhnliche* Differentialgleichungen, wir suchen also Lösungen  $t \mapsto \phi(t)$ , die Funktionen einer reellen Variablen sind. Sucht man hingegen Funktionen  $\phi$  mehrerer Variablen und stellt Gleichungen unter Benutzung partieller Ableitungen auf, so spricht man von einer *partiellen Differentialgleichung*; solche Gleichungen sind nicht Gegenstand dieser Vorlesung. Ein Beispiel ist die Laplace-Gleichung

$$0 = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2},$$

deren Lösungen sogenannte *harmonische Funktionen*  $(x, y) \mapsto \phi(x, y)$  sind.

(e) Wir kürzen “Differentialgleichung” mitunter als “DGL” ab und “Anfangswertproblem” als “AWP.” Im Englischen kürzt man “ordinary differential equation” als “ODE” ab, “partial differential equation” als “PDE”.

### 17.5 (Fortsetzung von Beispiel 17.1).

In Beispiel 17.1 haben wir also das Anfangswertproblem

$$\begin{cases} y'(t) &= f(t, y(t)), \\ y(t_0) &= N_0 \end{cases}$$

gelöst mit der durch  $f(t, y) = \alpha y$  gegebenen Funktion. Was wir als Definitionsbereich  $U$  von  $f$  nehmen, haben wir nun zu präzisieren. Natürliche Wahlen sind  $U := \mathbb{R} \times \mathbb{R}$  oder (weil die Lösung überall positiv sein soll)  $U := \mathbb{R} \times ]0, \infty[$ . Die Formulierung in Beispiel 17.1 war noch unpräzise und unvollständig, da der Definitionsbereich nicht angegeben wurde.

**Beispiel 17.6** (Radioaktiver Zerfall). Wir betrachten eine Probe einer radioaktiv zerfallenden Substanz. Die Anzahl der Atome der Substanz zur Zeit  $t$  sei  $N(t) > 0$  (wobei wir wieder näherungsweise  $N$  als stetig differenzierbare Funktion annehmen) und zu einer Anfangszeit  $t_0$  liegen  $N_0$  Atome vor. Modellierung: Wir nehmen an, dass  $N'(t)$  proportional zu  $N(t)$  ist, mit einer negativen reellen Proportionalitätskonstanten  $-k$  (mit  $k > 0$ ). Zu lösen haben wir also das Anfangswertproblem

$$\begin{cases} N'(t) &= -kN(t), \\ N(t_0) &= N_0. \end{cases}$$

Der gleiche Rechenweg wie in Beispiel 17.1 zeigt, dass es auf  $\mathbb{R}$  genau eine überall positive Lösung gibt,

$$N(t) = N_0 e^{-k(t-t_0)}.$$

Die Zahl der Atome nimmt also exponentiell ab.

Wir betrachten nun drei Arten von gewöhnlichen Differentialgleichungen, für welche elementare Lösungsmethoden bekannt sind. Dies sind zum einen homogene sowie inhomogene lineare Differentialgleichungen. Zum anderen betrachten wir sogenannte Differentialgleichungen mit getrennten Variablen, von der Form  $y'(t) = g(t)h(y(t))$ .

## Homogene lineare Differentialgleichungen 1. Ordnung

**Definition 17.7** Eine *lineare Differentialgleichung* erster Ordnung ist eine Differentialgleichung der Form

$$y'(t) = a(t)y(t) + b(t),$$

wobei  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall ist und  $a: J \rightarrow \mathbb{R}$  sowie  $b: J \rightarrow \mathbb{R}$  stetige Funktionen. Ist  $b(t) = 0$  für alle  $t \in J$ , die DGL also von der Form

$$y'(t) = a(t)y(t),$$

so spricht man von einer *homogenen* linearen Differentialgleichung erster Ordnung; andernfalls heißt die lineare Differentialgleichung *inhomogen*. Ist  $a$  eine konstante Funktion, also einfach

$$y'(t) = a y(t) + b(t),$$

so spricht man von einer linearen Differentialgleichung erster Ordnung *mit konstanten Koeffizienten*.

**17.8** (Homogene lineare Differentialgleichung 1. Ordnung). *Es sei  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $a: J \rightarrow \mathbb{R}$  eine stetige Funktion. Dann gilt:*

- (a) (Globale Existenz von Lösungen). *Für jedes  $t_0 \in J$  und  $y_0 \in \mathbb{R}$  ist  $\phi: J \rightarrow \mathbb{R}$ ,*

$$t \mapsto y_0 e^{\int_{t_0}^t a(s) ds} \tag{86}$$

*eine auf ganz  $J$  definierte Lösung des Anfangswertproblems*

$$\begin{cases} y'(t) &= a(t)y(t) \\ y(t_0) &= y_0. \end{cases}$$

(b) (Eindeutigkeit von Lösungen). Ist  $I \subseteq J$  ein nicht entartetes Intervall mit  $t_0 \in I$  und auch  $\gamma: I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems aus (a), so gilt  $\gamma = \phi|_I$ .

Beweis für (a): Es ist  $\phi(t_0) = y_0 e^0 = y_0$  und Ableiten mit der Kettenregel zeigt, dass

$$\phi'(t) = y_0 e^{\int_{t_0}^t a(s) ds} \frac{d}{dt} \int_{t_0}^t a(s) ds = y_0 e^{\int_{t_0}^t a(s) ds} a(t) = a(t) \phi(t).$$

Also löst  $\phi$  die Differentialgleichung und auch das Anfangswertproblem.

(b): Nach (a) ist

$$\psi: J \rightarrow \mathbb{R}, \quad t \mapsto e^{\int_{t_0}^t a(s) ds}$$

eine Lösung der DGL (mit  $\psi(t_0) = 1$ ); weiter ist  $\phi = y_0 \psi$ . Beachten Sie, dass  $\psi(t) \neq 0$  für alle  $t \in J$ . Wir können also die Funktion

$$h: I \rightarrow \mathbb{R}, \quad t \mapsto \frac{\gamma(t)}{\psi(t)}$$

betrachten. Wir leiten mit der Quotientenregel ab; da  $\gamma$  und  $\psi$  Lösungen der Differentialgleichung sind, ergibt sich für jedes  $t \in I$

$$h'(t) = \frac{\psi(t)\gamma'(t) - \gamma(t)\psi'(t)}{\psi(t)^2} = \frac{a(t)\psi(t)\gamma(t) - a(t)\gamma(t)\psi(t)}{\psi(t)^2} = 0.$$

Die Funktion  $h$  ist also konstant, mit dem konstanten Funktionswert

$$h(t_0) = \frac{\gamma(t_0)}{\psi(t_0)} = \frac{y_0}{1} = y_0.$$

Also ist  $\frac{\gamma(t)}{\psi(t)} = y_0$  und somit

$$\gamma(t) = y_0 \psi(t) = \phi(t).$$

Beispiel 17.1 und Beispiel 17.6 waren bereits Beispiele für homogene lineare Differentialgleichungen, mit konstanten Koeffizienten. Schauen wir uns ein Beispiel an, bei dem nicht konstante Koeffizienten vorliegen.

**Beispiel 17.9** Finden Sie die auf ganz  $\mathbb{R}$  definierte Lösung  $\phi$  des Anfangswertproblems

$$\begin{cases} y'(t) &= t y(t) \\ y(0) &= 5. \end{cases}$$

Lösung: Nach 17.8 (a) ist

$$\phi(t) = y_0 e^{\int_{t_0}^t a(s) ds}$$

mit  $t_0 = 0$ ,  $y_0 = 5$  und  $a(s) = s$ , also

$$\phi(t) = 5 e^{\int_0^t s ds} = 5 e^{t^2/2}.$$

## Inhomogene lineare Differentialgleichungen 1. Ordnung

**17.10** (Inhomogene lineare Differentialgleichungen 1. Ordnung). *Es seien  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $a: J \rightarrow \mathbb{R}$  sowie  $b: J \rightarrow \mathbb{R}$  stetige Funktionen. Dann gilt:*

- (a) (Globale Existenz von Lösungen). *Für jedes  $t_0 \in J$  und  $y_0 \in \mathbb{R}$  ist die auf ganz  $J$  definierte Funktion  $\phi: J \rightarrow \mathbb{R}$ ,*

$$\phi(t) = \left( y_0 + \int_{t_0}^t b(\tau) e^{-\int_{t_0}^{\tau} a(s) ds} d\tau \right) e^{\int_{t_0}^t a(s) ds}$$

*eine Lösung des Anfangswertproblems*

$$\begin{cases} y'(t) &= a(t)y(t) + b(t) \\ y(t_0) &= y_0. \end{cases} \quad (87)$$

- (b) (Eindeutigkeit von Lösungen). *Ist  $I \subseteq J$  ein nicht entartetes Intervall mit  $t_0 \in I$  und auch  $\gamma: I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems aus (a), so gilt  $\gamma = \phi|_I$ .*

Beweis für (a): Wir suchen eine Lösung der Form

$$\phi(t) = c(t)\eta(t),$$

wobei  $\eta(t) = e^{\int_{t_0}^t a(s) ds}$  eine Lösung der zugehörigen homogenen linearen Differentialgleichung  $y'(t) = a(t)y(t)$  ist (diesen Ansatz nennt man auch "Variation der Konstanten.") Dann ist  $c(t_0) = y_0$  und

$$\phi'(t) = c'(t)\eta(t) + c(t)\eta'(t) = c'(t)\eta(t) + c(t)a(t)\eta(t). \quad (88)$$

Wir setzen  $\phi'(t)$  aus (88) und  $\phi(t) = c(t)\eta(t)$  in (87) ein und erhalten die Bedingung

$$c'(t)\eta(t) + c(t)a(t)\eta(t) = a(t)c(t)\eta(t) + b(t),$$

also

$$c'(t)\eta(t) = b(t),$$

also

$$c'(t) = b(t)\eta(t)^{-1} = b(t)e^{-\int_{t_0}^t a(s) ds} \quad \text{für alle } t \in J.$$

Äquivalent hierzu ist

$$c(t) = y_0 + \int_{t_0}^t b(\tau)e^{-\int_{t_0}^{\tau} a(s) ds} \quad \text{für alle } t \in J,$$

was den Beweis für (a) beendet.

(b): Wir betrachten  $h: I \rightarrow \mathbb{R}$ ,  $t \mapsto \gamma(t) - \phi(t)$ . Dann ist  $h(t_0) = y_0 - y_0 = 0$ . Weiter gilt

$$h'(t) = \gamma'(t) - \phi'(t) = a(t)\gamma(t) + b(t) - (a(t)\phi(t) + b(t)) = a(t)(\gamma(t) - \phi(t)) = a(t)h(t).$$

Es ist  $h$  also die nach 17.8 eindeutige Lösung des Anfangswertproblems

$$\begin{cases} y'(t) &= a(t)y(t) \\ y(t_0) &= 0 \end{cases}$$

und diese ist nach 17.8  $h = 0$ . Also ist  $\gamma = \phi|_I$ .

**Bemerkung 17.11** (a) In der Situation von 17.10 sind die auf  $J$  definierten Lösungen der zugehörigen homogenen linearen Differentialgleichung  $y'(t) = a(t)y(t)$  von der Form

$$c e^{\int_{t_0}^t a(s) ds}$$

mit einer Konstanten  $c \in \mathbb{R}$  (dem Anfangswert  $y_0$  zur Zeit  $t_0$ ). Beim Ansatz der "Variation der Konstanten" wird die Konstante  $c$  durch eine Funktion  $c(t)$  ersetzt.

(b) Es dürfte vielen leichter fallen, sich den Ansatz der Variation der Konstanten zu merken und anzuwenden als die fertige Endformel aus 17.10 (die sich aus dem Ansatz jederzeit wieder herleiten lässt).

