

# Mathematik II für Bauwesen

geschrieben von Michael Dreher nach einer Vorlesung von Robert Haller–Dintelmann  
Fachbereich für Mathematik  
Technische Universität Darmstadt

Sommersemester 2013  
— wird kontinuierlich ergänzt —

Das Skript lebt auch durch die Mitarbeit der Leserinnen und Leser ! Hinweise auf Fehler und Verbesserungsmöglichkeiten sind unter der Adresse `dreher@mathematik.tu-darmstadt.de` immer willkommen; ein erster Dank geht an Benjamin Kuhn, Steffen Petker, Dirk Schröder und Jens Schmitt.

**Etwas Juristisches:**

Dieses Werk ist unter einem *Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported Lizenzvertrag* lizenziert. Um die Lizenz anzusehen, gehen Sie bitte zu <http://creativecommons.org/licenses/by-nc-nd/3.0/de/> oder schicken Sie einen Brief an Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

# Inhaltsverzeichnis

<b>1</b>	<b>Lineare Algebra</b>	<b>5</b>
1	Vektorrechnung . . . . .	5
2	Lineare Gleichungssysteme . . . . .	12
3	Matrizenrechnung . . . . .	21
4	Lineare Abbildungen . . . . .	28
5	Determinanten . . . . .	39
6	Eigenwerte und Eigenvektoren . . . . .	44
7	Quadratische Formen und Kegelschnitte . . . . .	49
<b>2</b>	<b>Differentialrechnung in mehreren Variablen</b>	<b>55</b>
1	Funktionen mehrerer Veränderlicher . . . . .	55
2	Differentialrechnung in mehreren Variablen . . . . .	63
3	Anwendungen . . . . .	73
<b>3</b>	<b>Integration in mehreren Variablen</b>	<b>87</b>
1	Parameterintegrale . . . . .	87
2	Kurvenintegrale . . . . .	90
3	Integrale im $\mathbb{R}^n$ . . . . .	100
4	Vektoranalysis . . . . .	108
	<b>Index</b>	<b>116</b>



# Kapitel 1

## Lineare Algebra

### 1 Vektorrechnung

Wir **erinnern** an die Vorlesung Mathematik I: der  $\mathbb{R}^n$  ist die Menge aller Punkte

$$P = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \dots, x_n]^T.$$

Die Einträge  $x_1, x_2, \dots, x_n$  sind reelle Zahlen und heißen *Koordinaten* von  $P$ . Für  $n = 1$  erhalten wir den Zahlenstrahl, für  $n = 2$  eine Ebene, und für  $n = 3$  den Raum.

Wenn wir zwei Punkte  $P$  und  $Q$  aus dem  $\mathbb{R}^n$  haben, dann können wir einen Vektor  $\vec{x}$  definieren als

$$\vec{x} = P - Q,$$

wobei wir  $P$  als Endpunkt (markiert durch die „Spitze des Pfeils“) und  $Q$  als Anfangspunkt haben. Ein solcher Vektor  $\vec{x}$  ist eindeutig bestimmt durch seine Richtung und seine Länge, die auch „Norm“ genannt wird. Verschiedene Anfangspunkte können denselben Vektor ergeben, siehe Abbildung 1.1.

Wenn wir  $Q = 0$  wählen, dann entsteht ein Pfeil, den wir *Ortsvektor* zum Punkt  $R$  nennen (siehe Abbildung 1.1). Auf diese Weise entsteht eine eindeutige Identifizierung von Punkten mit Vektoren.

Wir erinnern auch an zwei typische Rechenoperationen mit Vektoren: die Summe zweier Vektoren ergibt wieder einen Vektor, und die Multiplikation einer Zahl mit einem Vektor ergibt ebenfalls einen Vektor. Diese beiden Rechenoperationen haben folgende Eigenschaften:

#### Rechenregeln 1.1:

Seien  $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$  und  $\lambda, \mu \in \mathbb{R}$ . Dann gilt:

$$\vec{u} + \vec{v} = \vec{v} + \vec{u}, \quad (,,+“ \text{ ist kommutativ}) \quad (\text{V1})$$

$$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}), \quad (,,+“ \text{ ist assoziativ}) \quad (\text{V2})$$

$$\text{Es gibt ein } \vec{0} \in \mathbb{R}^n \text{ mit } \vec{u} + \vec{0} = \vec{u} \text{ für alle } \vec{u} \in \mathbb{R}^n, \quad (\text{Nullvektor}) \quad (\text{V3})$$

$$\text{nämlich } \vec{0} = [0, 0, \dots, 0]^T.$$

$$\text{Zu jedem } \vec{u} \in \mathbb{R}^n \text{ existiert ein } \vec{-u} \in \mathbb{R}^n \text{ mit } \vec{u} + (\vec{-u}) = \vec{0}. \quad (\text{inverses Element}) \quad (\text{V4})$$

$$1 \cdot \vec{u} = \vec{u}, \quad (\text{V5})$$

$$\lambda \cdot (\mu \cdot \vec{v}) = (\lambda \cdot \mu) \cdot \vec{v}, \quad (\text{V6})$$

$$\lambda \cdot (\vec{u} + \vec{v}) = \lambda \cdot \vec{u} + \lambda \cdot \vec{v}, \quad (\text{V7})$$

$$(\lambda + \mu) \cdot \vec{u} = \lambda \cdot \vec{u} + \mu \cdot \vec{u}. \quad (\text{V8})$$

Diese Rechenregeln bedeuten insbesondere, dass wir rechnen können wie gewohnt.

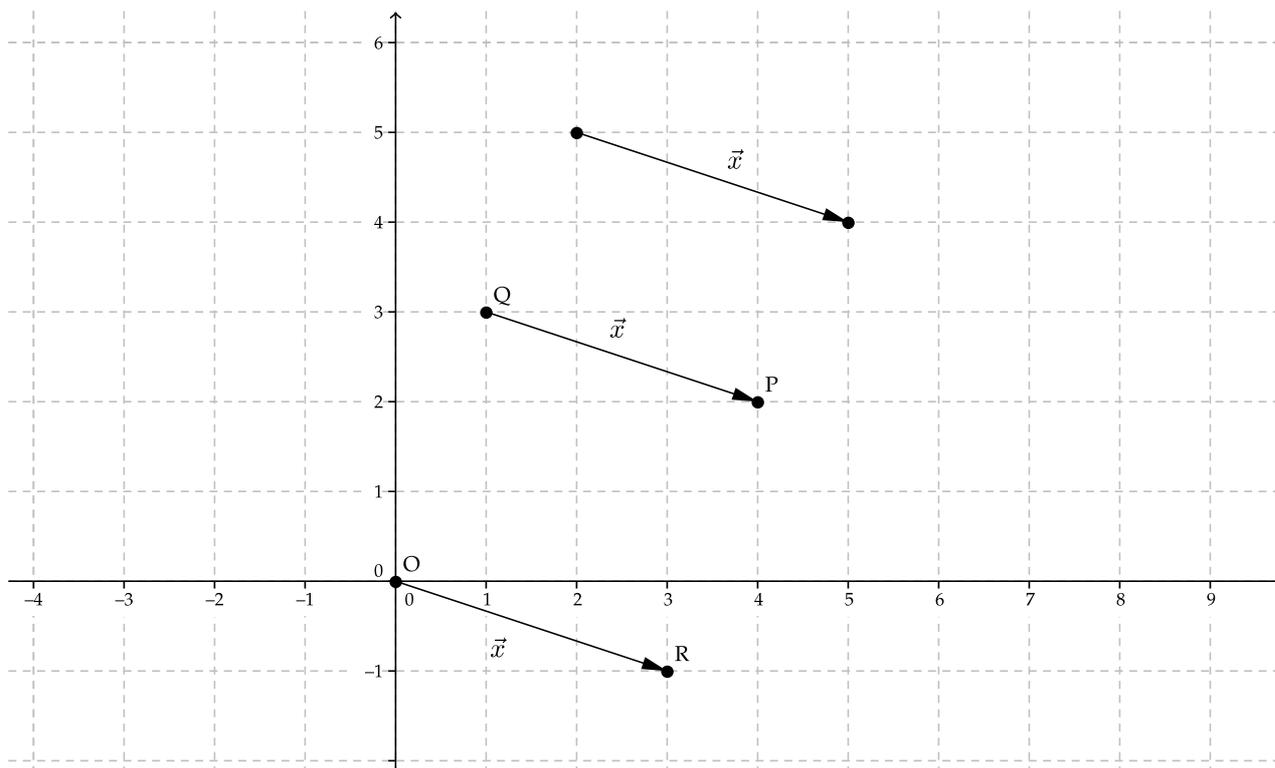


Abbildung 1.1: Jeder der drei Pfeile ist derselbe Vektor  $\vec{x}$ .

**Definition 1.2 (Vektorraum):**

Eine nichtleere Menge  $V$ , in der man Summe und zu jedem  $\lambda \in \mathbb{R}$  das  $\lambda$ -fache bilden kann, und die (V1)–(V8) erfüllt, heißt  $\mathbb{R}$ -Vektorraum.

**Beispiele 1.3:** (a)  $\mathbb{R}^n$  wie oben

(b)  $\mathbb{P}_n = \{a_0 + a_1x + \dots + a_nx^n : a_0, a_1, \dots, a_n \in \mathbb{R}\}$  (die Menge aller Polynome vom Grad  $\leq n$ ) ist ein  $\mathbb{R}$ -Vektorraum.

(c) Die Menge der stetigen reellwertigen Funktionen auf einem Intervall ist ein  $\mathbb{R}$ -Vektorraum.

**Bemerkung 1.4:** (a) Mit  $\mathbb{C}$  anstatt  $\mathbb{R}$  erhält man komplexe Vektorräume, auch  $\mathbb{C}$ -Vektorräume genannt. Die Rechenregeln sind die gleichen.

(b) Vektorräume, insbesondere  $\mathbb{R}^n$  und  $\mathbb{C}^n$ , sind grundlegend

- für geometrische Berechnungen (im Raum, in der Ebene, ...),
- als Definitionsbereich für Funktionen in mehreren Variablen,
- für lineare Gleichungssysteme,
- für lineare Differentialgleichungen (siehe Mathe III),
- ....

**Definition 1.5 (Untervektorraum):**

Sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Eine Menge  $U \subseteq V$  heißt *Untervektorraum* bzw. *Unterraum* bzw. *linearer Teilraum von  $V$* , falls  $U$  selbst ein  $\mathbb{R}$ -Vektorraum ist.