**Beispiel 17.12** (Lineare Differentialgleichungen erster Ordnung mit konstanten Koeffizienten). Wir betrachten den Spezialfall einer linearen Differentialgleichung der Form

$$y'(t) = a y(t) + b(t),$$

wobei  $b: J \rightarrow \mathbb{R}$  eine stetige Funktion auf einem nicht entarteten Intervall  $J \subseteq \mathbb{R}$  ist und  $a \in \mathbb{R}$  eine feste reelle Zahl. Gegeben  $t_0 \in J$  und  $y_0 \in \mathbb{R}$  ist die auf  $J$  definierte Lösung des Anfangswertproblems

$$\begin{cases} y'(t) &= a y(t) + b(t) \\ y(t_0) &= y_0 \end{cases}$$

dann die Funktion  $\gamma: J \rightarrow \mathbb{R}$ ,

$$\gamma(t) = \left( y_0 + \int_{t_0}^t b(\tau) e^{-(\tau-t_0)a} d\tau \right) e^{(t-t_0)a}.$$

Dies können wir (wenn gewünscht) noch umformen zu

$$\gamma(t) = e^{(t-t_0)a} y_0 + \int_{t_0}^t e^{(t-\tau)a} b(\tau) d\tau. \quad (89)$$

## Differentialgleichungen in getrennten Veränderlichen

**17.13** Es seien  $J \subseteq \mathbb{R}$  und  $K \subseteq \mathbb{R}$  offene Intervalle,  $g: J \rightarrow \mathbb{R}$  eine stetige Funktion und  $h: K \rightarrow \mathbb{R}$  eine stetige Funktion derart, dass  $h(y) > 0$  für alle  $y \in K$  oder  $h(y) < 0$  für alle  $y \in K$ . Gegeben  $t_0 \in I$  und  $y_0 \in K$  wollen wir das Anfangswertproblem

$$\begin{cases} y'(t) &= g(t) h(y(t)) \\ y(t_0) &= y_0 \end{cases}$$

lösen.

**17.14** Kochrezept. Man schreibe

$$\frac{dy}{dt} = g(t) h(y),$$

gehe formal zu

$$\frac{dy}{h(y)} = g(t) dt$$

über und integriere beide Seiten:

$$\int_{y_0}^{y(t)} \frac{1}{h(x)} dx = \int_{t_0}^t g(s) ds.$$

Man berechne die Integrale und löse nach  $y(t)$  auf.

Eine mathematische Begründung (die in der Vorlesung übersprungen wird) finden Sie in einem Anhang am Kapitelende. Unter der oben gemachten Voraussetzung, dass  $h(y)$  stets positiv oder stets negativ ist, sind die Lösungen des Anfangswertproblems übrigens wieder eindeutig.

**Beispiel 17.15** *Lösen Sie das Anfangswertproblem*

$$\begin{cases} y'(t) = t y(t)^2 \\ y(0) = 1. \end{cases}$$

*Lösung.* Wir schreiben

$$\frac{dy}{dt} = t y^2$$

und stellen formal um zu

$$\frac{dy}{y^2} = t dt.$$

Wir integrieren beide Seiten,

$$\int_{y_0}^{y(t)} \frac{1}{x^2} dx = \int_{t_0}^t s ds,$$

also mit  $t_0 = 0$ ,  $y_0 = 1$

$$\int_1^{y(t)} \frac{1}{x^2} dx = \int_0^t s ds.$$

Ausrechnen der Integrale mit Hilfe von Stammfunktionen liefert

$$\left[ -\frac{1}{x} \right]_1^{y(t)} = [s^2/2]_0^t,$$

also

$$1 - \frac{1}{y(t)} = t^2/2.$$

Auflösen nach  $y(t)$  ergibt

$$y(t) = \frac{1}{1 - t^2/2}.$$

## Anhang: Ein Beispiel ohne lokale Eindeutigkeit

**Beispiel 17.16** Das Anfangswertproblem

$$\begin{cases} y'(t) &= 3\sqrt[3]{y(t)^2} \\ y(0) &= 0 \end{cases}$$

hat die Lösungen  $\phi: \mathbb{R} \rightarrow \mathbb{R}, t \mapsto 0$  und  $\psi: \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t^3$  (Nachrechnen!). Es ist  $\phi \neq \psi$  und sogar  $\phi|_{]-\varepsilon, \varepsilon[} \neq \psi|_{]-\varepsilon, \varepsilon[}$  für jedes  $\varepsilon > 0$ .

## Anhang: Details zu DGLn in getrennten Veränderlichen

Der Abschnitt wird in der Vorlesung übersprungen.

**17.17** (Differentialgleichungen in getrennten Veränderlichen - Details). *Es seien  $J \subseteq \mathbb{R}$  und  $K \subseteq \mathbb{R}$  nicht entartete Intervalle,  $g: J \rightarrow \mathbb{R}$  eine stetige Funktion und  $h: K \rightarrow \mathbb{R}$  eine stetige Funktion derart, dass  $h(y) > 0$  für alle  $y \in K$  oder  $h(y) < 0$  für alle  $y \in K$ . Gegeben  $t_0 \in I$  und  $y_0 \in K$  definieren wir  $C^1$ -Funktionen*

$$G: J \rightarrow \mathbb{R}, \quad G(t) := \int_{t_0}^t g(s) ds$$

und

$$H: K \rightarrow \mathbb{R}, \quad H(y) := \int_{y_0}^y \frac{1}{h(x)} dx.$$

*Dann ist  $H$  eine streng monotone Funktion,  $H(K)$  ein nicht entartetes Intervall und  $H: K \rightarrow H(K)$  ein  $C^1$ -Diffeomorphismus. Weiter gilt:*

- (a) *Auf einem nicht entarteten Intervall  $I \subseteq J$  mit  $t_0 \in I$  gibt es genau dann eine Lösung  $\gamma: I \rightarrow \mathbb{R}$  des Anfangswertproblems*

$$\begin{cases} y' &= g(t)h(y) \\ y(t_0) &= y_0, \end{cases}$$

*wenn  $G(I) \subseteq H(K)$ . Die Lösung ist dann eindeutig und gegeben durch  $\gamma(t) = H^{-1}(G(t))$  für alle  $t \in I$ .*

- (b) *Ist  $K \subseteq \mathbb{R}$  ein offenes Intervall, so existiert ein in  $J$  offenes Intervall  $I \subseteq J$  mit  $t_0 \in I$  und eine Lösung  $\gamma: I \rightarrow \mathbb{R}$  des Anfangswertproblems aus (a).*

Beweis. Nach dem Zwischenwertsatz ist  $h(K)$  ein Intervall. Ist  $h(y) > 0$  für alle  $y \in K$ , so ist wegen  $H' = 1/h$  die  $C^1$ -Funktion  $H$  streng monoton wachsend; im Fall  $h(K) \subseteq ]-\infty, 0[$  ist  $H$  streng fallend. Zudem wissen wir aus der Mathematik 1, dass  $H^{-1}$  stetig differenzierbar ist mit

$$(H^{-1})'(H(y)) = \frac{1}{H'(y)} \quad \text{für alle } y \in K. \quad (90)$$

(a) Ist  $\gamma: I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems, so ist  $I \subseteq J$  und  $\gamma(t) \in K$  für alle  $t \in I$ . Für alle  $t \in I$  gilt weiter

$$\gamma'(t) = g(t)h(\gamma(t)),$$

also

$$\frac{\gamma'(t)}{h(\gamma(t))} = g(t). \quad (91)$$

Da  $(H \circ \gamma)'(t) = H'(\gamma(t))\gamma'(t) = \frac{\gamma'(t)}{h(\gamma(t))}$ , ist  $H \circ \gamma$  eine Stammfunktion der linken Seite von (91). Integration von  $t_0$  bis  $t$  ergibt

$$[H \circ \gamma]_{t_0}^t = [G]_{t_0}^t,$$

also

$$H(\gamma(t)) = G(t)$$

(wobei  $H(y_0) = G(t_0) = 0$  benutzt wurde). Insbesondere ist  $G(t) = H(\gamma(t)) \in H(K)$  für alle  $t \in I$ , also  $G(I) \subseteq H(K)$  und die Lösung  $\gamma$  erfüllt  $\gamma(t) = H^{-1}(G(t))$  für alle  $t \in I$ .

Ist umgekehrt  $I \subseteq J$  ein nicht entartetes Intervall mit  $t_0 \in I$  und  $G(I) \subseteq H(K)$ , so ist

$$\gamma: I \rightarrow \mathbb{R}, \quad t \mapsto H^{-1}(G(t))$$

eine  $C^1$ -Funktion mit  $\gamma(t_0) = H^{-1}(G(t_0)) = H^{-1}(0) = y_0$ . Wegen (90) ist

$$\gamma'(t) = (H^{-1})'(G(t))G'(t) = (H^{-1})'(H(\gamma(t)))G'(t) = \frac{1}{H'(\gamma(t))}g(t) = h(\gamma(t))g(t)$$

und somit  $\gamma$  eine Lösung des Anfangswertproblems aus (a).

(b) Ist  $K$  offen in  $\mathbb{R}$ , so auch  $H(K)$ . Also ist  $H(K)$  eine offene 0-Umgebung in  $\mathbb{R}$ . Da  $G: J \rightarrow \mathbb{R}$  stetig ist und  $G(t_0) = 0$ , ist  $G^{-1}(H(K))$  eine Umgebung von  $t_0$  in  $J$  und enthält als solche ein in  $J$  offenes Intervall  $I$  mit  $t_0 \in I$ .

## 18 Lineare Differentialgleichungssysteme erster Ordnung

Über Definition 17.7 hinaus betrachten wir nun auch lineare Differentialgleichungen für vektorwertige Funktionen.

**Definition 18.1** Eine *lineare Differentialgleichung* erster Ordnung ist eine Differentialgleichung der Form

$$y'(t) = A(t)y(t) + b(t), \quad (92)$$

wobei  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall ist,  $A: J \rightarrow \mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$  eine stetige matrixwertige Funktion und  $b: J \rightarrow \mathbb{R}^n$  eine stetige vektorwertige Funktion.<sup>20</sup> Ist  $b(t) = 0$  für alle  $t \in J$ , die DGL also von der Form

$$y'(t) = A(t)y(t),$$

so spricht man von einer *homogenen* linearen Differentialgleichung erster Ordnung; andernfalls heißt die lineare Differentialgleichung *inhomogen*. Ist  $A(t)$  eine konstante Funktion in  $t$ , also eine feste Matrix  $A$ , so nennt man (92) eine lineare Differentialgleichung *mit konstanten Koeffizienten*.

Betrachtet wird also die Differentialgleichung

$$y'(t) = f(t, y(t))$$

in  $\mathbb{R}^n$  mit  $f: J \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $(t, y) \mapsto A(t)y + b(t)$ .

Ist  $n > 1$ , werden solche Differentialgleichungen traditionell auch *Differentialgleichungssysteme* genannt. In diesem Kapitel beschränken wir uns auf den Fall von linearen Differentialgleichungssystemen *mit konstanten Koeffizienten*. Ein entscheidendes Hilfsmittel für uns ist die *Matrixexponentialfunktion*.

**18.2** (Matrixexponentialfunktion). Für jede Matrix  $A \in \mathbb{R}^{n \times n}$  konvergiert die Reihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

in  $\mathbb{R}^{n \times n}$ . Wir definieren

$$e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

---

<sup>20</sup>Es ist also  $A(t) = (a_{ij}(t))_{i,j=1,\dots,n}$  mit stetigen Funktionen  $a_{ij}: J \rightarrow \mathbb{R}$  für  $i, j \in \{1, \dots, n\}$  und  $b = (b_1, \dots, b_n)$  mit stetigen Funktionen  $b_1, \dots, b_n: J \rightarrow \mathbb{R}$ .

Konvergenz von Matrizen ist dabei wie folgt zu verstehen.

**18.3** Identifizieren wir eine quadratische Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  mit dem Spaltenvektor

$$(a_{11}, \dots, a_{n1}, a_{12}, \dots, a_{n2}, \dots, a_{1n}, \dots, a_{nn})^T \in \mathbb{R}^{n^2},$$

können wir den Vektorraum  $\mathbb{R}^{n \times n}$  aller  $n \times n$ -Matrizen mit  $\mathbb{R}^{n^2}$  identifizieren.