**Bemerkung 1.6:**

Es gilt wegen (V3) immer  $\vec{0} \in U$ , insbesondere ist jeder Untervektorraum nichtleer.

**Satz 1.7 (Kriterien für Untervektorräume):**

Eine nichtleere Teilmenge  $U \subseteq V$  eines  $\mathbb{R}$ -Vektorraums  $V$  ist Untervektorraum von  $V$  genau dann, wenn

(a) für alle  $\vec{u}, \vec{v} \in U$  und alle  $\lambda \in \mathbb{R}$  gilt, dass  $\vec{u} + \vec{v} \in U$  und  $\lambda\vec{u} \in U$ ,

und dies gilt genau dann, wenn

(b) für alle  $\vec{u}, \vec{v} \in U$  und alle  $\lambda, \mu \in \mathbb{R}$  gilt, dass  $\lambda\vec{u} + \mu\vec{v} \in U$ .

Das heißt, dass die beiden typischen Operationen eines Vektorraums (Vektor plus Vektor sowie Zahl mal Vektor) aus der Menge  $U$  niemals herausführen können.

**Beispiele 1.8:** (a) Offenkundig sind  $\{\vec{0}\}$  und  $V$  Untervektorräume von  $V$ . Diese heißen *triviale Untervektorräume*.

(b) Siehe Abbildungen 1.2, 1.3, 1.4.

(c) Alle Untervektorräume des  $\mathbb{R}^3$  sind:

- $\{\vec{0}\}$  und  $\mathbb{R}^3$ ,
- alle Geraden durch den Ursprung,
- alle Ebenen durch den Ursprung.

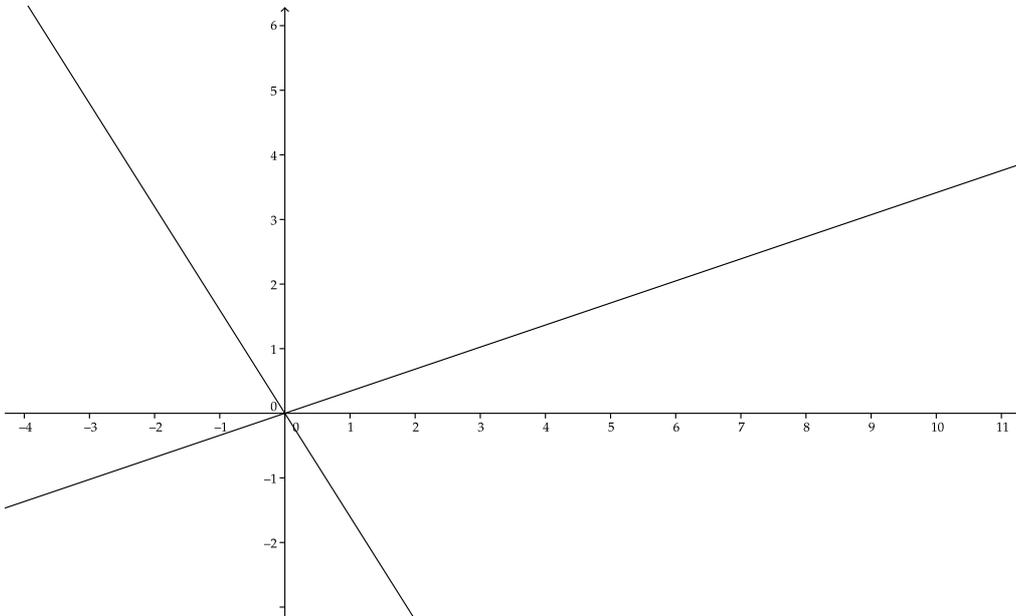


Abbildung 1.2: Die eine „schrägliegende Gerade“ ist ein Untervektorraum von  $V = \mathbb{R}^2$ , die andere ist ein weiterer UVR.

**Definition 1.9 (Linearkombination und lineare Hülle):**

Sei  $V$  ein  $\mathbb{R}$ -Vektorraum.

(a) Seien  $\vec{v}_1, \dots, \vec{v}_k \in V$ . Dann heißt  $\sum_{j=1}^k \alpha_j \vec{v}_j$  mit  $\alpha_1, \dots, \alpha_k \in \mathbb{R}$  *Linearkombination* von  $\vec{v}_1, \dots, \vec{v}_k$ .

(b)

$$\text{Lin}(\vec{v}_1, \dots, \vec{v}_k) := \left\{ \sum_{j=1}^k \alpha_j \vec{v}_j : \alpha_1, \dots, \alpha_k \in \mathbb{R} \right\},$$

also die Menge aller möglichen Linearkombinationen der Vektoren  $\vec{v}_1, \dots, \vec{v}_k$ , heißt *lineare Hülle* von  $\vec{v}_1, \dots, \vec{v}_k$ .

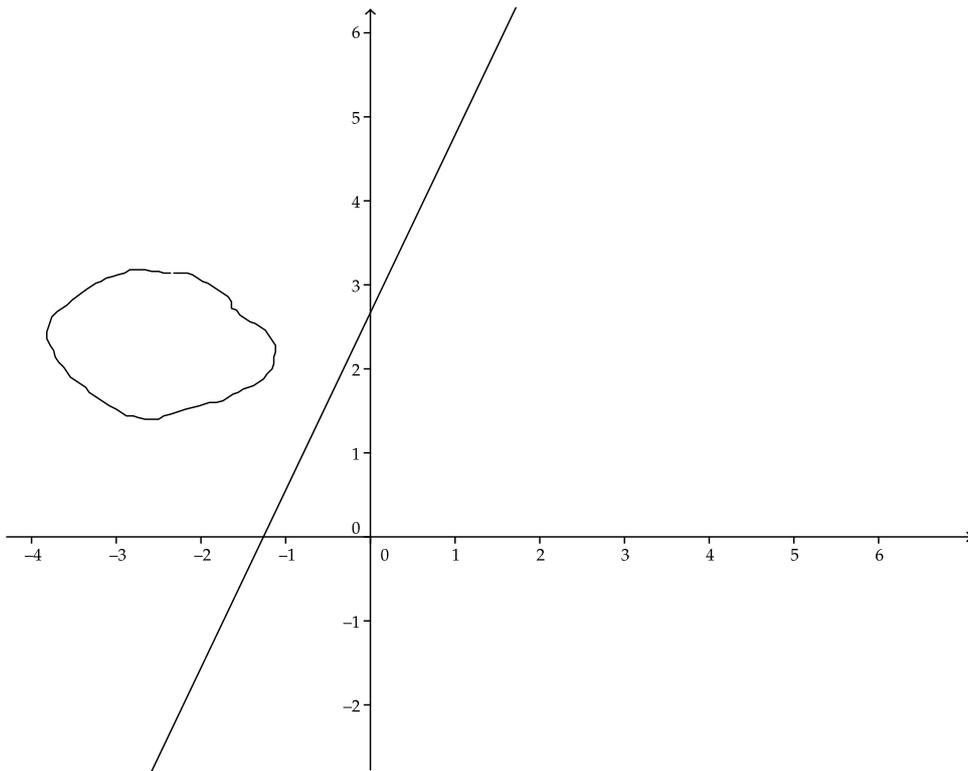


Abbildung 1.3: Dies sind keine Untervektorräume des  $\mathbb{R}^2$ .

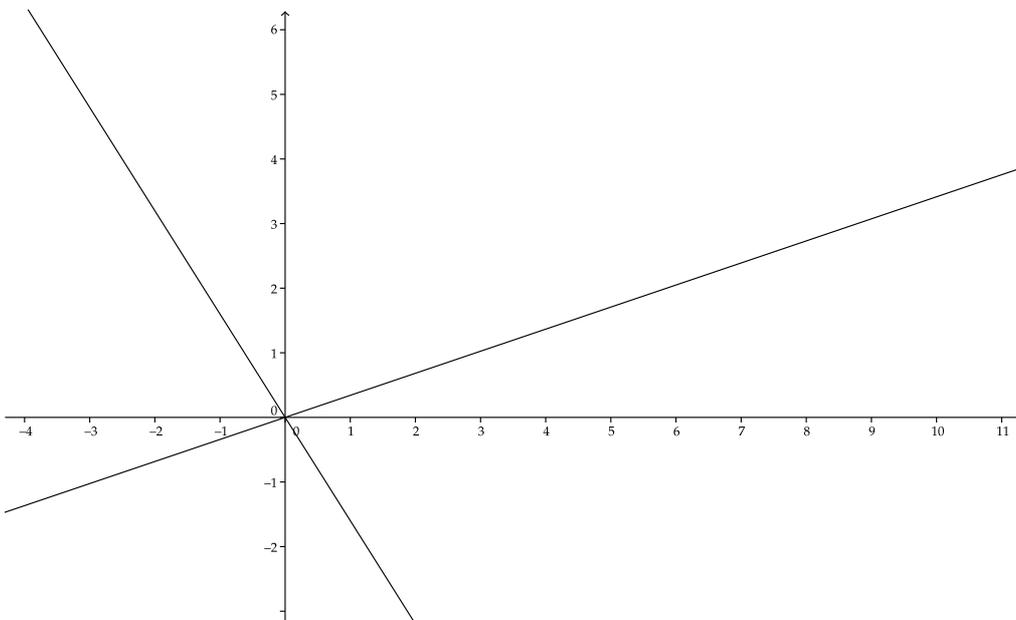


Abbildung 1.4: Beide „schrägliegenden Geraden“ zusammen ergeben keinen Untervektorraum von  $V = \mathbb{R}^2$ .

**Satz 1.10:**

$\text{Lin}(\vec{v}_1, \dots, \vec{v}_k)$  ist ein Untervektorraum von  $V$ .

*Beweis.* Wir benutzen Teil (b) von Satz 1.7.

**Schritt 1:** Es ist  $\text{Lin}(\vec{v}_1, \dots, \vec{v}_k) \neq \emptyset$ , denn wenigstens der Nullvektor  $\vec{0}$  ist in  $\text{Lin}(\vec{v}_1, \dots, \vec{v}_k)$  enthalten wegen  $\vec{0} = \sum_{j=1}^k 0 \cdot \vec{v}_j$ . Also dürfen wir im nächsten Schritt Teil (b) von Satz 1.7 benutzen.