- Wir können daher von Konvergenz einer Folge  $(A_k)_{k \in \mathbb{N}_0}$  von  $n \times n$ -Matrizen gegen eine  $n \times n$ -Matrix  $A = (a_{ij})$  sprechen: Verlangt ist  $|A_k - A| \rightarrow 0$  für  $k \rightarrow \infty$ , oder äquivalent Konvergenz des  $(i, j)$ -Eintrags  $(A_k)_{ij}$  von  $A_k$  gegen  $a_{ij}$  für  $k \rightarrow \infty$ , für alle  $i, j \in \{1, \dots, n\}$ .
- Ist  $(A_k)_{k \in \mathbb{N}_0}$  eine Folge von  $n \times n$ -Matrizen, können wir die Anfangssummen  $S_m := \sum_{k=0}^m A_k \in \mathbb{R}^{n \times n}$  betrachten für  $m \in \mathbb{N}_0$ . Konvergiert die Folge  $(S_m)_{m \in \mathbb{N}_0}$  gegen eine Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ , schreiben wir

$$A = \sum_{k=0}^{\infty} A_k := \lim_{m \rightarrow \infty} \sum_{k=0}^m A_k.$$

Schreiben wir  $(A_k)_{ij}$  für den  $(i, j)$ -Eintrag von  $A_k$ , bedeutet dies also  $a_{ij} = \sum_{k=0}^{\infty} (A_k)_{ij}$  für alle  $i, j \in \{1, \dots, n\}$ .

Die Identifizierung von  $\mathbb{R}^{n \times n}$  mit  $\mathbb{R}^{n^2}$  erlaubt uns zudem, von stetigen bzw. stetig differenzierbaren Matrizenwertigen Funktionen zu sprechen.

- Eine matrixwertige Funktion  $J \rightarrow \mathbb{R}^{n \times n}$ ,  $t \mapsto A(t) = (a_{ij}(t))$  auf einem nicht entarteten Intervall ist stetig genau dann, wenn jede Komponente  $a_{ij}: J \rightarrow \mathbb{R}$  stetig ist.
- Eine matrixwertige Funktion  $J \rightarrow \mathbb{R}^{n \times n}$ ,  $t \mapsto A(t) = (a_{ij}(t))$  auf einem nicht entarteten Intervall ist genau dann stetig differenzierbar, wenn jede Komponente  $a_{ij}: J \rightarrow \mathbb{R}$  stetig differenzierbar ist; in diesem Fall setzen wir  $A'(t) = (a'_{ij}(t)) \in \mathbb{R}^{n \times n}$  für  $t \in J$ .

Die in 18.2 und gleich in 18.8 behaupteten Fakten benutzen wir zunächst als Black Box. Am Ende des Kapitels werden sie nachgewiesen.

**Beispiel 18.4** Für die 0-Matrix  $\mathbf{0} \in \mathbb{R}^{n \times n}$  ist  $e^{\mathbf{0}} = \mathbf{1}_n$  die Einheitsmatrix in  $\mathbb{R}^{n \times n}$ .

Es ist nämlich  $\mathbf{0}^0 = \mathbf{1}_n$  per Definition und  $\mathbf{0}^k = \mathbf{0}$  für alle  $k \in \mathbb{N}$ . Also ist

$$\sum_{k=0}^m \frac{1}{k!} \mathbf{0}^k = \mathbf{1}_n + \mathbf{0} = \mathbf{1}_n$$

für alle  $m \in \mathbb{N}_0$ , mit Limes  $\mathbf{1}_n$  für  $m \rightarrow \infty$ .

**Beispiel 18.5** Für die Diagonalmatrix

$$D = \text{diag}(2, 7) = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}$$

ist

$$e^D = \text{diag}(e^2, e^7) = \begin{pmatrix} e^2 & 0 \\ 0 & e^7 \end{pmatrix}$$

und für  $t \in \mathbb{R}$

$$e^{tD} = \text{diag}(e^{2t}, e^{7t}) = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{7t} \end{pmatrix}.$$

Es gilt nämlich

$$\begin{aligned} \sum_{k=0}^m \frac{1}{k!} \begin{pmatrix} 2t & 0 \\ 0 & 7t \end{pmatrix}^k &= \sum_{k=0}^m \frac{1}{k!} \begin{pmatrix} (2t)^k & 0 \\ 0 & (7t)^k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=0}^m \frac{1}{k!} (2t)^k & 0 \\ 0 & \sum_{k=0}^m \frac{1}{k!} (7t)^k \end{pmatrix} \rightarrow \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{7t} \end{pmatrix} \end{aligned}$$

für  $m \rightarrow \infty$ . Analog sieht man, dass

$$e^{tD} = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$$

gilt für  $D = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ , für alle  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  und  $t \in \mathbb{R}$ .

**Beispiel 18.6** Für die Matrix

$$N := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

gilt

$$e^N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad e^{tN} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{für alle } t \in \mathbb{R}.$$

Es ist nämlich  $N^2 = \mathbf{0}$ , also auch  $N^k = \mathbf{0}$  für alle  $k \geq 2$  und somit

$$\sum_{k=0}^m \frac{1}{k!} (tN)^k = \sum_{k=0}^m \frac{t^k}{k!} N^k = \sum_{k=0}^1 \frac{t^k}{k!} N^k = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

für alle  $m \geq 1$ . Die rechte Seite ist von  $m \geq 1$  unabhängig und konvergiert gegen

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

für  $m \rightarrow \infty$ .

**Beispiel 18.7** Man nennt ein Matrix  $N \in \mathbb{R}^{n \times n}$  *nilpotent*, wenn  $N^\ell = \mathbf{0}$  für ein  $\ell \in \mathbb{N}$  (die vorige  $2 \times 2$ -Matrix war also nilpotent mit  $\ell = 2$ ). Dann sind die Komponenten von

$$e^{tN} = \sum_{k=0}^{\ell-1} \frac{t^k}{k!} N^k$$

einfach Polynome in  $t$  vom Grad  $\leq \ell - 1$ .

**18.8** (Eigenschaften der Matrixexponentialfunktion) Für  $A \in \mathbb{R}^{n \times n}$  gilt:

- (a) Für alle  $s, t \in \mathbb{R}$  ist  $e^{sA} e^{tA} = e^{(s+t)A}$ .
- (b) Die  $n \times n$ -Matrix  $e^A$  ist invertierbar, mit inverser Matrix  $(e^A)^{-1} = e^{-A}$ .
- (c) Die Abbildung  $\mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ ,  $t \mapsto e^{tA}$  ist stetig differenzierbar mit

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A \quad \text{für alle } t \in \mathbb{R}.$$

- (d) (Vertauschende Matrizen). Sind  $A, B \in \mathbb{R}^{n \times n}$  Matrizen derart, dass  $AB = BA$ , so gilt  $e^{A+B} = e^A e^B$ .

Daraus folgt:

**18.9** (Homogene lineare Differentialgleichungssysteme 1. Ordnung mit konstanten Koeffizienten). Für  $A \in \mathbb{R}^{n \times n}$ ,  $t_0 \in \mathbb{R}$  und  $y_0 \in \mathbb{R}^n$  gilt folgendes:

(a) (Globale Existenz von Lösungen). *Die Funktion*

$$\phi: \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto e^{(t-t_0)A} y_0$$

*ist eine auf ganz  $\mathbb{R}$  definierte Lösung des Anfangswertproblems*

$$\begin{cases} y'(t) &= A y(t) \\ y(t_0) &= y_0. \end{cases}$$

(b) (Eindeutigkeit). *Ist  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall mit  $t_0 \in J$  und auch  $\gamma: J \rightarrow \mathbb{R}^n$  eine Lösung des Anfangswertproblems aus (a), so gilt  $\gamma = \phi|_J$ .*

Zum Nachweis benutzen wir folgende Tatsache.

**18.10** (Produktregel für Matrix mal Vektor). *Es sei  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall,  $J \rightarrow \mathbb{R}^{n \times n}$ ,  $t \mapsto A(t)$  eine stetig differenzierbare Matrixwertige Funktion und  $J \rightarrow \mathbb{R}^n$ ,  $t \mapsto v(t) = (v_1(t), \dots, v_n(t))$  eine stetig differenzierbare vektorwertige Funktion. Dann ist die vektorwertige Funktion*

$$J \rightarrow \mathbb{R}^n, \quad t \mapsto A(t)v(t)$$

*stetig differenzierbar und*

$$\frac{d}{dt} (A(t)v(t)) = A'(t)v(t) + A(t)v'(t).$$

Beweis: Sei  $A(t) = (a_{ij}(t))$ . Die  $i$ te Komponente des Vektors  $A(t)v(t) \in \mathbb{R}^n$  ist

$$\sum_{j=1}^n a_{ij}(t)v_j(t).$$

Mit der Produktregel erhalten wir die Ableitung

$$\sum_{j=1}^n (a'_{ij}(t)v_j(t) + a_{ij}(t)v'_j(t)) = \sum_{j=1}^n a'_{ij}(t)v_j(t) + \sum_{j=1}^n a_{ij}(t)v'_j(t).$$

Diese stimmt mit der  $i$ ten Komponente von  $A'(t)v(t) + A(t)v'(t)$  überein.

Beweis von 18.9 (a): Nach 18.8 (c) und 18.10 ist

$$\phi: \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto e^{(t-t_0)A} y_0$$

stetig differenzierbar mit

$$\phi'(t) = A e^{(t-t_0)A} y_0 = A \phi(t).$$

Da zudem  $\phi(t_0) = e^0 y_0 = \mathbf{1}_n y_0 = y_0$ , ist  $\phi$  eine Lösung des vorgelegten Anfangswertproblems.

(b) Wir betrachten die Hilfsfunktion

$$h: \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto e^{-(t-t_0)A} \gamma(t),$$

welche nach 18.10 stetig differenzierbar ist mit

$$h'(t) = -e^{-(t-t_0)A} \underbrace{A\gamma(t)}_{=\gamma'(t)} + e^{-(t-t_0)A} \gamma'(t) = 0$$

für alle  $t \in J$ . Also ist  $h$  konstant mit Funktionswert

$$h(t_0) = e^0 \gamma(t_0) = \gamma(t_0) = y_0.$$

Da  $e^{-(t-t_0)A} = (e^{(t-t_0)A})^{-1}$ , liefert Multiplikation von  $h(t) = y_0$  von links mit  $e^{(t-t_0)A}$

$$\gamma(t) = e^{(t-t_0)A} y_0,$$

**Beispiel 18.11** Man finde die auf ganz  $\mathbb{R}$  definierte Lösung des Anfangswertproblems

$$\begin{aligned} y_1'(t) &= y_2(t) \\ y_2'(t) &= 0 \\ y_1(0) &= y_2(0) = 1. \end{aligned}$$

Lösung: Wir suchen die Lösung von

$$\begin{cases} y'(t) = Ay \\ y(0) = y_0 \end{cases}$$

mit

$$A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Nach 18.9 und Beispiel 18.6 ist die Lösung

$$y(t) = e^{tA} y_0 = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+t \\ 1 \end{pmatrix}.$$

Das hätten wir natürlich auch von Hand geschafft: Wegen  $y_2' = 0$  ist  $y_2$  konstant mit Wert 1. Aus  $y_1' = y_2 = 1$  folgt dann  $y_1(t) = y_1(0) + t = 1 + t$ .

Auch das folgende Beispiel ist noch nicht so spannend, da man direkt die DGLn für  $y_1$  und  $y_2$  lösen könnte.

**Beispiel 18.12** Man finde die auf ganz  $\mathbb{R}$  definierte Lösung des Anfangswertproblems

$$\begin{aligned}y_1'(t) &= 2y_1(t) \\y_2'(t) &= -y_2(t) \\y_1(0) &= 1 \\y_2(0) &= 3.\end{aligned}$$

Lösung: Wir suchen die Lösung von

$$\begin{cases}y'(t) = Ay \\y(0) = y_0\end{cases}$$

mit

$$A := \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \quad y_0 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Nach 18.9 und Beispiel 18.5 ist die Lösung

$$y(t) = e^{tA} y_0 = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ 3e^{-t} \end{pmatrix}.$$

Der folgende Sachverhalt ermöglicht, schwierigere Systeme zu behandeln.