**Schritt 2:** Seien  $\vec{u}, \vec{v} \in \text{Lin}(\vec{v}_1, \dots, \vec{v}_k)$  und  $\lambda, \mu \in \mathbb{R}$ . Das heißt, dass es Zahlen  $\alpha_1, \dots, \alpha_k$  und  $\beta_1, \dots, \beta_k \in \mathbb{R}$  gibt mit

$$\vec{u} = \sum_{j=1}^k \alpha_j \vec{v}_j \quad \text{und} \quad \vec{v} = \sum_{j=1}^k \beta_j \vec{v}_j,$$

und deshalb haben wir

$$(\lambda \vec{u} + \mu \vec{v}) = \lambda \sum_{j=1}^k \alpha_j \vec{v}_j + \mu \sum_{j=1}^k \beta_j \vec{v}_j = \sum_{j=1}^k \underbrace{(\lambda \alpha_j + \mu \beta_j)}_{=: \gamma_j} \vec{v}_j = \sum_{j=1}^k \gamma_j \vec{v}_j,$$

also ist tatsächlich  $\lambda \vec{u} + \mu \vec{v} \in \text{Lin}(\vec{v}_1, \dots, \vec{v}_k)$ .

Dann garantiert uns Teil (b) von Satz 1.7, dass  $\text{Lin}(\vec{v}_1, \dots, \vec{v}_k)$  ein Untervektorraum von  $V$  ist.  $\square$

**Beispiele 1.11:** (a) Wir haben

$$\text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) = \left\{ \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} : \alpha \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} \alpha \\ 0 \\ \alpha \end{bmatrix} : \alpha \in \mathbb{R} \right\}.$$

(b) Weiterhin ist

$$\text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = \mathbb{R}^3,$$

denn für jeden Vektor  $\vec{x} = [x_1, x_2, x_3]^\top \in \mathbb{R}^3$  gilt

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(c) Analog haben wir

$$\text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right) = \mathbb{R}^n.$$

Das letzte Beispiel hat die anschauliche Bedeutung, dass man den ganzen Raum  $\mathbb{R}^n$  aus  $n$  Vektoren zusammenbasteln kann (es sei denn, man wählt diese  $n$  Vektoren zu ungeschickt, nämlich linear abhängig).

**Definition 1.12 (Lineare (Un)abhängigkeit):**

Sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Die Vektoren  $\vec{v}_1, \dots, \vec{v}_k \in V$  heißen

**linear abhängig**, falls Zahlen  $\alpha_1, \dots, \alpha_k \in \mathbb{R}$  existieren, die nicht alle gleich Null sind und  $\sum_{j=1}^k \alpha_j \vec{v}_j = \vec{0}$  erfüllen.

**linear unabhängig**, falls sie nicht linear abhängig sind.

Eine äquivalente Formulierung ist: die Vektoren  $\vec{v}_1, \dots, \vec{v}_k$  sind linear abhängig, wenn es einen dieser Vektoren gibt, den man als Linearkombination der übrigen darstellen kann.

**Beispiele 1.13:** (a) Die Vektoren

$$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^3$$

sind linear abhängig, weil

$$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 2 \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \vec{0}.$$

(b) Die Vektoren  $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$  und  $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2$  sind linear unabhängig, denn aus

$$\alpha_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

folgt

$$\begin{cases} 3\alpha_1 + \alpha_2 = 0, \\ \alpha_1 + 2\alpha_2 = 0 \end{cases} \iff \begin{cases} \alpha_2 = -3\alpha_1, \\ \alpha_1 - 6\alpha_1 = 0 \end{cases} \iff \begin{cases} \alpha_2 = -3\alpha_1, \\ -5\alpha_1 = 0 \end{cases} \iff \begin{cases} \alpha_2 = 0, \\ \alpha_1 = 0. \end{cases}$$

Also kann  $\vec{0}$  aus  $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$  und  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$  nur trivial linear kombiniert werden.

(c)  $\{\vec{0}\}$  ist linear abhängig, denn  $23 \cdot \vec{0} = \vec{0}$ , und  $23 \neq 0$ .

**Satz 1.14:**

Es gilt:

- (a) Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.
- (b) Jede Obermenge einer linear abhängigen Menge ist linear abhängig.

**Definition 1.15 (Basis):**

Eine Menge  $\{\vec{b}_1, \dots, \vec{b}_n\} \subseteq V$  heißt *Basis von V*, falls

(B1)  $\vec{b}_1, \dots, \vec{b}_n$  linear unabhängig ist und

(B2)  $\text{Lin}(\vec{b}_1, \dots, \vec{b}_n) = V$  ist.

**Beispiele 1.16:** (a) Die Einheitsvektoren

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

bilden eine Basis des  $\mathbb{R}^n$  (auch *Standardbasis* genannt), denn:

(B1) gilt: Ist

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{j=1}^n \alpha_j \vec{e}_j = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + \alpha_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix},$$

so folgt  $\alpha_1 = \alpha_2 = \dots = \alpha_n$ . Also ist  $\{\vec{e}_1, \dots, \vec{e}_n\}$  eine linear unabhängige Menge.

**(B2) gilt:** Siehe Teil (c) von Beispiel 1.11.

(b) Es ist

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\}$$

eine Basis des  $\mathbb{R}^3$  (Übungsaufgabe).

**Satz 1.17 (Wichtige Eigenschaften von Basen):** (a) Sei  $V$  ein  $\mathbb{R}$ -Vektorraum mit Basis  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ . Dann gibt es zu jedem  $\vec{v} \in V$  **eindeutig bestimmte** Koeffizienten  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  mit  $\vec{v} = \sum_{j=1}^n \alpha_j \vec{b}_j$ . Diese Zahlen  $\alpha_1, \dots, \alpha_n$  heißen *Koordinaten* von  $\vec{v}$  bezüglich der Basis  $B$ .

(b) Ist  $B$  eine Basis von  $V$  und  $\vec{v}$  ein Vektor aus  $V$ , so ist  $B \cup \{\vec{v}\}$  linear abhängig (es sei denn,  $\vec{v}$  wäre bereits in  $B$  enthalten). Man sagt auch: „Basen sind maximal große linear unabhängige Mengen“.

(c) Ist  $B$  eine Basis von  $V$  und  $C \subsetneq B$ , so ist  $\text{Lin}(C) \subsetneq V$ . Man sagt auch: „Basen sind minimal kleine erzeugende Mengen“.

(d) Je zwei Basen von  $V$  enthalten gleich viele Vektoren.

**Bemerkung 1.18 (Dimension):**

Teil (d) von Satz 1.17 sagt: Ist  $V$  ein Vektorraum, so enthalten alle seine Basen die gleiche Anzahl von Vektoren. Diese Anzahl heißt *Dimension von  $V$* , geschrieben als  $\dim(V)$ .

Im Sonderfall  $V = \{\vec{0}\}$  setzt man  $\dim(V) = 0$ .

**Beispiel 1.19:**

Wir betrachten

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^3$$

und definieren  $U := \text{Lin}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ .

Wegen Satz 1.10 ist  $U$  ein Untervektorraum des  $\mathbb{R}^3$ , und nach Teil (a) von Beispiel 1.13 ist  $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$  eine linear abhängige Menge, denn wir hatten  $\vec{v}_1 - \vec{v}_2 - 2\vec{v}_3 = \vec{0}$  gefunden, bzw.  $\vec{v}_1 = \vec{v}_2 + 2\vec{v}_3$ .

Andererseits ist  $\{\vec{v}_2, \vec{v}_3\}$  linear unabhängig, denn falls

$$\vec{0} = \alpha_2 \vec{v}_2 + \alpha_3 \vec{v}_3 = \begin{bmatrix} \alpha_2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha_3 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_3 \end{bmatrix},$$

so ist  $\alpha_2 = \alpha_3 = 0$ .

Außerdem ist sogar  $U = \text{Lin}(\vec{v}_2, \vec{v}_3)$ , das heißt, dass in der Definition von  $U$  der Vektor  $\vec{v}_1$  eigentlich gar nicht benötigt wird. Denn falls  $\vec{x} \in U$  ein beliebiger Vektor ist, so existieren nach Definition von  $U$  Zahlen  $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$  mit

$$\begin{aligned} \vec{x} &= \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \alpha_3 \vec{v}_3 = \alpha_1 (\vec{v}_2 + 2\vec{v}_3) + \alpha_2 \vec{v}_2 + \alpha_3 \vec{v}_3 \\ &= (\alpha_1 + \alpha_2) \vec{v}_2 + (2\alpha_1 + \alpha_3) \vec{v}_3 \in \text{Lin}(\vec{v}_2, \vec{v}_3). \end{aligned}$$

Wir fassen zusammen: es ist  $\{\vec{v}_2, \vec{v}_3\}$  eine Basis von  $U$ , also ist  $\dim(U) = 2$  und wir haben

$$U = \{ \alpha_2 \vec{v}_2 + \alpha_3 \vec{v}_3 : \alpha_2, \alpha_3 \in \mathbb{R} \} = \left\{ \begin{bmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_3 \end{bmatrix} : \alpha_2, \alpha_3 \in \mathbb{R} \right\}.$$

**Satz 1.20:**

Ist  $V$  ein  $\mathbb{R}$ -Vektorraum und  $U$  ein Untervektorraum von  $V$  mit  $\dim(U) = \dim(V)$ , so ist  $U = V$ .

**Definition 1.21 (Orthogonalbasen und Orthonormalbasen):**

Sei  $V$  ein  $\mathbb{R}$ -Vektorraum und  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  sei eine Basis von  $V$ .

- (a)  $B$  heißt *Orthogonalbasis*, falls  $\langle \vec{b}_i, \vec{b}_j \rangle = 0$  für alle  $i \neq j$  gilt.
- (b)  $B$  heißt *Orthonormalbasis*, falls  $\langle \vec{b}_i, \vec{b}_j \rangle = 0$  für alle  $i \neq j$  gilt und zusätzlich  $\langle \vec{b}_i, \vec{b}_i \rangle = 1$  für alle  $i$ .

Im Falle einer Orthonormalbasis werden beide Bedingungen oft zusammengefasst:

$$\langle \vec{b}_i, \vec{b}_j \rangle = \delta_{ij} = \begin{cases} 0 & : \text{wenn } i \neq j, \\ 1 & : \text{wenn } i = j. \end{cases}$$

Der Ausdruck  $\delta_{ij}$  wird *Kronecker-Delta* genannt.

**Beispiele 1.22:** (a) Für den  $\mathbb{R}^n$  haben wir die Standardbasis  $\vec{e}_1, \dots, \vec{e}_n$ . Diese ist eine Orthonormalbasis für den  $\mathbb{R}^n$ .

(b) Weiterhin gilt:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\} \quad \text{ist eine Orthogonalbasis des } \mathbb{R}^3,$$

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\} \quad \text{ist eine Orthonormalbasis des } \mathbb{R}^3.$$

**Bemerkung 1.23 (Koordinaten bezüglich einer ONB):**

Sei  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  eine Orthonormalbasis von  $\mathbb{R}^n$  und  $\vec{x} \in \mathbb{R}^n$ . Es ist  $B$  eine Basis des  $\mathbb{R}^n$ , also existieren wegen Teil (a) von Satz 1.17 eindeutig bestimmte Zahlen  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  mit  $\vec{x} = \sum_{j=1}^n \alpha_j \vec{b}_j$ . Diese  $\alpha_j$  heißen *Koordinaten von  $\vec{x}$  bezüglich  $B$* .

Für alle  $k = 1, \dots, n$  gilt dann

$$\langle \vec{x}, \vec{b}_k \rangle = \left\langle \sum_{j=1}^n \alpha_j \vec{b}_j, \vec{b}_k \right\rangle \quad \left| \quad \text{das Skalarprodukt ist linear} \right.$$

$$= \sum_{j=1}^n \alpha_j \langle \vec{b}_j, \vec{b}_k \rangle = \alpha_k.$$

Also sind die Koordinaten von  $\vec{x}$  bezüglich der Orthonormalbasis  $B$  leicht zu bestimmen:

$$\alpha_k = \langle \vec{x}, \vec{b}_k \rangle, \quad k = 1, \dots, n.$$

Dies gilt nur für ONB ! Anders formuliert: wegen dieser Eigenschaft sind Orthonormalbasen besonders wertvolle Basen.

## 2 Lineare Gleichungssysteme

**Beispiel 2.1:** (a) Wir suchen die Schnittpunkte dreier Ebenen  $E_1, E_2, E_3$ , die sich im  $\mathbb{R}^3$  befinden:

$$\begin{array}{l} E_1: \quad 2x + y + z = 1 \quad \boxed{1} \\ E_2: \quad 3x + y + z = 2 \quad \boxed{2} \\ E_3: \quad 4x + 2y + 3z = 0 \quad \boxed{3} \end{array}$$

**Gesucht** sind alle gemeinsamen Lösungen aller drei Gleichungen, d.h. alle Lösungen des *Gleichungssystems*. Siehe Abbildung 2.5.

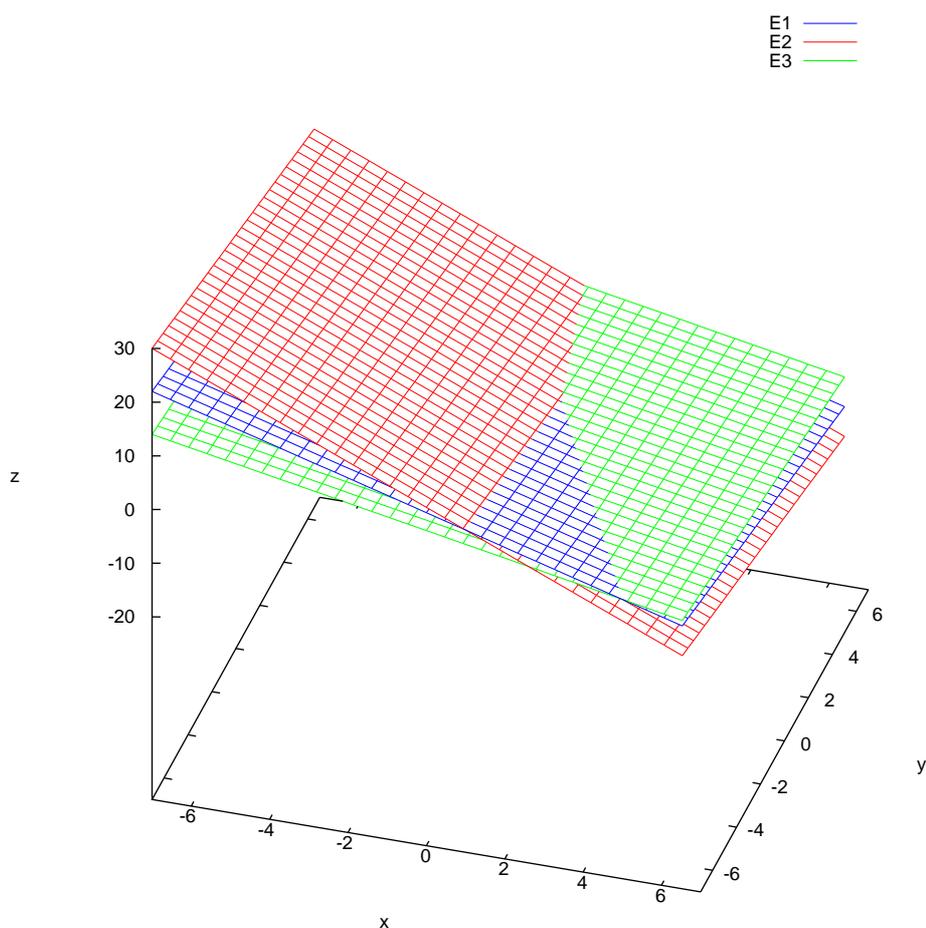


Abbildung 2.5: Die drei Ebenen  $E_1$ ,  $E_2$ ,  $E_3$  schneiden einander in genau einem Punkt. Dieser ist die Lösung des Gleichungssystems in Fall (a).

Wir erinnern an das **Schulvorgehen**:  $2 \cdot \boxed{2} - 3 \cdot \boxed{1}$  liefert:

$$\begin{array}{rcl}
 2 \cdot \boxed{2} & : & 6x + 2y + 2z = 4 \\
 3 \cdot \boxed{1} & : & 6x + 3y + 3z = 3 \\
 \hline
 & & -y - z = 1 \quad \boxed{4}
 \end{array}$$

Weiterhin rechnen wir wie folgt:

$$\begin{array}{rcl}
 \boxed{3} & : & 4x + 2y + 3z = 0 \\
 2 \cdot \boxed{1} & : & 4x + 2y + 2z = 2 \\
 \hline
 & & z = -2 \quad \boxed{5}
 \end{array}$$

Einsetzen von  $\boxed{5}$  in  $\boxed{4}$  ergibt  $-y + 2 = 1$ , d.h.  $y = 1$ . Einsetzen in  $\boxed{1}$  liefert  $2x + 1 - 2 = 1$ , d.h.  $x = 1$ .

Der Schnittpunkt ist also  $\vec{x} = [1, 1, -2]^T$ , und das Gleichungssystem ist eindeutig lösbar. Geometrisch bedeutet dies, dass es genau einen Punkt gibt, der auf den drei Ebenen  $E_1$ ,  $E_2$ ,  $E_3$  gleichzeitig liegt.

(b) Jetzt suchen wir alle Schnittpunkte der beiden Ebenen  $E_1$  und  $E_2$ , siehe Abbildung 2.6:

$$\begin{aligned} E_1: \quad 2x + y + z &= 1 & \boxed{1} \\ E_2: \quad 3x + y + z &= 2 & \boxed{2} \end{aligned}$$

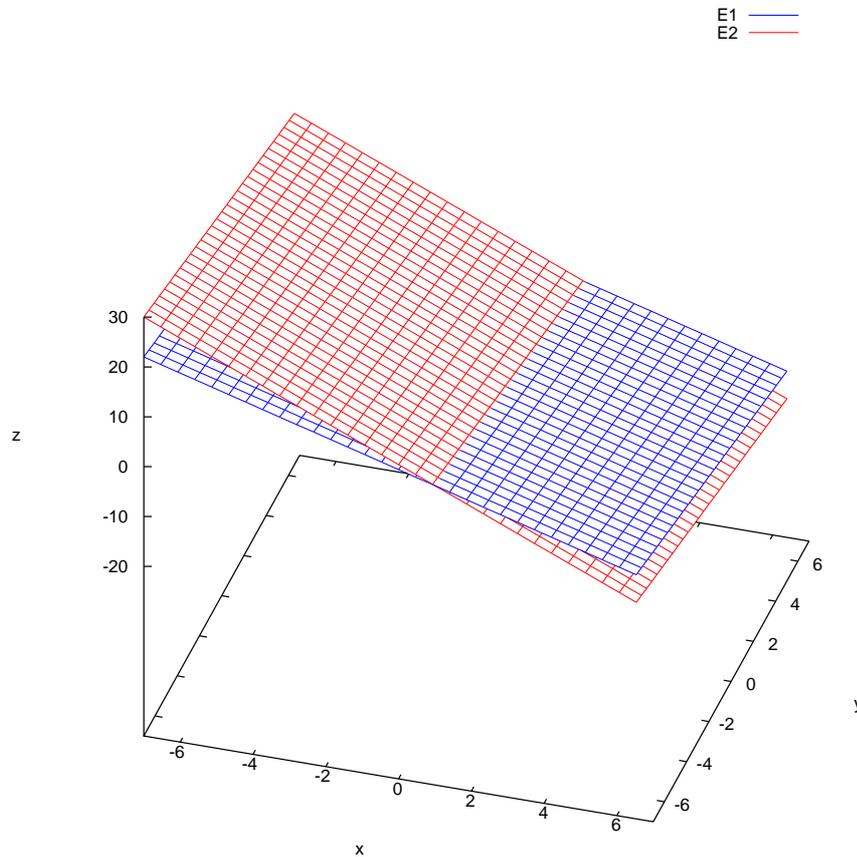


Abbildung 2.6: Die beiden Ebenen  $E_1$  und  $E_2$  schneiden einander in einer Geraden. Diese ist die Lösungsmenge des Gleichungssystems in Fall (b).

Wie oben ergibt  $2 \cdot \boxed{2} - 3 \cdot \boxed{1}$  dann die Gleichung

$$-y - z = 1 \quad \boxed{4},$$

und weitere Bedingungen sind nicht vorhanden. Wir benennen  $z = t$  für  $t \in \mathbb{R}$ . Dann liefert  $\boxed{4}$ , dass

$$-y - t = 1, \quad \text{d.h.} \quad y = -1 - t.$$

Und aus  $\boxed{1}$  erhalten wir

$$1 = 2x + (-1 - t) + t = 2x - 1, \quad \text{d.h.} \quad x = 1.$$

Also sind alle Schnittpunkte von  $E_1$  und  $E_2$  gegeben durch

$$\vec{x} = \begin{bmatrix} 1 \\ -1 - t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Die Lösungsmenge des Gleichungssystems ist also eine Gerade. Entlang dieser Geraden ist immer  $x = 1$ , was mit Abbildung 2.6 recht gut übereinstimmt.