**18.13** Ist  $B$  eine  $n \times n$ -Matrix und  $S$  eine invertierbare  $n \times n$ -Matrix, so ist

$$e^{SBS^{-1}} = Se^BS^{-1}.$$

Also ist  $e^{tSBS^{-1}} = Se^{tB}S^{-1}$  für alle  $t \in \mathbb{R}$ .

Wir wissen nämlich, dass jede lineare Abbildung  $\mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$  stetig ist. Die Abbildung

$$\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, \quad A \mapsto SA$$

entspricht einer solchen, wenn wir  $\mathbb{R}^{n \times n}$  mit  $\mathbb{R}^{n^2}$  identifizieren. Analog ist die Abbildung  $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ,  $A \mapsto AS^{-1}$  stetig. Also ist

$$\begin{aligned}Se^BS^{-1} &= S \left( \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{1}{k!} B^k \right) S^{-1} = \lim_{m \rightarrow \infty} S \left( \sum_{k=0}^m \frac{1}{k!} B^k \right) S^{-1} \\ &= \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{1}{k!} SB^k S^{-1} = \sum_{k=0}^{\infty} \frac{1}{k!} (SBS^{-1})^k = e^{SBS^{-1}}.\end{aligned}$$

Hierbei wurde benutzt, dass

$$(SBS^{-1})^k = (SBS^{-1})(SBS^{-1}) \cdots (SBS^{-1}) = SBB \cdots BS = SB^k S^{-1},$$

da sich im Inneren  $S^{-1}$  und  $S$  stets kürzen lassen,  $S^{-1}S = \mathbf{1}_n$ .

**Beispiel 18.14** Finden Sie die auf ganz  $\mathbb{R}$  definierte Lösung des Anfangswertproblems

$$\begin{aligned} y_1'(t) &= 5y_1(t) - 2y_2(t) \\ y_2'(t) &= -\frac{7}{2}y_1(t) + \frac{7}{2}y_2(t) \\ y_1(0) &= 1 \\ y_2(0) &= 2. \end{aligned}$$

Lösung: Wir suchen die Lösung von

$$\begin{cases} y'(t) = Ay \\ y(0) = y_0 \end{cases}$$

mit

$$A := \begin{pmatrix} 5 & -2 \\ -\frac{7}{2} & \frac{7}{2} \end{pmatrix}, \quad y_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Berechnung der Eigenwerte und Eigenvektoren zeigt, dass  $A$  diagonalisierbar ist; es ist

$$A = SDS^{-1}$$

mit

$$S = \begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \quad \text{und} \quad D = \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix},$$

wobei

$$S^{-1} := \begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix}.$$

Nach 18.9, 18.13 und Beispiel 18.5 ist die Lösung

$$\begin{aligned} y(t) &= e^{tA} y_0 = e^{S(tD)S^{-1}} y_0 = S e^{tD} S^{-1} y_0 \\ &= \begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{7t} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2}e^{3t} - \frac{1}{2}e^{7t} \\ \frac{3}{2}e^{3t} + \frac{1}{2}e^{7t} \end{pmatrix}. \end{aligned}$$

**18.15** Ist  $A \in \mathbb{R}^{n \times n}$  eine diagonalisierbare Matrix, so ist  $A = SDS^{-1}$  mit einer invertierbaren  $n \times n$ -Matrix  $S$  und einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Für  $t \in \mathbb{R}$  ist nach 18.13 und Beispiel 18.5 dann also

$$e^{tA} = e^{StDS^{-1}} = Se^{tD}S^{-1},$$

wobei  $e^{tD} = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$ .

Was passiert, wenn  $A$  nicht diagonalisierbar ist? Kann auch dann  $e^{tA}$  im Prinzip immer berechnet werden?

Dies ist in der Tat möglich, mit Hilfe der sogenannten Jordanschen Normalform von  $A$ . Wir überspringen den Sachverhalt in der Vorlesung und fahren mit 18.17 fort.

**Bemerkung 18.16** Ist  $A \in \mathbb{R}^{n \times n}$  beliebig (also nicht notwendig diagonalisierbar), so gibt es immer noch eine komplexe invertierbare Matrix  $S \in \mathbb{C}^{n \times n}$  derart, dass

$$A = SJS^{-1}$$

mit einer gewissen Block-Diagonalmatrix  $J = \text{diag}(J_1, \dots, J_m)$  (einer Matrix  $J$  in "Jordanscher Normalform"). Hierbei ist

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & & 0 \\ & & \ddots & \ddots & \\ & & & \lambda_k & 1 \\ & & & 0 & \lambda_k \end{pmatrix}$$

eine  $n_k \times n_k$ -Matrix mit einem komplexen Eigenwert  $\lambda_k$  von  $A$  auf der Diagonalen und Einsen auf der Parallelen zur Diagonalen (ein sogenanntes "Jordan-Kästchen"). Es ist also

$$J_k = \lambda_k \mathbf{1}_{n_k} + N_{n_k}$$

mit

$$N_{n_k} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & & 0 \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

Die Matrix  $N_{n_k}$  ist nilpotent mit  $(N_{n_k})^{n_k} = \mathbf{0}$ . Da die Matrizen  $t\lambda_k \mathbf{1}_{n_k}$  und  $tN_{n_k}$  miteinander vertauschen, ist dann

$$e^{tJ_k} = e^{t\lambda_k \mathbf{1}_{n_k} + tN_{n_k}} = e^{t\mathbf{1}_{n_k}} e^{tN_{n_k}} = e^{\lambda_k t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n_k-1}}{(n_k-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{n_k-2}}{(n_k-2)!} \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & t \\ & & & 0 & 1 \end{pmatrix}.$$

Weiter ist

$$e^{tA} = S \operatorname{diag}(e^{tJ_1}, \dots, e^{tJ_m}) S^{-1}$$

mit  $e^{tJ_1}, \dots, e^{tJ_m}$  wie gerade berechnet.

Etwas mehr Details: Ist

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{\ell \times \ell},$$

so rutscht in den Potenzen  $N^j$  die Nebendiagonale mit Einsen in jedem Schritt eins weiter in die rechte obere Ecke; ist etwa  $\ell = 4$ , so ist

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^4 = \mathbf{0}$$

und somit

$$e^{tN} = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 \\ 0 & 1 & t & t^2/2 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Für allgemeines  $\ell \in \mathbb{N}$  ist analog

$$e^{tN} = \sum_{j=0}^{\ell-1} \frac{1}{j!} t^j N^j = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 & \cdots & t^{\ell-1}/(\ell-1)! \\ 0 & 1 & t & t^2/2 & \cdots & t^{\ell-2}/(\ell-2)! \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & t \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die Methode der ‘‘Variation der Konstanten’’ funktioniert in geeigneter Form auch f ur Systeme linearer Differentialgleichungen.

**18.17** (Inhomogene lineare Systeme erster Ordnung mit konstanten Koeffizienten). *Es seien  $A$  eine  $n \times n$ -Matrix,  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall,  $b: J \rightarrow \mathbb{R}^n$  eine stetige vektorwertige Funktion und  $t_0 \in J$ ,  $y_0 \in \mathbb{R}^n$ . Dann gilt:*

- (a) (Globale Existenz von L osungen). *Die auf ganz  $J$  definierte Funktion  $\phi: J \rightarrow \mathbb{R}^n$ ,*

$$\phi(t) = e^{(t-t_0)A} y_0 + \int_{t_0}^t e^{(t-s)A} b(s) ds \quad (93)$$

*ist eine L osung des Anfangswertproblems*

$$\begin{cases} y'(t) &= A y(t) + b(t), \\ y(t_0) &= y_0. \end{cases}$$

- (b) (Eindeutigkeit). *Ist  $I \subseteq J$  ein nicht entartetes Intervall mit  $t_0 \in I$  und auch  $\gamma: I \rightarrow \mathbb{R}^n$  eine L osung des Anfangswertproblems aus (a), so ist  $\gamma = \phi|_I$ .*

Nachweis: F ur die L osung  $\phi: J \rightarrow \mathbb{R}^n$  des obigen Anfangswertproblems machen wir den Ansatz

$$\phi(t) = e^{(t-t_0)A} c(t)$$

mit einer  $C^1$ -Funktion  $c: J \rightarrow \mathbb{R}^n$  (und werden sehen, dass dies funktioniert!). Die Anfangsbedingung ist  aquivalent zu

$$y_0 = \phi(t_0) = c(t_0).$$

Nun ist

$$\phi'(t) = \left( \frac{d}{dt} e^{(t-t_0)A} \right) c(t) + e^{(t-t_0)A} c'(t) = A e^{(t-t_0)A} c(t) + e^{(t-t_0)A} c'(t).$$

Einsetzen in die inhomogene DGL f uhrt auf die  aquivalente Bedingung

$$A e^{(t-t_0)A} c(t) + e^{(t-t_0)A} c'(t) = A e^{(t-t_0)A} c(t) + b(t),$$

also

$$e^{(t-t_0)A} c'(t) = b(t),$$

also

$$c'(t) = e^{(t_0-t)A} b(t).$$

Für die Stammfunktion erfordert dies

$$c(t) = c(t_0) + \int_{t_0}^t e^{(t_0-s)A} b(s) ds = y_0 + \int_{t_0}^t e^{(t_0-s)A} b(s) ds.$$

Also ist

$$\begin{aligned} \phi(t) &= e^{(t-t_0)A} c(t) = e^{(t-t_0)A} y_0 + e^{(t-t_0)A} \int_{t_0}^t e^{(t_0-s)A} b(s) ds \\ &= e^{(t-t_0)A} y_0 + \int_{t_0}^t \underbrace{e^{(t-t_0)A} e^{(t_0-s)A}}_{e^{(t-t_0+t_0-s)A} = e^{(t-s)A}} b(s) ds \end{aligned}$$

von der behaupteten Form.

(b) Die  $C^1$ -Funktion  $h: I \rightarrow \mathbb{R}^n$ ,  $t \mapsto \gamma(t) - \phi(t)$  erfüllt

$$h'(t) = \gamma'(t) - \phi'(t) = A\gamma(t) + b(t) - (A\phi(t) + b(t)) = A(\gamma(t) - \phi(t)) = Ah(t),$$

da  $\phi$  und  $\gamma$  Lösungen der inhomogenen DGL  $y'(t) = Ay(t) + b(t)$  sind. Da  $h(t_0) = y_0 - y_0 = 0$ , ist  $h$  also eine Lösung des Anfangswertproblems

$$\begin{cases} y'(t) &= Ay(t) \\ y(t_0) &= 0. \end{cases}$$

Nach 18.9 ist  $h = 0$ , also  $\gamma = \phi|_I$ .

## Anhang: Beweise zur Matrixexponentialfunktion

Wir benutzen eine simple Tatsache.

**18.18** (Submultiplikativität der Operatornorm). *Sind  $A, B \in \mathbb{R}^{n \times n}$ , so ist  $\|AB\|_{\text{op}} \leq \|A\|_{\text{op}} \|B\|_{\text{op}}$ . Insbesondere gilt  $\|A^k\|_{\text{op}} \leq (\|A\|_{\text{op}})^k$  für alle  $k \in \mathbb{N}_0$ .*

Für jedes  $v \in \mathbb{R}^n$  mit  $|v| \leq 1$  gilt nämlich

$$|ABv| \leq \|A\|_{\text{op}} |Bv| \leq \|A\|_{\text{op}} \|B\|_{\text{op}}.$$

Da die rechte Seite nicht von  $v$  abhängt, liefert Übergang zum Supremum über alle  $v$  die gewünschte Ungleichung,  $\|AB\|_{\text{op}} \stackrel{\text{def}}{=} \sup_v |ABv| \leq \|A\|_{\text{op}} \|B\|_{\text{op}}$ .

**18.19** (Subadditivität der Operatornorm / Dreiecksungleichung) Sind  $A, B \in \mathbb{R}^{n \times n}$ , so ist  $\|A + B\|_{\text{op}} \leq \|A\|_{\text{op}} + \|B\|_{\text{op}}$ .

Für jedes  $v \in \mathbb{R}^n$  mit  $|v| \leq 1$  ist nämlich

$$|(A + B)v| = |Av + Bv| \leq |Av| + |Bv| \leq \|A\|_{\text{op}} + \|B\|_{\text{op}}.$$

Übergang zum Supremum über  $v$  liefert  $\|A + B\|_{\text{op}} \stackrel{\text{def}}{=} \sup_v |(A + B)v| \leq \|A\|_{\text{op}} + \|B\|_{\text{op}}$ .