- (c) Wir bestimmen wieder die Schnittpunkte von drei Ebenen, wobei wir die Ebene  $E_3$  gegen eine Ebene  $E_4$  austauschen:

$$E_1: \quad 2x + y + z = 1 \quad \boxed{1}$$

$$E_2: \quad 3x + y + z = 2 \quad \boxed{2}$$

$$E_4: \quad x + y + z = 3 \quad \boxed{3'}$$

Wie oben bilden wir  $2 \cdot \boxed{2} - 3 \cdot \boxed{1}$  und erhalten

$$-y - z = 1 \quad \boxed{4}.$$

Jetzt können wir aber auch  $2 \cdot \boxed{3'} - \boxed{1}$  bilden und es entsteht

$$y + z = 5 \quad \boxed{5}.$$

Wir bilden  $\boxed{4} + \boxed{5}$  und es ergibt sich

$$0 = 6.$$

Das kann nicht sein, also ist das Gleichungssystem unlösbar.

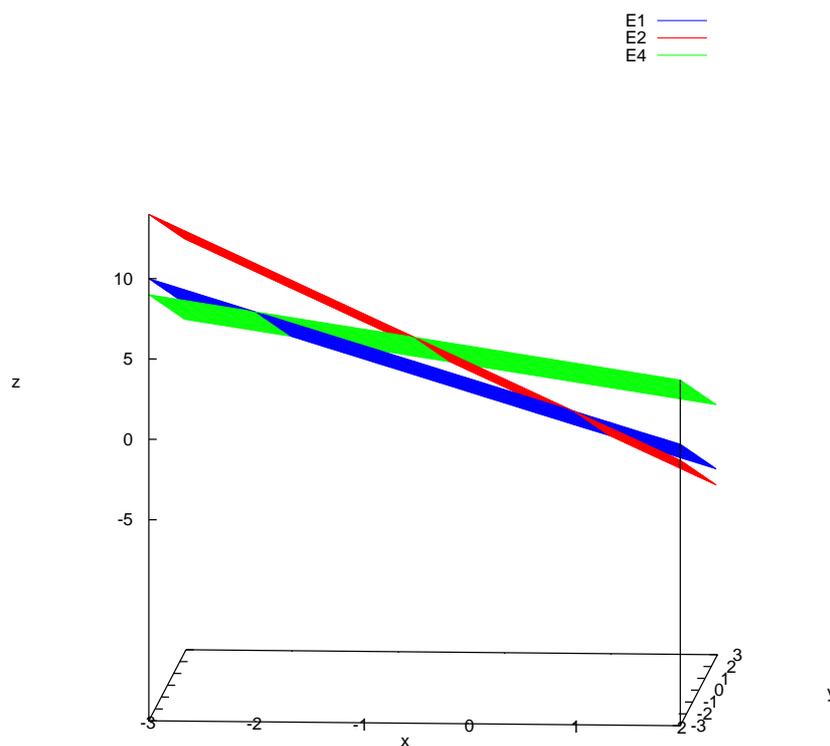


Abbildung 2.7: Die drei Ebenen  $E_1$ ,  $E_2$ ,  $E_4$  haben keinen Punkt gemeinsam. Es gibt zwar eine Schnittgerade zwischen z.B.  $E_2$  und  $E_4$ , aber  $E_1$  ist zu dieser Schnittgeraden parallel. Deshalb ist die Lösungsmenge des Gleichungssystems in Fall (c) die leere Menge.

Unser **Ziel** ist es jetzt, ein systematisches algorithmisches Verfahren zum Lösen beliebiger solcher linearer Gleichungssysteme zu entwickeln.

**Definition 2.2 (Lineares Gleichungssystem):**

Die Gleichungen

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

mit gegebenen Koeffizienten  $a_{ij}$  (wobei  $i = 1, \dots, m$  und  $j = 1, \dots, n$ ) und gegebenen rechten Seiten  $b_i$  (wobei  $i = 1, \dots, m$ ) und gesuchten Unbekannten  $x_j$  (wobei  $j = 1, \dots, n$ ) heißt *lineares Gleichungssystem (LGS)*.

Dieses Gleichungssystem heißt

**unterbestimmt**, falls  $m < n$  (es gibt weniger Gleichungen als Variable, siehe Beispiel 2.1 (b)),

**quadratisch**, falls  $m = n$  (es gibt gleich viele Gleichungen wie Variable, siehe Beispiel 2.1 (a) und (c)),

**überbestimmt**, falls  $m > n$  (es gibt mehr Gleichungen als Variable).

**Bemerkung 2.3 (Matrixform des LGS):**

Die wesentliche Information steckt in den  $a_{ij}$  und  $b_i$ , mit  $i = 1, \dots, m$  und  $j = 1, \dots, n$ . Zur Einsparung von Schreibarbeit fassen wir zusammen:

$$\vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (\text{Matrix des LGS}).$$

Dann schreiben wir das LGS in der Kurzform  $A\vec{x} = \vec{b}$ .

**Beispiele 2.4:** (a) Beispiel 2.1 (a) lautete

$$\begin{aligned} 2x + y + z &= 1, \\ 3x + y + z &= 2, \\ 4x + 2y + 3z &= 0, \end{aligned}$$

und die Kurzschreibweise davon ist  $A\vec{x} = \vec{b}$  mit

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \\ 4 & 2 & 3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Es liegt ein quadratisches LGS vor.

(b) Beispiel 2.1 (b) lautete

$$\begin{aligned} 2x + y + z &= 1, \\ 3x + y + z &= 2, \end{aligned}$$

die Kurzschreibweise ist erneut  $A\vec{x} = \vec{b}$ , aber jetzt haben wir

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Das Gleichungssystem ist unterbestimmt.

**Bemerkung 2.5 (Elementarumformungen):**

Das LGS  $A\vec{x} = \vec{b}$  kann schematisch wie folgt dargestellt werden:

$$\begin{array}{l}
 \boxed{1} : \\
 \boxed{2} : \\
 \vdots \\
 \boxed{m} :
 \end{array}
 \begin{array}{cccc|c}
 & x_1 & x_2 & \cdots & x_n & \vec{b} \\
 \hline
 & a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
 & a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
 & \vdots & \vdots & \ddots & \vdots & \vdots \\
 & a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
 \end{array}$$

In der Kopfzeile stehen die Bezeichnungen der Variablen, und in der linken Spalte die Zeilennamen.

Die folgenden Aktionen (*Elementarumformungen* genannt) ändern die Lösungsmenge des LGS nicht und können deshalb verwendet werden, das LGS zu vereinfachen und schlussendlich zu lösen:

- (1) **Zeilen vertauschen:** wir tauschen  $\boxed{i} \leftrightarrow \boxed{j}$
- (2) **Spalten tauschen:** wir tauschen Spalte  $i$  mit Spalte  $j$ , wobei wir die Einträge  $x_i$  und  $x_j$  in der Kopfzeile nicht vergessen dürfen.
- (3) **Linearkombination:** Wir ersetzen die  $i$ -te Zeile durch  $p$  mal die  $i$ -te Zeile plus  $q$  mal die  $j$ -Zeile für  $p \neq 0$  und  $j \neq i$ :

$$p \cdot \boxed{i} + q \cdot \boxed{j} \rightsquigarrow \boxed{i}.$$

Beachte:  $q = 0$  ist erlaubt und liefert dann die Multiplikation der  $i$ -ten Zeile mit  $p$ :

$$p \cdot \boxed{i} \rightsquigarrow \boxed{i}, \quad p \neq 0.$$

**Beispiel 2.6:**

Das LGS aus Beispiel 2.1 (a) war

$$2x + y + z = 1,$$

$$3x + y + z = 2,$$

$$4x + 2y + 3z = 0.$$

Im Schema wird dies zu

$$\begin{array}{l}
 \boxed{1} : \\
 \boxed{2} : \\
 \boxed{3} :
 \end{array}
 \begin{array}{ccc|c}
 & x & y & z & \\
 \hline
 & 2 & 1 & 1 & 1 \\
 & 3 & 1 & 1 & 2 \\
 & 4 & 2 & 3 & 0
 \end{array}$$

Jetzt benutzen wir zweimal die Elementarumformung (3):

$$\begin{array}{l}
 2 \cdot \boxed{2} - 3 \cdot \boxed{1} =: \boxed{4} \\
 1 \cdot \boxed{3} - 2 \cdot \boxed{1} =: \boxed{5}
 \end{array}
 \begin{array}{ccc|c}
 & x & y & z & \\
 \hline
 & 2 & 1 & 1 & 1 \\
 & 0 & -1 & -1 & 1 \\
 & 0 & 0 & 1 & -2
 \end{array}$$

Wir beobachten im  $3 \times 3$ -Feld eine schöne treppenförmige Anordnung von Nullen; man redet auch von der *Zeilenstufenform*.

Nun haben wir zwei Möglichkeiten, wie wir weiter verfahren können.

1. **Möglichkeit, sukzessives Auflösen:** wir behandeln die Gleichungen im Schema von unten nach oben:

$$\boxed{5} : z = -2,$$

$$\boxed{4} : -y - z = 1 \implies -y + 2 = 1 \implies y = 1,$$

$$\boxed{1} : 2x + y + z = 1 \implies 2x + 1 - 2 = 1 \implies x = 1.$$

Die Lösung ist  $\vec{x} = [1, 1, -2]^T$ .

**2. Möglichkeit, weitere Elementarumformungen:** wir erzeugen noch mehr Nullen im Schema:

$$\begin{array}{l} \boxed{1} + \boxed{4} =: \boxed{6} : \\ \boxed{4} + \boxed{5} =: \boxed{7} : \\ \boxed{5} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 0 & 0 & 2 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{array}$$

und jetzt bringen wir Einsen auf die Diagonale:

$$\begin{array}{l} \frac{1}{2} \cdot \boxed{6} =: \boxed{8} : \\ -1 \cdot \boxed{7} =: \boxed{9} : \\ \boxed{5} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}$$

Damit ist die Lösung dann  $\vec{x} = [1, 1, -2]^\top$ .