**18.20** (Absolut konvergente Reihen von Matrizen). Ist  $(A_k)_{k \in \mathbb{N}_0}$  eine Folge von Matrizen  $A_k \in \mathbb{R}^{n \times n}$  derart, dass  $\sum_{k=0}^{\infty} \|A_k\|_{\text{op}} < \infty$ , so konvergiert die Reihe  $\sum_{k=0}^{\infty} A_k$  in  $\mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$ , d.h. es existiert der Limes

$$\sum_{k=0}^{\infty} A_k := \lim_{K \rightarrow \infty} \sum_{k=0}^K A_k.$$

Für  $i, j \in \{1, \dots, n\}$  und eine Matrix  $B = (b_{ij}) \in \mathbb{R}^{n \times n}$  schreiben wir im Beweis auch  $B_{ij} := b_{ij}$ . Wir brauchen nur für alle  $i, j \in \{1, \dots, n\}$  die Konvergenz des  $(i, j)$ -Matrixeintrags

$$\left( \sum_{k=0}^K A_k \right)_{ij} = \sum_{k=0}^K (A_k)_{ij}$$

zu zeigen. Da  $|(A_k)_{ij}| \leq \|A_k\|_{\text{op}}$  nach 16.7, folgt die absolute Konvergenz der Reihe  $\sum_{k=0}^{\infty} (A_k)_{ij}$  mit dem Majorantenkriterium aus  $\sum_{k=0}^{\infty} \|A_k\|_{\text{op}} < \infty$ .

**18.21** (Konvergenz der Matrix-Exponentialreihe). Wir beweisen für  $A \in \mathbb{R}^{n \times n}$  die in 18.2 behauptete Konvergenz der Reihe

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

Nach 18.20 brauchen wir nur zu zeigen, dass  $\sum_{k=0}^{\infty} \left\| \frac{1}{k!} A^k \right\|_{\text{op}} < \infty$ . Dies ist erfüllt, denn nach 18.18 gilt

$$\left\| \frac{1}{k!} A^k \right\|_{\text{op}} = \frac{1}{k!} \|A^k\|_{\text{op}} \leq \frac{1}{k!} (\|A\|_{\text{op}})^k,$$

wobei  $\sum_{k=0}^{\infty} \frac{1}{k!} (\|A\|_{\text{op}})^k = e^{\|A\|_{\text{op}}} < \infty$ .

**18.22** (Cauchysche Produktformel für Reihen von Matrizen). Sind  $\sum_{k=0}^{\infty} A_k$  und  $\sum_{\ell=0}^{\infty} B_\ell$  Reihen von  $n \times n$ -Matrizen derart, dass  $\sum_{k=0}^{\infty} \|A_k\|_{\text{op}} < \infty$  und  $\sum_{\ell=0}^{\infty} \|B_\ell\|_{\text{op}} < \infty$ , so gilt auch  $\sum_{m=0}^{\infty} \|C_m\|_{\text{op}} < \infty$  für

$$C_m := \sum_{k=0}^m A_k B_{m-k} \in \mathbb{R}^{n \times n}.$$

Es gilt

$$\left( \sum_{k=0}^{\infty} A_k \right) \left( \sum_{\ell=0}^{\infty} B_\ell \right) = \sum_{m=0}^{\infty} C_m. \quad (94)$$

Von der Cauchyschen Produktformel für absolut konvergente Reihen von Zahlen wissen wir zunächst, dass

$$\sum_{m=0}^{\infty} c_m = \left( \sum_{k=0}^{\infty} \|A_k\|_{\text{op}} \right) \left( \sum_{\ell=0}^{\infty} \|B_\ell\|_{\text{op}} \right) < \infty$$

gilt für

$$c_m := \sum_{k=0}^m \|A_k\|_{\text{op}} \|B_{m-k}\|_{\text{op}}.$$

Nun gilt  $\|C_m\|_{\text{op}} \leq c_m$ ; wegen 18.18 und 18.19 ist nämlich

$$\left\| \sum_{k=0}^m A_k B_{m-k} \right\|_{\text{op}} \leq \sum_{k=0}^m \|A_k B_{m-k}\|_{\text{op}} \leq \sum_{k=0}^m \|A_k\|_{\text{op}} \|B_{m-k}\|_{\text{op}} = c_m.$$

Also ist  $\sum_{m=0}^{\infty} \|C_m\|_{\text{op}} \leq \sum_{m=0}^{\infty} c_m < \infty$ , die Matrixreihe  $\sum_{m=0}^{\infty} C_m$  nach 18.20 somit konvergent. Die noch zu zeigende Formel (94) gilt genau dann, wenn

$$\mathbf{0} = \left( \sum_{k=0}^{\infty} A_k \right) \left( \sum_{\ell=0}^{\infty} B_\ell \right) - \sum_{m=0}^{\infty} C_m,$$

also

$$\mathbf{0} = \lim_{N \rightarrow \infty} \left( \left( \sum_{k=0}^N A_k \right) \left( \sum_{\ell=0}^N B_\ell \right) - \sum_{m=0}^N C_m \right). \quad (95)$$

Für  $N \in \mathbb{N}$  sei  $\Delta_N$  die Menge aller  $(k, \ell) \in \{0, \dots, N\} \times \{0, \dots, N\}$  derart, dass  $k + \ell > N$ . Die Matrix zu gegebenem  $N$ , die im Limes in (95) steht, lässt sich schreiben als

$$M_N = \sum_{(k, \ell) \in \Delta_N} A_k B_\ell.$$

Da  $\sum_{k=0}^{\infty} \|A_k\|_{\text{op}} \sum_{\ell=0}^{\infty} \|B_\ell\|_{\text{op}} = \sum_{m=0}^{\infty} c_m$ , gilt

$$0 = \lim_{N \rightarrow \infty} \sum_{(k,\ell) \in \Delta_N} \|A_k\|_{\text{op}} \|B_\ell\|_{\text{op}}.$$

Da nach der Dreiecksungleichung

$$0 \leq \|M_N\|_{\text{op}} \leq \sum_{(k,\ell) \in \Delta_N} \|A_k\|_{\text{op}} \|B_\ell\|_{\text{op}},$$

folgt mit dem Sandwichkriterium  $\|M_N\|_{\text{op}} \rightarrow 0$  für  $N \rightarrow \infty$ , also  $M_N \rightarrow \mathbf{0}$ .

**18.23** (Beweis von 18.8).

(d) Nach 18.22 ist

$$e^A e^B = \left( \sum_{k=0}^{\infty} \frac{1}{k!} A^k \right) \left( \sum_{\ell=0}^{\infty} \frac{1}{\ell!} B^\ell \right) = \sum_{m=0}^{\infty} C_m$$

mit

$$\begin{aligned} C_m &= \sum_{k=0}^m \frac{1}{k!} A^k \frac{1}{(m-k)!} B^{m-k} = \frac{1}{m!} \sum_{k=0}^m \frac{m!}{k!(m-k)!} A^k B^{m-k} \\ &= \frac{1}{m!} \sum_{k=0}^m \binom{m}{k} A^k B^{m-k} \\ &= \frac{1}{m!} (A+B)^m. \end{aligned}$$

Beim Übergang zur letzten Zeile wurde hierbei der binomische Lehrsatz für  $(A+B)^m$  angewandt. Dieser ist hier gültig (und kann wie für Zahlen bewiesen werden), da  $AB = BA$  per Voraussetzung.

(a) Ist ein Spezialfall von (d).

(b) Nach (a) ist

$$\mathbf{1}_n = e^{\mathbf{0}} = e^{A+(-A)} = e^A e^{-A}.$$

Analog ist  $e^{-A} e^A = \mathbf{1}_n$ , die Matrix  $e^A$  also invertierbar mit  $e^{-A}$  als inverser.

(c) Für den  $(i, j)$ -Eintrag von  $e^{tA}$  gilt

$$(e^{tA})_{ij} = \left( \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k \right)_{ij} = \sum_{k=0}^{\infty} \frac{1}{k!} (A^k)_{ij} t^k.$$

Da wir konvergente Potenzreihen gliedweise ableiten dürfen, folgt

$$\begin{aligned} \frac{d}{dt} (e^{tA})_{ij} &= \sum_{k=1}^{\infty} \frac{1}{k!} (A^k)_{ij} k t^{k-1} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} (A^k)_{ij} t^{k-1} \\ &= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (A^{\ell+1})_{ij} t^{\ell} = \left( \sum_{\ell=0}^{\infty} \frac{1}{\ell!} t^{\ell} A^{\ell+1} \right)_{ij}, \end{aligned}$$

wobei  $\ell := k-1, k = \ell+1$  substituiert wurde. Also ist  $\frac{d}{dt} e^{tA} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} t^{\ell} A^{\ell+1} = A e^{tA} = e^{tA} A$ . Da Linksmultiplikation mit  $A$  stetig ist und linear, können wir  $A$  nämlich nach links aus der vorigen Summe herausziehen (ebenso nach rechts). Somit ist  $t \mapsto e^{tA}$  differenzierbar. Da  $A e^{tA}$  ebenso differenzierbar ist mit Ableitung  $A^2 e^{tA}$ , ist  $t \mapsto A e^{tA}$  insbesondere stetig, folglich  $t \mapsto e^{tA}$  stetig differenzierbar.

## 19 Differentialgleichungen zweiter Ordnung, insb. lineare mit konstanten Koeffizienten

In diesem Kapitel führen wir Differentialgleichungen zweiter Ordnung ein. Diese treten sehr natürlich in der Physik auf. Nach dem zweiten Newtonschen Gesetz über die Bewegung eines Körpers ist nämlich die Änderung des Impulses  $m x'(t)$  gleich der auf den Körper wirkenden Kraft  $F$ . Ist die Masse  $m$  unabhängig von  $t$ , so ist also

$$m x''(t) = F$$

mit der Beschleunigung  $x''(t)$ . Die Kraft hängt häufig nur von der Zeit  $t$ , der Position  $x(t) \in \mathbb{R}^3$  des Körpers und seiner Geschwindigkeit  $x'(t)$  ab. In diesem Fall erhalten wir die Differentialgleichung

$$m x''(t) = F(t, x(t), x'(t))$$

(wobei man in der Physik natürlich  $\dot{x}(t)$  und  $\ddot{x}(t)$  für erste und zweite Zeitableitungen schreibt).

Wir beschränken uns auf Differentialgleichungen für reellwertige Funktionen.

**Definition 19.1** Es sei  $f: U \rightarrow \mathbb{R}$  eine Funktion auf einer Teilmenge  $U \subseteq \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ . Eine *Differentialgleichung 2ter Ordnung mit rechter Seite*  $f$  ist eine Gleichung der Form

$$y''(t) = f(t, y(t), y'(t)). \quad (96)$$

Eine *Lösung* der Differentialgleichung (96) ist eine  $C^2$ -Funktion  $\phi: I \rightarrow \mathbb{R}$  auf einem nicht entarteten Intervall  $I \subseteq \mathbb{R}$  derart, dass

$$(t, \phi(t), \phi'(t)) \in U \quad \text{für alle } t \in I$$

und

$$\phi''(t) = f(t, \phi(t), \phi'(t)).$$

Sind  $(t_0, y_0, v_0) \in U$ , so nennen wir eine Funktion  $\phi: I \rightarrow \mathbb{R}$  eine *Lösung des Anfangswertproblems*

$$\begin{cases} y''(t) &= f(t, x(t), y'(t)) \\ y(t_0) &= y_0 \\ y'(t_0) &= v_0, \end{cases}$$

wenn  $\phi$  eine Lösung der Differentialgleichung  $y''(t) = f(t, y(t), y'(t))$  ist und zudem  $t_0 \in I$  ist und die Bedingungen  $\phi(t_0) = y_0$  und  $\phi'(t_0) = v_0$  erfüllt sind.

**Beispiel 19.2** Da  $\sin'(t) = \cos(t)$ ,  $\sin''(t) = \cos'(t) = -\sin(t)$ ,  $\sin(0) = 0$  und  $\sin'(0) = \cos(0) = 1$ , ist die Funktion  $\sin: \mathbb{R} \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems

$$\begin{cases} y''(t) &= -y(t) \\ y(0) &= 0 \\ y'(0) &= 1. \end{cases}$$

Im folgenden diskutieren wir ausschließlich lineare Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten, eine besonders wichtige Klasse von Differentialgleichungen.

**Definition 19.3** Eine *lineare Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten* ist eine Differentialgleichung der Form

$$y''(t) + a_1 y'(t) + a_0 y(t) = b(t)$$

mit  $a_0, a_1 \in \mathbb{R}$  und einer stetigen Funktion  $b: J \rightarrow \mathbb{R}$  auf einem nicht entarteten Intervall  $J \subseteq \mathbb{R}$ . Ist  $b = 0$ , so sprechen wir von einer *homogenen* linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten; andernfalls heißt die DGL *inhomogen*.

Wir betrachten also die Differentialgleichung

$$y''(t) = -a_0 y(t) - a_1 y'(t) + b(t),$$

das heißt

$$y''(t) = f(t, y(t), y'(t))$$

mit

$$f: J \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (t, y, v) \mapsto -a_0 y - a_1 v + b(t).$$

Im Falle homogener linearer DGLn können wir stets  $J = \mathbb{R}$  wählen.

**19.4** (Globale Existenz von Lösungen und Eindeutigkeit) *Es seien  $a_0, a_1 \in \mathbb{R}$  reelle Zahlen und  $b: J \rightarrow \mathbb{R}$  eine stetige Funktion auf einem nicht entarteten Intervall  $J \subseteq \mathbb{R}$ . Für alle  $t_0 \in J$  und  $y_0, v_0 \in \mathbb{R}$  gilt:*

- (a) *Es existiert eine auf ganz  $J$  definierte Lösung  $\phi: J \rightarrow \mathbb{R}$  des Anfangswertproblems*

$$\begin{cases} y''(t) + a_1 y'(t) + a_0 y(t) &= b(t) \\ y(t_0) &= y_0 \\ y'(t_0) &= v_0. \end{cases}$$

(b) Ist  $I$  ein nicht entartetes Intervall mit  $t_0 \in I \subseteq J$  und  $\psi: I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems aus (a), so ist  $\psi = \phi|_I$ .

Beweis. (a) Nach 18.17 (a) gibt es eine Lösung

$$(\phi, \eta): J \rightarrow \mathbb{R}^2$$

des linearen Differentialgleichungssystems

$$\begin{pmatrix} y'(t) \\ v'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix} \begin{pmatrix} y(t) \\ v(t) \end{pmatrix} + \begin{pmatrix} 0 \\ b(t) \end{pmatrix} \quad (97)$$

mit  $(\phi(t_0), \eta(t_0)) = (y_0, v_0)$ . Dann ist also  $\phi: J \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion mit

$$\phi'(t) = \eta(t)$$

und  $\eta: J \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion mit

$$\eta'(t) = -a_0 \phi(t) - a_1 \eta(t) + b(t).$$

Insbesondere ist  $\phi$  eine  $C^2$ -Funktion und Einsetzen von  $\eta = \phi'$ ,  $\eta' = \phi''$  in die vorige Gleichung liefert

$$\phi''(t) = -a_0 \phi(t) - a_1 \phi'(t) + b(t).$$

Da zudem  $\phi(t_0) = y_0$  und  $\phi'(t_0) = \eta(t_0) = v_0$ , ist  $\phi$  eine Lösung des Anfangswertproblems in (a).

(b) Ist auch  $\psi: I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems in (a), so ist  $(\psi, \psi')$  eine Lösung des Differentialgleichungssystems (97) mit  $(\psi(t_0), \psi'(t_0)) = (y_0, v_0)$ , also  $(\psi, \psi') = (\phi, \eta)|_I$  nach 18.17 (b) und somit insbesondere  $\psi = \phi|_I$ .

**19.5** (Lösungsmenge einer homogenen linearen DGL zweiter Ordnung). *Es seien  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $a_0, a_1 \in \mathbb{R}$ . Dann gilt:*

- (a) Die Menge  $L_h$  aller auf  $J$  definierten Lösungen  $\phi: J \rightarrow \mathbb{R}$  der DGL  $y'' + a_1 y' + a_0 y = 0$  ist ein 2-dimensionaler Untervektorraum des Vektorraums  $\mathbb{R}^J$  aller reellwertigen Funktionen auf  $J$ .
- (b) Es sei  $t_0 \in J$ . Zwei Lösungen  $\phi_1, \phi_2 \in L_h$  der DGL bilden genau dann eine Basis für  $L_h$ , wenn die Vektoren

$$\begin{pmatrix} \phi_1(t_0) \\ \phi_1'(t_0) \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \phi_2(t_0) \\ \phi_2'(t_0) \end{pmatrix}$$

in  $\mathbb{R}^2$  linear unabhängig (und somit eine Basis für  $\mathbb{R}^2$ ) sind.

**Bemerkung 19.6** Die in (b) angegebenen Vektoren bilden genau dann eine Basis von  $\mathbb{R}^2$ , wenn die Matrix

$$\begin{pmatrix} \phi_1(t_0) & \phi_2(t_0) \\ \phi_1'(t_0) & \phi_2'(t_0) \end{pmatrix}$$

invertierbar ist, also die Determinante

$$W(\phi_1, \phi_2)(t_0) := \det \begin{pmatrix} \phi_1(t_0) & \phi_2(t_0) \\ \phi_1'(t_0) & \phi_2'(t_0) \end{pmatrix} = \phi_1(t_0)\phi_2'(t_0) - \phi_1'(t_0)\phi_2(t_0)$$

von Null verschieden ist. Man nennt  $W(\phi_1, \phi_2)(t_0)$  die *Wronski-Determinante* von  $\phi_1$  und  $\phi_2$  zur Zeit  $t_0$ .

Nachweis für 19.5. (a) Sind  $\phi_1, \phi_2 \in L_h$ , so ist

$$(\phi_1 + \phi_2)'' + a_1(\phi_1 + \phi_2)' + a_0(\phi_1 + \phi_2) = \underbrace{\phi_1'' + a_1\phi_1' + a_0\phi_1}_{=0} + \underbrace{\phi_2'' + a_1\phi_2' + a_0\phi_2}_{=0} = 0,$$

also  $\phi_1 + \phi_2 \in L_h$ . Ähnlich sieht man, dass  $\lambda\phi \in L_h$  für alle  $\phi \in L_h$  und  $\lambda \in \mathbb{R}$ .

Für  $t_0 \in J$  ist die Abbildung

$$\alpha_{t_0}: L_h \rightarrow \mathbb{R}^2, \quad \phi \mapsto \begin{pmatrix} \phi(t_0) \\ \phi'(t_0) \end{pmatrix}$$

linear. Da nach 19.4 für alle

$$\begin{pmatrix} y_0 \\ v_0 \end{pmatrix} \in \mathbb{R}^2$$

das Anfangswertproblem

$$\begin{cases} y'' + a_1y' + a_0y = 0 \\ y(t_0) = y_0 \\ y'(t_0) = v_0 \end{cases}$$

eine Lösung  $\phi \in L_h$  besitzt, ist  $\alpha_{t_0}$  surjektiv. Da die Lösung eindeutig ist, ist  $\alpha_{t_0}$  auch injektiv und somit ein Isomorphismus von Vektorräumen. Insbesondere ist

$$\dim(L_h) = \dim \mathbb{R}^2 = 2.$$

(b) Da  $\alpha_{t_0}: L_h \rightarrow \mathbb{R}^2$  ein Isomorphismus von Vektorräumen ist, sind  $\phi_1, \phi_2 \in L_h$  genau dann linear unabhängig (und somit eine Basis für  $L_h$ ), wenn

$$\alpha_{t_0}(\phi_1) = \begin{pmatrix} \phi_1(t_0) \\ \phi_1'(t_0) \end{pmatrix} \quad \text{und} \quad \alpha_{t_0}(\phi_2) = \begin{pmatrix} \phi_2(t_0) \\ \phi_2'(t_0) \end{pmatrix}$$

in  $\mathbb{R}^2$  linear unabhängig (und somit eine Basis) sind.

**Definition 19.7** Bilden  $\phi_1, \phi_2 \in L_h$  eine Basis für  $L_h$ , so nennt man das Paar  $(\phi_1, \phi_2)$  auch ein *Lösungsfundamentalsystem* für die homogene lineare DGL  $y'' + a_1y' + a_0y = 0$ .

**Beispiel 19.8** (Harmonischer Oszillator). Für  $\omega_0 > 0$  betrachten wir die homogene lineare Differentialgleichung

$$y'' + \omega_0^2 y = 0$$

zweiter Ordnung, die sogenannte *Schwingungsgleichung*. Dann bilden die durch

$$\phi_1(t) := \sin(\omega_0 t) \quad \text{und} \quad \phi_2(t) := \cos(\omega_0 t)$$

gegebenen Funktionen  $\phi_1, \phi_2: \mathbb{R} \rightarrow \mathbb{R}$  ein Lösungsfundamentalsystem. Wegen

$$\phi_1'(t) = \omega_0 \cos(\omega_0 t), \quad \phi_1''(t) = -\omega_0^2 \sin(\omega_0 t) = -\omega_0^2 \phi_1(t)$$

ist  $\phi_1$  eine Lösung der Schwingungsgleichung. Da

$$\phi_2'(t) = -\omega_0 \sin(\omega_0 t) \quad \text{und} \quad \phi_2''(t) = -\omega_0^2 \cos(\omega_0 t) = -\omega_0^2 \phi_2(t),$$

ist auch  $\phi_2$  eine Lösung der Schwingungsgleichung. Weil die Vektoren

$$\begin{pmatrix} \phi_1(0) \\ \phi_1'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \omega_0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \phi_2(0) \\ \phi_2'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

linear unabhängig sind, ist  $(\phi_1, \phi_2)$  ein Lösungsfundamentalsystem.

Beliebige Linearkombinationen

$$\phi(t) = \alpha \sin(\omega_0 t) + \beta \cos(\omega_0 t)$$

mit  $\alpha, \beta \in \mathbb{R}$  sind dann ebenfalls Lösungen der Schwingungsgleichung. All diese sind übrigens harmonische Schwingungen der Form

$$\phi(t) = \sqrt{\alpha^2 + \beta^2} \sin(\omega_0 t + c)$$

mit einem  $c \in \mathbb{R}$ , als Konsequenz des (aus der Mathematik 1 bekannten) Additionstheorems

$$\cos(\sigma) \sin(\tau) + \sin(\sigma) \cos(\tau) = \sin(\sigma + \tau)$$

für  $\sigma, \tau \in \mathbb{R}$ .

**19.9** Für alle  $a, b \in \mathbb{R}$  und  $\omega_0 > 0$  gibt es ein  $c \in \mathbb{R}$  derart, dass für alle  $t \in \mathbb{R}$

$$a \sin(\omega_0 t) + b \cos(\omega_0 t) = \sqrt{a^2 + b^2} \sin(\omega_0 t + c).$$

Die Aussage ist trivial, wenn  $a = b = 0$ . Seien nun nicht  $a$  und  $b$  beide 0, also  $a^2 + b^2 > 0$ . Der Punkt

$$\left( \frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} \right)$$

liegt auf dem Einheitskreis, ist also von der Form  $(\cos(c), \sin(c))$  für ein  $c \in [0, 2\pi[$ . Folglich ist

$$\begin{aligned} a \sin(\omega_0 t) + b \cos(\omega_0 t) &= \sqrt{a^2 + b^2} \left( \frac{a}{\sqrt{a^2 + b^2}} \sin(\omega_0 t), \frac{b}{\sqrt{a^2 + b^2}} \cos(\omega_0 t) \right) \\ &= \sqrt{a^2 + b^2} (\cos(c) \sin(\omega_0 t) + \sin(c) \cos(\omega_0 t)) \\ &= \sqrt{a^2 + b^2} \sin(\omega_0 t + c). \end{aligned}$$

**19.10** Wir betrachten eine homogene lineare Differentialgleichung

$$y'' + a_1 y' + a_0 y = 0 \tag{98}$$

zweiter Ordnung mit konstanten Koeffizienten. Wir bilden das Polynom

$$p(z) = z^2 + a_1 z + a_0$$

mit  $z \in \mathbb{C}$ , das sogenannte charakteristische Polynom.