**Beispiel 2.7:**

Das Beispiel 2.1 (b) bringt uns auf das Schema

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \end{array}$$

Mit einer Elementarumformung stellen wir die Zeilenstufenform her:

$$2 \cdot \boxed{2} - 3 \cdot \boxed{1} =: \boxed{3} : \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \end{array}$$

Jetzt ist eine der Variablen  $y$  und  $z$  frei wählbar. Wir entscheiden uns dafür,  $z$  frei zu wählen. Also schreiben wir  $z = t$ . Dann liefert uns  $\boxed{3}$ , dass  $-y - t = 1$ , also  $y = -1 - t$ . Anschließend schließen wir aus  $\boxed{1}$ , dass  $2x - 1 - t + t = 1$ , also  $2x = 2$ , also  $x = 1$ .

Die Lösungsmenge ist

$$\left\{ \begin{bmatrix} 1 \\ -1-t \\ t \end{bmatrix} : t \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

Geometrisch ist dies eine Gerade im dreidimensionalen Raum, die nicht durch den Ursprung geht.

**Beispiel 2.8:**

Wir behandeln Beispiel 2.1 (c) erneut, diesmal allerdings in der neuen Schreibweise:

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \boxed{3} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{array} \quad \rightsquigarrow \quad \begin{array}{l} 2 \cdot \boxed{2} - 3 \cdot \boxed{1} =: \boxed{4} : \\ 2 \cdot \boxed{3} - 1 \cdot \boxed{1} =: \boxed{5} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & 5 \end{array}$$

Die Zeilenstufenform entsteht nach einer weiteren Elementarumformung:

$$\begin{array}{l} \boxed{4} + \boxed{5} =: \boxed{6} : \end{array} \quad \begin{array}{ccc|c} x & y & z & \\ \hline 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 6 \end{array}$$

Die letzte Zeile besagt  $0 = 6$ , also gibt es keine Lösung.

**Bemerkung 2.9 (Zeilenstufenform):**

Das Ziel der Elementarumformungen ist

$\tilde{x}_1$	$\tilde{x}_2$	$\dots$	$\tilde{x}_r$	$\tilde{x}_{r+1}$	$\dots$	$\tilde{x}_n$	$\vec{b}$
•	*	$\dots$	*	*	$\dots$	*	*
0	•	$\dots$	*	*	$\dots$	*	*
0	0	$\dots$	*	*	$\dots$	*	*
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
0	0	$\dots$	•	*	$\dots$	*	*
0	0	$\dots$	0	0	$\dots$	0	×
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
0	0	$\dots$	0	0	$\dots$	0	×

Hierbei sind die  $\tilde{x}_1, \dots, \tilde{x}_n$  eine Umordnung der gesuchten  $x_1, \dots, x_n$ , die durch Spaltenvertauschungen sich ergibt. Die Einträge \* und × sind beliebige Werte, und • sind Werte ungleich Null.

Aus dieser Zeilenstufenform kann die Lösungsmenge folgendermaßen bestimmt werden:

1. **Fall: einer der Einträge × ist  $\neq 0$ :** dann gibt es keine Lösung.
2. **Fall: alle Einträge × sind Null:** dann setzen wir  $\tilde{x}_{r+1}, \dots, \tilde{x}_n = t_{n-r}$  und lösen die Gleichungen schrittweise von unten nach oben auf, womit wir nacheinander  $\tilde{x}_r, \dots, \tilde{x}_1$  bestimmen.

**Verfahren 2.10 (Gauß-Algorithmus):**

(Verfahren zum Erreichen der Zeilenstufenform)

1. **Schritt:** wir suchen ein  $a_{ij} \neq 0$ . Wir vertauschen die 1. und die  $j$ -te Spalte; und wir vertauschen die 1. und die  $i$ -te Zeile (das bewirkt, dass das vorherige  $a_{ij}$  in der Ecke oben links steht).
2. **Schritt:** wir ersetzen alle Zeilen  $\boxed{i}$  ab der zweiten durch  $a_{11} \cdot \boxed{i} - a_{i1} \cdot \boxed{1}$ . Damit erreichen wir, dass die erste Spalte jetzt  $(a_{11}, 0, 0, \dots, 0)^\top$  lautet.
3. **Schritt:** wir behandeln das Rest-LGS (ohne die erste Zeile und ohne die erste Spalte) wie vor.

**Beispiel 2.11:**

Für einen reellen Parameter  $\alpha \in \mathbb{R}$  untersuchen wir das folgende LGS:

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\vec{b}$
$\boxed{1}$ :	1	1	3	-2	4	-1
$\boxed{2}$ :	-1	-1	-3	2	-4	$\alpha$
$\boxed{3}$ :	0	0	-2	1	0	3
$\boxed{4}$ :	0	0	4	-2	0	-6

Wir behandeln die erste Spalte:

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
$\boxed{1}$ :	1	1	3	-2	4	-1
$\boxed{1} + \boxed{2}$ =:	$\boxed{5}$	0	0	0	0	$\alpha - 1$
$\boxed{3}$ :	0	0	-2	1	0	3
$\boxed{4}$ :	0	0	4	-2	0	-6

Um die zweite Zeile in die gewünschte Form zu bringen, wird in den Zeilen  $\boxed{5}, \boxed{3}, \boxed{4}$  ein von Null verschiedener Eintrag gesucht. Wir wählen z.B. den Eintrag  $a_{34} = 1$  (Einsen sind rechentechnisch schöner).

Vertauschung der zweiten und der vierten Spalte sowie der Zeilen  $\boxed{5}$  und  $\boxed{3}$  ergibt

	$x_1$	$x_4$	$x_3$	$x_2$	$x_5$	
$\boxed{1}$ :	1	-2	3	1	4	-1
$\boxed{3}$ :	0	1	-2	0	0	3
$\boxed{5}$ :	0	0	0	0	0	$\alpha - 1$
$\boxed{4}$ :	0	-2	4	0	0	-6

Nun werden die Einträge der Zeilen  $\boxed{5}$ ,  $\boxed{4}$  in der zweiten Spalte zu Null gemacht:

$$\begin{array}{r}
 \boxed{1} : \\
 \boxed{3} : \\
 \boxed{5} : \\
 2 \cdot \boxed{3} + \boxed{4} =: \boxed{6} :
 \end{array}
 \begin{array}{ccccc|c}
 x_1 & x_4 & x_3 & x_2 & x_5 & \\
 \hline
 1 & -2 & 3 & 1 & 4 & -1 \\
 0 & 1 & -2 & 0 & 0 & 3 \\
 0 & 0 & 0 & 0 & 0 & \alpha - 1 \\
 0 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

Damit ist die Zeilenstufenform erreicht, und jetzt ermitteln wir die Lösungen  $\vec{x}$  (soweit vorhanden):

**1. Fall: wenn  $\alpha \neq 1$ ,** dann gibt es keine Lösung.

**2. Fall: wenn  $\alpha = 1$ ,** dann setzen wir  $x_3 = t_1$ ,  $x_2 = t_2$ ,  $x_5 = t_3$  und bestimmen die noch fehlenden Lösungsbestandteile  $x_1$  und  $x_4$  wie folgt:

$$\begin{array}{l}
 \boxed{3} : \quad x_4 - 2t_1 = 3 \quad \implies \quad x_4 = 3 + 2t_1, \\
 \boxed{1} : \quad x_1 - 2x_4 + 3t_1 + t_2 + 4t_3 = -1 \\
 \implies \quad x_1 - 6 - 4t_1 + 3t_1 + t_2 + 4t_3 = -1 \\
 \implies \quad x_1 = 5 + t_1 - t_2 - 4t_3.
 \end{array}$$

Somit erhalten wir insgesamt (im Falle  $\alpha = 1$ )

$$\vec{x} = \begin{bmatrix} 5 + t_1 - t_2 - 4t_3 \\ t_2 \\ t_1 \\ 3 + 2t_1 \\ t_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad t_1, t_2, t_3 \in \mathbb{R}.$$

**Definition 2.12 (Homogene und inhomogene LGS):**

Das LGS  $A\vec{x} = \vec{b}$  heißt *homogen*, falls  $\vec{b} = \vec{0}$ . Ansonsten heißt das LGS *inhomogen*.

**Satz 2.13 (Struktur der Lösungsmenge von LGS):**

Sei  $A\vec{x} = \vec{b}$  ein LGS und  $n$  die Spaltenzahl der Matrix  $A$ .

(a) Ist  $\vec{b} = \vec{0}$  (d.h. es liegt ein homogenes LGS vor), so ist die Lösungsmenge

$$\ker(A) := \left\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0} \right\}$$

ein Untervektorraum von  $\mathbb{R}^n$ , der als *Kern von A* bezeichnet wird.

(b) Ist  $\vec{x}_p$  eine Lösung von  $A\vec{x} = \vec{b}$  (genannt *Partikulärlösung* oder *spezielle Lösung*), so ist die gesamte Lösungsmenge von  $A\vec{x} = \vec{b}$  gegeben durch

$$\{ \vec{x}_p + \vec{y} : \vec{y} \in \ker(A) \} \quad \text{„} = \vec{x}_p + \ker(A)\text{“}.$$

Insbesondere ist die Lösungsmenge jedes LGS entweder leer oder ein verschobener Untervektorraum, d.h. ein Punkt, eine Gerade, eine Ebene, ... Die Dimension von  $\ker(A)$  ist gleich der üblichen Dimension des verschobenen Unterraums (Punkte sind nulldimensional, Geraden sind eindimensional, Ebenen sind zweidimensional ...).

(c) Ist  $r$  die maximale Anzahl linear unabhängiger Spalten von  $A$ , so gilt die Dimensionsformel

$$n = r + \dim(\ker(A)),$$

mit  $n$  als der Gesamtzahl der Spalten von  $A$ .

Dabei ist  $r$  gleich der Anzahl der Zeilen, die nicht nur Nullen enthalten, in der Zeilenstufenform, vgl. Bemerkung 2.9. Diese Zahl  $r$  heißt *Rang von A*, geschrieben als  $\text{rang}(A)$ .