- (a) Hat  $p$  zwei verschiedene reelle Nullstellen  $\lambda_1, \lambda_2 \in \mathbb{R}$ , so bilden die durch

$$\phi_1(t) = e^{\lambda_1 t}, \quad \phi_2(t) = e^{\lambda_2 t}$$

gegebenen Funktion  $\phi_1, \phi_2: \mathbb{R} \rightarrow \mathbb{R}$  ein Lösungsfundamentalsystem für die homogene lineare DGL (98).

(b) Hat  $p$  eine doppelte reelle Nullstelle  $\lambda \in \mathbb{R}$ , so bilden die durch

$$\phi_1(t) = e^{\lambda t}, \quad \phi_2(t) = t e^{\lambda t}$$

gegebenen Funktion  $\phi_1, \phi_2: \mathbb{R} \rightarrow \mathbb{R}$  ein Lösungsfundamentalsystem für die homogene lineare DGL (98).

(c) Hat  $p$  ein Paar nicht reeller, komplex konjugierter komplexer Nullstellen  $\alpha \pm i\omega$  mit  $\alpha \in \mathbb{R}$  und  $0 \neq \omega \in \mathbb{R}$ , so bilden die durch

$$\phi_1(t) = e^{\alpha t} \sin(\omega t), \quad \phi_2(t) = e^{\alpha t} \cos(\omega t)$$

gegebenen Funktion  $\phi_1, \phi_2: \mathbb{R} \rightarrow \mathbb{R}$  ein Lösungsfundamentalsystem für die homogene lineare DGL (98).

Nachweis: (a) Für  $\phi_j(t) = e^{\lambda_j t}$  mit  $j \in \{1, 2\}$  und  $t \in \mathbb{R}$  ist

$$\phi_j'(t) = \lambda_j e^{\lambda_j t} = \lambda_j \phi_j(t) \quad \text{und} \quad \phi_j''(t) = \lambda_j^2 e^{\lambda_j t} = \lambda_j^2 \phi_j(t),$$

also

$$\phi_j''(t) + a_1 \phi_j'(t) + a_0 \phi_j(t) = (\lambda_j^2 + a_1 \lambda_j + a_0) \phi_j(t) = p(\lambda_j) \phi_j(t) = 0.$$

Somit sind  $\phi_1$  und  $\phi_2$  Lösungen der DGL  $y'' + a_1 y' + a_0 y = 0$ . Da die Vektoren

$$\begin{pmatrix} \phi_1(0) \\ \phi_1'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \phi_2(0) \\ \phi_2'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$$

linear unabhängig sind (mit Wronski-Determinante  $\lambda_2 - \lambda_1 \neq 0$ ), ist  $(\phi_1, \phi_2)$  ein Lösungsfundamentalsystem.

(b) Wie im Beweis von (a) sieht man, dass

$$\phi_1'' + a_1 \phi_1' + a_0 \phi_1 = p(\lambda) \phi_1 = 0.$$

Da  $p: \mathbb{R} \rightarrow \mathbb{R}$ ,  $t \mapsto (t - \lambda)^2$  die doppelte reelle Nullstelle  $\lambda$  hat, ist  $p'(\lambda) = 2(\lambda - \lambda) = 0$ . Andererseits ist  $p(t) = t^2 + a_1 t + a_0$ , also

$$p'(t) = 2t + a_1,$$

also

$$0 = p'(\lambda) = 2\lambda + a_1.$$

Nun ist

$$\phi_2'(t) = e^{\lambda t} + t\lambda e^{\lambda t}$$

und

$$\phi_2''(t) = \lambda e^{\lambda t} + \lambda e^{\lambda t} + t\lambda^2 e^{\lambda t} = 2\lambda e^{\lambda t} + t\lambda^2 e^{\lambda t},$$

also

$$\begin{aligned}\phi_2''(t) + a_1 \phi_2'(t) + a_0 \phi_2(t) &= (2\lambda + t\lambda^2 + a_1 + a_1 t\lambda + a_0) e^{\lambda t} \\ &= \underbrace{(2\lambda + a_1)}_{=p'(\lambda)=0} + \underbrace{(\lambda^2 + a_1\lambda + a_0)}_{=p(\lambda)=0} t e^{\lambda t} \\ &= 0.\end{aligned}$$

Folglich ist auch  $\phi_2$  eine Lösung der homogenen DGL. Man beachte, dass  $\phi_2(0) = 0$  und  $\phi_2'(0) = 1$ . Also ist

$$\begin{pmatrix} \phi_1(0) \\ \phi_1'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \phi_2(0) \\ \phi_2'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

mit Wronski-Determinante  $1 \neq 0$ , somit  $(\phi_1, \phi_2)$  ein Lösungsfundamentalsystem.

(c) Es ist

$$\begin{aligned}p(z) &= (z - \alpha - i\omega)(z - \alpha + i\omega) = z^2 + (-\alpha - i\omega - \alpha + i\omega)z + (\alpha + i\omega)(\alpha - i\omega) \\ &= z^2 - 2\alpha z + |\alpha + i\omega|^2 = z^2 - 2\alpha z + (\alpha^2 + \omega^2).\end{aligned}$$

Vergleich mit

$$p(z) = z^2 + a_1 z + a_0$$

zeigt, dass

$$-2\alpha = a_1 \quad \text{und} \quad \alpha^2 + \omega^2 = a_0.$$

Für  $\phi_1(t) = e^{\alpha t} \sin(\omega t)$  und  $\phi_2(t) = e^{\alpha t} \cos(\omega t)$  gilt

$$\begin{aligned}\phi_1'(t) &= \alpha e^{\alpha t} \sin(\omega t) + \omega e^{\alpha t} \cos(\omega t), \\ \phi_2'(t) &= \alpha e^{\alpha t} \cos(\omega t) - \omega e^{\alpha t} \sin(\omega t), \\ \phi_1''(t) &= \alpha^2 e^{\alpha t} \sin(\omega t) + \alpha\omega e^{\alpha t} \cos(\omega t) + \alpha\omega e^{\alpha t} \cos(\omega t) - \omega^2 e^{\alpha t} \sin(\omega t), \\ \phi_2''(t) &= \alpha^2 e^{\alpha t} \cos(\omega t) - \alpha\omega e^{\alpha t} \sin(\omega t) - \alpha\omega e^{\alpha t} \sin(\omega t) - \omega^2 e^{\alpha t} \cos(\omega t).\end{aligned}$$

Also ist  $\phi_1''(t) + a_1 \phi_1'(t) + a_0 \phi_1(t)$  gleich

$$\alpha^2 e^{\alpha t} \sin(\omega t) + \alpha \omega e^{\alpha t} \cos(\omega t) + \alpha \omega e^{\alpha t} \cos(\omega t) - \omega^2 e^{\alpha t} \sin(\omega t) \\ + \underbrace{a_1}_{=-2\alpha} (\alpha e^{\alpha t} \sin(\omega t) + \omega e^{\alpha t} \cos(\omega t)) + \underbrace{a_0}_{=\alpha^2 + \omega^2} e^{\alpha t} \sin(t) = 0,$$

somit  $\phi_1$  eine Lösung der DGL. Analog sieht man, dass auch  $\phi_2$  die DGL löst. Nun sind

$$\begin{pmatrix} \phi_1(0) \\ \phi_1'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \omega \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \phi_2(0) \\ \phi_2'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}$$

linear unabhängig, mit Wronski-Determinante  $W(\phi_1, \phi_2)(0) = -\omega \neq 0$ . Also ist  $(\phi_1, \phi_2)$  ein Lösungsfundamentalsystem.

**Beispiel 19.11** Seien  $\nu, \omega_0 > 0$ . In der Physik beschreibt die homogene lineare Differentialgleichung

$$y'' + \nu y' + \omega_0^2 y = 0 \tag{99}$$

einen gedämpften Oszillator.<sup>21</sup> Das charakteristische Polynom lautet

$$p(z) = z^2 + \nu z + \omega_0^2$$

und wir haben

$$0 = p(z) = z^2 + \nu z + \omega_0^2 = (z + \nu/2)^2 + \omega_0^2 - \nu^2/4$$

genau dann, wenn

$$(z + \nu/2)^2 = \nu^2/4 - \omega_0^2. \tag{100}$$

1. Fall (Schwingfall): Ist  $\nu^2/4 - \omega_0^2 < 0$ , so können wir (100) schreiben als

$$(z + \nu/2)^2 = -(\omega_0^2 - \nu^2/4)$$

mit  $\omega_0^2 + \nu^2/4 > 0$ . Wir haben also ein Paar komplex konjugierter, nicht reeller Nullstellen,

$$z_1 = -\nu/2 + i\sqrt{\omega_0^2 - \nu^2/4}, \quad z_2 = -\nu/2 - i\sqrt{\omega_0^2 - \nu^2/4}$$

---

<sup>21</sup>Es handelt sich um sogenannte "lineare" Dämpfung, die proportional zu  $y'(t)$  ist. Bei Gleitreibung oder Haftreibung eines Körpers auf einer Oberfläche ist dies *nicht* (!) der Fall.

der Vielfachheit 1. Nach 19.10 (c) bilden also die Funktionen

$$t \mapsto e^{-\frac{\nu}{2}t} \sin(\omega t) \quad \text{und} \quad t \mapsto e^{-\frac{\nu}{2}t} \cos(\omega t)$$

mit  $\omega := \sqrt{\omega_0^2 - \nu^2/4}$  ein Lösungsfundamentalsystem.

2. Fall (Kriechfall). Ist  $\nu^2/4 - \omega_0^2 > 0$ , so hat  $p$  die zwei reellen Nullstellen

$$\lambda_1 = -\nu/2 - \sqrt{\nu^2/4 - \omega_0^2} \quad \text{und} \quad \lambda_2 = -\nu/2 + \sqrt{\nu^2/4 - \omega_0^2}$$

der Vielfachheit 1. Offenbar ist  $\lambda_1 < 0$  und es ist auch  $\lambda_2 < 0$ , weil  $\sqrt{\nu^2/4 - \omega_0^2} < \sqrt{\nu^2/4} = \nu/2$ . Nach 19.10 (a) bilden die Funktionen

$$t \mapsto e^{t\lambda_1} \quad \text{und} \quad t \mapsto e^{t\lambda_2}$$

ein Lösungsfundamentalsystem.

3. Fall (aperiodischer Grenzfall). Ist  $\nu^2/4 - \omega_0^2 = 0$ , so hat  $p$  die doppelte Nullstelle

$$\lambda = -\nu/2.$$

Nach 19.10 (b) bilden also die Funktionen

$$t \mapsto e^{-\frac{\nu}{2}t} \quad \text{und} \quad t \mapsto t e^{-\frac{\nu}{2}t}$$

ein Lösungsfundamentalsystem.

**Bemerkung 19.12** Beachten Sie, dass  $\phi_1(t) \rightarrow 0$  und  $\phi_2(t) \rightarrow 0$  für  $t \rightarrow \infty$ . Da jede Lösung  $\phi$  der DGL eine Linearkombination  $\phi = \alpha\phi_1 + \beta\phi_2$  von  $\phi_1$  und  $\phi_2$  ist, folgt  $\phi(t) \rightarrow \alpha \cdot 0 + \beta \cdot 0 = 0$  für  $t \rightarrow \infty$ .