**Beispiel 2.14:**

Wir betrachten erneut das LGS aus Beispiel 2.11 mit  $\alpha = 1$ , d.h.

$$A = \begin{bmatrix} 1 & 1 & 3 & -2 & 4 \\ -1 & -1 & -3 & 2 & -4 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 0 \end{bmatrix}.$$

Wir untersuchen das homogene System, also  $A\vec{x} = \vec{0}$ . Dies führt uns auf die Zeilenstufenform (siehe Beispiel 2.11 für den Rechnungsverlauf)

$$\begin{array}{ccccc|c} x_1 & x_4 & x_3 & x_2 & x_5 & \\ \hline 1 & -2 & 3 & 1 & 4 & 0 \\ 0 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Also ist  $\text{rang}(A) = r = 2$  und  $\dim(\ker(A)) = n - r = 5 - 2 = 3$ . Es ist (siehe Beispiel 2.11)

$$\ker(A) = \text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right).$$

Jetzt untersuchen wir das inhomogene System  $A\vec{x} = \vec{b}$ . Zum Beispiel ist

$$\vec{x}_p = \begin{bmatrix} 0 \\ 0 \\ -5 \\ -7 \\ 0 \end{bmatrix}$$

eine spezielle Lösung, wie man durch eine Probe nachprüfen kann. Also sind alle Lösungen des inhomogenen Systems  $A\vec{x} = \vec{b}$  nach Satz 2.13, Teil (b) genau die Elemente der Menge

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ -5 \\ -7 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} : t_1, t_2, t_3 \in \mathbb{R} \right\}.$$

Im Vergleich zur Lösungsdarstellung in Beispiel 2.11 haben wir den Aufpunkt gewechselt, aber die Lösungsmengen sind gleich.

### 3 Matrizenrechnung

**Definition 3.1 (Matrizen):** (a) Ein rechteckiges Schema reeller (bzw. komplexer) Zahlen mit  $m$  Zeilen und  $n$  Spalten heißt  $m \times n$ -Matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]_{i=1, \dots, m; j=1, \dots, n}$$

(b) Die Einträge  $a_{ij}$  heißen *Koeffizienten* oder *Komponenten* der Matrix.

(c) Die Menge aller reellen  $m \times n$ -Matrizen schreiben wir als  $\mathbb{R}^{m \times n}$ . Es ist also  $A \in \mathbb{R}^{m \times n}$ .

(d) Wir untergliedern diese Matrix  $A$  spaltenweise,

$$A = [\vec{a}^1 \quad \vec{a}^2 \quad \dots \quad \vec{a}^n]$$

mit den *Spaltenvektoren*

$$\vec{a}^j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{R}^m, \quad j = 1, \dots, n.$$

Und wir untergliedern die Matrix  $A$  zeilenweise,

$$A = \begin{bmatrix} \vec{a}_1^\top \\ \vec{a}_2^\top \\ \vdots \\ \vec{a}_m^\top \end{bmatrix}$$

mit den *Zeilenvektoren*

$$\vec{a}_i^\top = [a_{i1} \quad a_{i2} \quad \dots \quad a_{in}], \quad i = 1, \dots, m$$

wobei  $\vec{a}_i \in \mathbb{R}^n$ .

(e) Ist  $m = n$ , so heißt  $A$  *quadratisch*.

(f) Die Matrix

$$0_{m,n} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}$$

heißt *Nullmatrix*.

### Bemerkung 3.2:

Vektoren aus dem  $\mathbb{R}^m$  sind spezielle Matrizen mit einer Spalte und  $m$  Zeilen, also  $(m \times 1)$ -Matrizen.

### Beispiel 3.3:

Sei

$$A = \begin{pmatrix} 3 & 1 & 4 & 5 \\ 1 & 5 & 2 & 0 \\ 0 & 7 & 3 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 4}.$$

Dann haben wir

$$a_{32} = 7, \quad \vec{a}^3 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{a}_2^\top = [1 \quad 5 \quad 2 \quad 0].$$

### Bemerkung 3.4 (Addition und Skalarmultiplikation):

Seien  $A = [a_{ij}]_{i=1, \dots, m; j=1, \dots, n}$  und  $B = [b_{ij}]_{i=1, \dots, m; j=1, \dots, n} \in \mathbb{R}^{m \times n}$  sowie  $\lambda \in \mathbb{R}$ . Dann definieren wir  $A + B \in \mathbb{R}^{m \times n}$  und  $\lambda A \in \mathbb{R}^{m \times n}$  durch komponentenweise Verknüpfung:

$$A + B := [a_{ij} + b_{ij}]_{i=1, \dots, m; j=1, \dots, n}, \quad \lambda A := [\lambda a_{ij}]_{i=1, \dots, m; j=1, \dots, n}.$$

### Beispiel 3.5:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 4 & 6 & 9 \end{bmatrix},$$

$$3 \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{ist nicht definiert !}$$

**Rechenregeln 3.6:**

Es ist  $\mathbb{R}^{m \times n}$  ein Vektorraum. Es gelten also (V1)–(V8) aus den Regeln 1.1. Insbesondere gilt für  $A, B, C \in \mathbb{R}^{m \times n}$  und  $\lambda, \mu \in \mathbb{R}$ , dass

$$A + B = B + A, \quad (, + \text{ ist kommutativ}) \quad (\text{V1})$$

$$(A + B) + C = A + (B + C), \quad (, + \text{ ist assoziativ}) \quad (\text{V2})$$

$$\lambda \cdot (\mu \cdot A) = (\lambda \cdot \mu) \cdot A, \quad (\text{V6})$$

$$\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B, \quad (\text{V7})$$

$$(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A. \quad (\text{V8})$$

**Definition 3.7 (Matrixmultiplikation):**

Seien  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{n \times p}$ , mit  $A = [a_{ij}]_{i=1, \dots, m; j=1, \dots, n}$  und  $B = [b_{ij}]_{j=1, \dots, n; k=1, \dots, p}$ .

Dann definieren wir das Matrixprodukt  $AB = A \cdot B \in \mathbb{R}^{m \times p}$  mit  $AB = [c_{ik}]_{i=1, \dots, m; k=1, \dots, p}$  durch

$$c_{ik} := \langle \vec{a}_i, \vec{b}^k \rangle = \sum_{j=1}^n a_{ij} b_{jk}. \quad (*)$$

Es ist also  $c_{ik}$  das Skalarprodukt der  $i$ -ten Zeile von  $A$  mit der  $k$ -ten Spalte von  $B$ .

Die **Merkregel lautet**

$$m \times n \quad \cdot \quad n \times p = m \times p$$

beziehungsweise (in einer zweiten, etwas graphischeren Formulierung)

$$m \underbrace{\left\{ \begin{array}{c} \left[ \begin{array}{c} \rightarrow \\ \hline \end{array} \right] \\ n \end{array} \right\}} \cdot \underbrace{\left\{ \begin{array}{c} \left[ \begin{array}{c} \downarrow \\ \hline \end{array} \right] \\ p \end{array} \right\}}_p \quad \text{ergibt} \quad m \underbrace{\left\{ \begin{array}{c} \left[ \begin{array}{c} * \\ \hline \end{array} \right] \\ p \end{array} \right\}}$$

Die beiden Pfeile werden im Sinne eines Skalarprodukts miteinander multipliziert, und das Ergebnis dieses Skalarprodukts schreibt man an diejenige Stelle  $*$ , wo die beiden Pfeile einander kreuzen.

Man beachte, dass von den vier Zahlen  $m, n, n, p$  (auf der linken Seite der ersten Formulierung der Merkregel) die beiden inneren Zahlen gleich sind; ansonsten gäbe es das Produkt  $AB$  gar nicht. Es ist also das Matrixprodukt nur dann definiert, wenn die Anzahl der Spalten der linken Matrix gleich der Anzahl der Zeilen der rechten Matrix ist. Und auch in Gleichung  $(*)$  sehen wir auf der rechten Seite die Indizes  $i, j, j, k$  in dieser Reihenfolge. Die beiden inneren heißen gleich und sind der Laufindex.

**Beispiel 3.8:**

Wir betrachten

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 3}, \quad B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{3 \times 3}.$$

Dann haben wir

$$AB = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -3 & 3 & 14 \\ 6 & -1 & 4 \end{bmatrix},$$

$$BB = B^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 7 \\ -10 & 4 & 7 \\ -15 & -2 & 3 \end{bmatrix}.$$

Die Produkte  $AA$  und  $BA$  sind nicht definiert. Man beachte, daß  $A \cdot B$  sich völlig anders verhält als  $B \cdot A$ !

**Bemerkung 3.9 (Spezialfall eines Vektors):**

Wir überlegen uns, was passiert, wenn einer der beiden Faktoren im Matrixprodukt ein Vektor ist.

(a) Sei  $A \in \mathbb{R}^{m \times n}$  und  $\vec{x} \in \mathbb{R}^n = \mathbb{R}^{n \times 1}$ . Dann ist

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}$$

$\cong$  linke Seite des LGS  $A\vec{x} = \vec{b}$  mit  $\vec{b} \in \mathbb{R}^m$ .

Das rechtfertigt im Nachhinein die Schreibweise  $A\vec{x} = \vec{b}$  als Matrixprodukt.

(b) Seien  $\vec{x}, \vec{y} \in \mathbb{R}^n$ . Dann ist  $\vec{x}^\top \in \mathbb{R}^{1 \times n}$  und  $\vec{y} \in \mathbb{R}^{n \times 1}$ , also  $\vec{x}^\top \cdot \vec{y} \in \mathbb{R}^{1 \times 1} = \mathbb{R}$ , definiert mit

$$\vec{x}^\top \cdot \vec{y} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \dots + x_ny_n = \sum_{j=1}^n x_jy_j = \langle \vec{x}, \vec{y} \rangle.$$

Das ist genau das bereits bekannte Skalarprodukt im  $\mathbb{R}^n$ .