**Bemerkung 19.13** Betrachten wir im Kriechfall oder aperiodischen Grenzfall eine Lösung  $\phi$  der DGL mit  $\phi(t_0) > 0$ . Drei Fälle können auftreten (Skizzen siehe Vorlesung!):

- (a) Es ist  $\phi'(t_0) > 0$ . Dann ist  $\phi$  auf einem Intervall  $[t_0, T]$  monoton wachsend, ab  $T$  dann monoton fallend mit  $\phi(t) \rightarrow 0$  für  $t \rightarrow \infty$ .
- (b) Es ist  $\phi'(t_0) \leq 0$  und  $\phi$  ist auf  $[t_0, \infty[$  monoton fallend mit  $\phi(t) \rightarrow 0$  für  $t \rightarrow \infty$ .
- (c) Es ist  $\phi'(t_0) \leq 0$  und  $\phi$  ist auf einem Intervall  $[t_0, T]$  monoton fallend mit  $\phi(T) < 0$ , anschließend monoton wachsend mit  $\phi(t) \rightarrow 0$  für  $t \rightarrow \infty$ .

Eine Linearkombination  $\phi = \alpha\phi_1 + \beta\phi_2 \neq 0$  hat nämlich höchstens eine lokale Extremalstelle. In der Tat ist im Kriechfall

$$\phi(t) = \alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t}$$

mit

$$\phi'(t) = \alpha\lambda_1 e^{\lambda_1 t} + \beta\lambda_2 e^{\lambda_2 t} = e^{\lambda_1 t}(\alpha\lambda_1 + \beta\lambda_2 e^{(\lambda_2 - \lambda_1)t}),$$

also  $\phi'(t) = 0$  genau dann, wenn

$$\beta\lambda_2 e^{(\lambda_2 - \lambda_1)t} = -\alpha\lambda_1,$$

was für höchstens ein  $t \geq t_0$  erfüllt ist. Im aperiodischen Grenzfall ist

$$\phi(t) = \alpha e^{-\frac{\nu}{2}t} + \beta t e^{-\frac{\nu}{2}t} = (\alpha + \beta t) e^{-\frac{\nu}{2}t},$$

also

$$\phi'(t) = \beta e^{-\frac{\nu}{2}t} - (\alpha + \beta t) \frac{\nu}{2} e^{-\frac{\nu}{2}t} = \left( \beta - \frac{\nu\alpha}{2} - \frac{\nu\beta}{2}t \right) e^{-\frac{\nu}{2}t}.$$

Folglich ist  $\phi'(t) = 0$  genau dann, wenn

$$\frac{\nu\beta}{2}t = \beta - \frac{\nu\alpha}{2},$$

was für höchstens ein  $t \in [t_0, \infty[$  der Fall ist.

## Inhomogene lineare DGLn zweiter Ordnung

**19.14** Es seien  $a_0, a_1 \in \mathbb{R}$ ,  $J \subseteq \mathbb{R}$  ein nicht entartetes Intervall und  $b: J \rightarrow \mathbb{R}$  eine stetige Funktion. Es sei  $L_h$  die Menge aller Lösungen  $\phi: J \rightarrow \mathbb{R}$  der homogenen linearen DGL

$$y'' + a_1 y' + a_0 y = 0$$

zweiter Ordnung mit konstanten Koeffizienten und  $L_i$  die Menge aller Lösungen der inhomogenen linearen DGL

$$y''(t) + a_1 y'(t) + a_0 y(t) = b(t)$$

zweiter Ordnung mit konstanten Koeffizienten. Ist  $\phi_p \in L_i$  eine beliebige Lösung der inhomogenen DGL, so ist

$$L_i = \phi_p + L_h = \{\phi_p + \phi: \phi \in L_h\}.$$

**Bemerkung 19.15** Traditionell nennt man  $\phi_p$  eine *partikuläre Lösung* der inhomogenen linearen DGL und kann 19.14 dann wie folgt zusammenfassen:

“Die allgemeine Lösung einer inhomogenen linearen DGL erhält man, indem man zu einer partikulären Lösung der inhomogenen DGL die allgemeine Lösung der zugehörigen homogenen linearen DGL addiert.”

Begründung für 19.14: Für jedes  $\phi \in L_h$  rechnet man sofort nach, dass  $\phi_p + \phi$  die inhomogene lineare DGL erfüllt. Für jedes  $\phi \in L_i$  rechnet man sofort nach, dass  $\phi - \phi_p$  die zugehörige homogene DGL erfüllt, also  $\phi - \phi_p \in L_h$  ist und somit

$$\phi = \phi_p + (\phi - \phi_p) \in \phi_p + L_h.$$

**Beispiel 19.16** (Getriebener Oszillator). Lässt man auf einen harmonischen Oszillator zusätzlich eine äußere Kraft wirken, die eine harmonische Schwingung mit Winkelfrequenz  $\Omega$  ist, erhält man eine inhomogene lineare DGL zweiter Ordnung mit konstanten Koeffizienten von der Form

$$y''(t) + \omega_0^2 y(t) = a \sin(\Omega t + b). \quad (101)$$

(a) Ist  $\Omega \neq \omega_0$ , so ist

$$\phi_p(t) = \frac{a}{\omega_0^2 - \Omega^2} \sin(\Omega t + b)$$

eine partikuläre Lösung, wie man durch Einsetzen von  $\phi_p(t)$  und

$$\phi_p''(t) = -\frac{a \Omega^2}{\omega_0^2 - \Omega^2} \sin(\Omega t + b)$$

in die inhomogene lineare DGL (101) sofort verifiziert.

(b) (Resonanzkatastrophe). Ist  $\Omega = \omega_0$ , wird der Oszillator also mit seiner Resonanzfrequenz angeregt, so ist

$$\phi_p(t) = -\frac{a}{2\omega_0} t \cos(\omega_0 t + b)$$

eine partikuläre Lösung der inhomogenen linearen DGL (101). Es ist nämlich

$$\phi_p'(t) = -\frac{a}{2\omega_0} \cos(\omega_0 t + b) + \frac{a}{2} t \sin(\omega_0 t + b).$$

Durch Einsetzen von  $\phi_p(t)$  und

$$\begin{aligned}\phi_p''(t) &= \frac{a}{2} \sin(\omega_0 t + b) + \frac{a}{2} \sin(\omega_0 t + b) + \frac{a\omega_0}{2} t \cos(\omega_0 t + b) \\ &= a \sin(\omega_0 t + b) + \frac{a\omega_0}{2} t \cos(\omega_0 t + b)\end{aligned}$$

in (101) verifiziert man sofort, dass  $\phi_p$  die inhomogene lineare DGL löst, also eine partikuläre Lösung ist.

Beachten Sie, dass für  $t = \frac{2\pi n - b}{\omega_0}$  mit  $n \in \mathbb{N}$

$$\phi_p(t) = -\frac{a}{2\omega_0^2} (2\pi n - b)$$

ist, was für  $n \rightarrow \infty$  bestimmt gegen  $-\infty$  divergiert. Es gibt also beliebig große Auslenkungen und gleiches gilt für jede Lösung  $\phi + \phi_p$  der inhomogenen DGL mit  $\phi \in L_h$ , da  $\phi$  eine Linearkombination von  $\sin(\omega_0 t)$  und  $\cos(\omega_0 t)$  und somit eine beschränkte Funktion ist. Zwar wird die Modellierung bei beliebig großen Auslenkungen irgendwann unpräzise; dennoch muss man befürchten, dass ein schwingendes Werkstück dann zerstört werden könnte.

**Beispiel 19.17** Wir betrachten einen getriebenen gedämpften Oszillator, der durch die inhomogene lineare Differentialgleichung

$$y''(t) + \nu y'(t) + \omega_0^2 y(t) = a \sin(\Omega t + b) \quad (102)$$

zweiter Ordnung beschrieben wird mit  $\nu > 0$ ,  $\omega_0 \neq 0$ ,  $a \neq 0$ ,  $\Omega \neq 0$  und  $b \in \mathbb{R}$ . Wir prüfen nach, dass eine partikuläre Lösung gegeben ist durch

$$\phi_p(t) = -\frac{a\nu\Omega}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b) - \frac{a(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b).$$

Wir berechnen

$$\phi_p'(t) = \frac{a\nu\Omega^2}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b) - \frac{a\Omega(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b)$$

und

$$\phi_p''(t) = \frac{a\nu\Omega^3}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b) + \frac{a\Omega^2(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b).$$

Einsetzen in die linke Seite von (102) liefert

$$\begin{aligned} & \frac{a\nu\Omega^3}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b) + \frac{a\Omega^2(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b) \\ & + \frac{a\nu^2\Omega^2}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b) - \frac{a\nu\Omega(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b) \\ & - \frac{a\nu\omega_0^2\Omega}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \cos(\Omega t + b) - \frac{a\omega_0^2(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} \sin(\Omega t + b). \end{aligned}$$

Summieren der Koeffizienten vor  $\cos(\Omega t + b)$  liefert

$$\frac{a\nu\Omega^3 - a\nu\Omega(\Omega^2 - \omega_0^2) - a\nu\omega_0^2\Omega}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} = 0$$

Summieren der Koeffizienten vor  $\sin(\Omega t + b)$  liefert

$$a \frac{\Omega^2(\Omega^2 - \omega_0^2) + \nu^2\Omega^2 - \omega_0^2(\Omega^2 - \omega_0^2)}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} = a.$$

Also erfüllt  $\phi_p$  die inhomogene lineare DGL (102).

Mit 19.9 schließen wir, dass

$$\phi_p(t) = A \sin(\Omega t + c)$$

mit einem  $c \in \mathbb{R}$  und

$$A = \frac{\sqrt{a^2\nu^2\Omega^2 + a^2(\Omega^2 - \omega_0^2)^2}}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} = |a| \frac{\sqrt{\nu^2\Omega^2 + (\Omega^2 - \omega_0^2)^2}}{(\Omega^2 - \omega_0^2)^2 + \nu^2\Omega^2} = \frac{|a|}{\sqrt{\nu^2\Omega^2 + (\Omega^2 - \omega_0^2)^2}}.$$

**Bemerkung 19.18** Die allgemeine Lösung der DGL (102) ist von der Form

$$\psi = \phi + \phi_p$$

mit  $\phi_p$  wie zuvor und einer Lösung  $\phi$  der homogenen linearen DGL (99). Da  $\phi(t) \rightarrow 0$  für  $t \rightarrow \infty$  (siehe Bemerkung 19.12), geht die Bewegung des gedämpften getriebenen Oszillators mehr und mehr in die harmonische Schwingung  $\phi_p$  mit der antreibenden Kreisfrequenz  $\Omega$  über.

**Bemerkung 19.19** Wie groß kann  $A$  werden als Funktion von  $\Omega$ ?

Wir beantworten diese Frage für einen nur leicht gedämpften Oszillator, mit

$$\omega_0^2 - \frac{\nu^2}{2} > 0.$$

Dann gilt insbesondere

$$\omega_0^2 - \frac{\nu^2}{4} > 0,$$

es liegt also der Schwingfall vor. Für  $\Omega \rightarrow 0$  geht  $A \rightarrow \frac{|a|}{\omega_0^2}$ . Für  $\Omega \rightarrow \infty$  geht  $A \rightarrow 0$ . Es wird  $A$  maximal, wenn der Nenner minimal wird, also sein Quadrat

$$\nu^2 \Omega^2 + (\Omega^2 - \omega_0^2)^2 = \nu^2 \Omega^2 + \Omega^4 - 2\Omega^2 \omega_0^2 + \omega_0^4 = \left( \Omega^2 + \frac{\nu^2}{2} - \omega_0^2 \right)^2 + \omega_0^4 - \left( \frac{\nu^2}{2} - \omega_0^2 \right)^2$$

minimal wird. Dies ist der Fall, wenn

$$\Omega^2 + \frac{\nu^2}{2} - \omega_0^2 = 0,$$

also

$$\Omega^2 = \omega_0^2 - \frac{\nu^2}{2}.$$

Die maximale Amplitude ist dann

$$\begin{aligned} \frac{|a|}{\sqrt{\omega_0^4 - \left(\frac{\nu^2}{2} - \omega_0^2\right)^2}} &= \frac{|a|}{\sqrt{(\omega_0^2 - (\frac{\nu^2}{2} - \omega_0^2))(\omega_0^2 + (\frac{\nu^2}{2} - \omega_0^2))}} \\ &= \frac{|a|}{\sqrt{(2\omega_0^2 - \frac{\nu^2}{2})\frac{\nu^2}{2}}} = \frac{|a|}{\nu\sqrt{\omega_0^2 - \frac{\nu^2}{4}}}. \end{aligned}$$