(c) Sind  $\vec{x} \in \mathbb{R}^n$ ,  $\vec{y} \in \mathbb{R}^m$ , so ist  $\vec{x} \in \mathbb{R}^{n \times 1}$  und  $\vec{y}^\top \in \mathbb{R}^{1 \times m}$ , also ist  $\vec{x}\vec{y}^\top \in \mathbb{R}^{n \times m}$  definiert durch

$$\vec{x} \cdot \vec{y}^\top = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & \dots & y_m \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_m \\ x_2y_1 & x_2y_2 & \dots & x_2y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \dots & x_ny_m \end{bmatrix}.$$

**Beispiel 3.10:**

Wir betrachten

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Dann berechnen wir

$$A\vec{x} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 3 \end{bmatrix},$$

$$\vec{y}^\top A = \begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 5 \end{bmatrix},$$

$$(\vec{y}^\top A)\vec{x} = \begin{bmatrix} 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = 19,$$

$$\vec{x} \cdot \vec{y}^\top = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ 4 & -2 \\ 2 & -1 \end{bmatrix}.$$

**Rechenregeln 3.11:**

Solange die Produkte definiert sind, gilt für Matrizen  $A, B, C$  und Skalare  $\lambda \in \mathbb{R}$ :

$$(A \cdot B) \cdot C = A \cdot (B \cdot C), \quad (\text{Assoziativität})$$

$$\lambda \cdot (A \cdot B) = (\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B),$$

$$A \cdot (B + C) = AC + AC \quad \text{und}$$

$$(A + B) \cdot C = AC + BC, \quad (\text{Distributivgesetze})$$

$$A \cdot 0 = 0 \quad \text{und} \quad 0 \cdot A = 0 \quad \text{mit } 0 \text{ als Nullmatrix.}$$

**Warnung 3.12:**

Es gilt im Allgemeinen nicht  $AB = BA$ . Und es gilt auch nicht: wenn  $AB = 0$ , dann  $A = 0$  oder  $B = 0$ .

**Beispiel 3.13:**

Wir betrachten

$$A = \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Dann haben wir

$$\begin{aligned} \text{einerseits } AB &= \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \text{andererseits } BA &= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix}. \end{aligned}$$

**Definition 3.14 (Transposition):**

Sei  $A = [a_{ij}]_{i=1, \dots, m; j=1, \dots, n} \in \mathbb{R}^{m \times n}$ . Dann heißt

$$A^\top := [a_{ji}]_{i=1, \dots, n; j=1, \dots, m} \in \mathbb{R}^{n \times m}$$

die *transponierte Matrix* von  $A$ . Es entsteht  $A^\top$  aus  $A$  durch Vertauschen von Zeilen und Spalten.

$$\text{Ist } A = \begin{bmatrix} \vec{a}_1^\top \\ \vec{a}_2^\top \\ \vdots \\ \vec{a}_m^\top \end{bmatrix}, \quad \text{so ist } A^\top = [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_m].$$

Es wird also die erste Zeile von  $A$  zur ersten Spalte von  $A^\top$  gemacht, die zweite Zeile von  $A$  wird zur zweiten Spalte von  $A^\top$  gemacht usw.

**Rechenregeln 3.15:**

Seien  $A, B$  Matrizen und  $\vec{x}, \vec{y}$  Vektoren. Dann gilt (solange die Produkte bzw. Summen definiert sind):

- (a)  $(A^\top)^\top = A,$
- (b)  $(\lambda A)^\top = \lambda \cdot A^\top,$
- (c)  $(A + B)^\top = A^\top + B^\top,$
- (d)  $(AB)^\top = B^\top A^\top$  (auf Reihenfolge achten!),
- (e)  $\langle A\vec{x}, \vec{y} \rangle = (A\vec{x})^\top \vec{y} = \vec{x}^\top A^\top \vec{y} = \langle \vec{x}, A^\top \vec{y} \rangle.$

**Definition 3.16 (Symmetrische und antisymmetrische Matrizen):**

Eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *symmetrisch*, wenn  $A^\top = A$ . Sie heißt *antisymmetrisch*, wenn  $A^\top = -A$ .

Es gelten folgende **Regeln**:

- Wenn  $A$  und  $B$  beide symmetrisch sind (bzw. beide antisymmetrisch), dann ist auch  $A + B$  symmetrisch (bzw. antisymmetrisch), denn wegen Regeln 3.15 (c) haben wir, dass

$$(A + B)^\top = A^\top + B^\top = A + B.$$

- Wenn  $A$  eine beliebige quadratische Matrix ist, dann sind  $A + A^\top$  und  $AA^\top$  beide immer symmetrisch, denn wir haben

$$\begin{aligned} (A + A^\top)^\top &\stackrel{(c)}{=} A^\top + A^{\top\top} \stackrel{(a)}{=} A^\top + A = A + A^\top, \\ (AA^\top)^\top &\stackrel{(d)}{=} A^{\top\top} A^\top \stackrel{(a)}{=} AA^\top. \end{aligned}$$

Es gibt eine **Stolperfalle**: wenn  $A$  und  $B$  symmetrisch sind, dann braucht  $AB$  nicht symmetrisch sein.

**Beispiel 3.17:**

Für die quadratische Matrix

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{haben wir} \quad A^\top = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

und es ergibt sich

$$A + A^\top = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

$$A^\top A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 1 \\ 6 & 8 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Es sind also  $A^\top + A$  und  $A^\top A$  tatsächlich beide symmetrisch.

**Verfahren 3.18 (Lösung von Matrixgleichungssystemen):**

Seien  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{m \times p}$  gegeben. Gesucht ist  $X \in \mathbb{R}^{n \times p}$  mit  $AX = B$ . Die Rechnung verfolgt analog zu LGSen.

**Beispiel 3.19:**

Wir betrachten

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{3 \times 3}, \quad B = \begin{bmatrix} 2 & 7 \\ 3 & 1 \\ 3 & 5 \end{bmatrix}$$

und suchen  $X \in \mathbb{R}^{3 \times 2}$  mit  $AX = B$ . Dazu schreiben wir unser übliches Schema auf, und in die Kopfzeile schreiben wir die Namen der Zeilenvektoren der Matrix  $X$ :

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \boxed{3} : \end{array} \begin{array}{ccc|cc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & & \\ \hline 2 & 0 & 1 & 2 & 7 \\ 1 & 2 & 1 & 3 & 1 \\ 2 & 1 & 2 & 3 & 5 \end{array}$$

Wir stellen durch die bekannten Elementarumformungen die Zeilenstufenform her:

$$\begin{array}{l} 2 \cdot \boxed{2} - \boxed{1} =: \boxed{4} : \\ \boxed{3} - \boxed{1} =: \boxed{5} : \end{array} \begin{array}{ccc|cc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & & \\ \hline 2 & 0 & 1 & 2 & 7 \\ 0 & 4 & 1 & 4 & -5 \\ 0 & 1 & 1 & 1 & -2 \end{array}$$

$$\rightsquigarrow \begin{array}{l} \boxed{1} : \\ \boxed{4} : \\ 4 \cdot \boxed{5} - \boxed{4} =: \boxed{6} : \end{array} \begin{array}{ccc|cc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & & \\ \hline 2 & 0 & 1 & 2 & 7 \\ 0 & 4 & 1 & 4 & -5 \\ 0 & 0 & 3 & 0 & -3 \end{array}$$

Und jetzt lösen wir von unten nach oben:

$$\begin{array}{l} \boxed{6} \implies 3\vec{x}_3^\top = [0 \quad -3] \implies \vec{x}_3^\top = [0 \quad -1], \\ \boxed{4} \implies 4\vec{x}_2^\top + \vec{x}_3^\top = [4 \quad -5] \implies 4\vec{x}_2^\top = [4 \quad -5] - [0 \quad -1] = [4 \quad -4], \\ \implies \vec{x}_2^\top = [1 \quad -1], \\ \boxed{1} \implies 2\vec{x}_1^\top + \vec{x}_3^\top = [2 \quad 7] \implies 2\vec{x}_1^\top = [2 \quad 7] - [0 \quad -1] = [2 \quad 8], \\ \implies \vec{x}_1^\top = [1 \quad 4]. \end{array}$$

Somit erhalten wir insgesamt

$$X = \begin{bmatrix} 1 & 4 \\ 1 & -1 \\ 0 & -1 \end{bmatrix}.$$

**Definition 3.20 (Einheitsmatrix):**

Die quadratische Matrix

$$E_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = [\vec{e}_1 \quad \vec{e}_2 \quad \dots \quad \vec{e}_n] \in \mathbb{R}^{n \times n}$$

heißt *Einheitsmatrix*.

**Bedeutung:** für alle  $A \in \mathbb{R}^{m \times n}$  gilt  $AE_n = E_m A = A$ , also hat  $E_m$  bei der Matrixmultiplikation die Rolle der Eins.

**Ziel:** wir wollen jetzt durch eine Matrix „teilen“. Gesucht ist also eine Matrix  $A^{-1}$ , die zu einer gegebenen Matrix  $A$  gehört, sodass  $AA^{-1} = E_n$ .

**Definition 3.21 (Inverse Matrix):**

Sei eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  gegeben. Wenn es dann eine Matrix  $A^{-1} \in \mathbb{R}^{n \times n}$  gibt mit  $AA^{-1} = E_n$  und  $A^{-1}A = E_n$ , so heißt  $A^{-1}$  die *Inverse von A*. In diesem Fall heißt  $A$  *invertierbar* oder *regulär*.

Wenn es eine solche inverse Matrix  $A^{-1}$  nicht gibt, dann heißt  $A$  *singulär*.

**Bemerkung 3.22:** (a) Wenn eine Inverse existiert, dann ist sie eindeutig.

(b) Es gibt noch mehr singuläre Matrizen als nur die Nullmatrix, wie wir später behandeln werden.

**Rechenregeln 3.23:**

Seien  $A, B \in \mathbb{R}^{n \times n}$  invertierbar. Dann gilt

$$(A^{-1})^{-1} = A, \tag{a}$$

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (\text{auf Reihenfolge achten!}) \tag{b}$$

$$(A^T)^{-1} = (A^{-1})^T. \tag{c}$$

**Bemerkung 3.24 (LGS mit invertierbarer Matrix):**

Ist die Matrix  $A$  invertierbar, so ist das LGS  $A\vec{x} = \vec{b}$  für jedes  $\vec{b}$  eindeutig lösbar durch  $\vec{x} := A^{-1}\vec{b}$ , denn wir haben  $A(A^{-1}\vec{b}) = AA^{-1}\vec{b} = E_n\vec{b} = \vec{b}$ . Und falls es eine weitere Lösung geben sollte (wir können sie  $\vec{y}$  nennen), dann ist  $A\vec{y} = \vec{b}$ , und es ist  $A^{-1}\vec{b} = A^{-1}A\vec{y} = E_n\vec{y} = \vec{y}$ , also  $\vec{x} = \vec{y}$ .

**Verfahren 3.25 (Bestimmung der Inversen):**

Sei  $A \in \mathbb{R}^{n \times n}$  gegeben. Die inverse Matrix  $A^{-1} =: X$  löst  $AX = E_n$ , also können wir das Verfahren 3.18 anwenden, wenn wir die inverse Matrix  $A^{-1}$  suchen.

**Beispiel 3.26:** (a) Wir betrachten die Matrix

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}$$

und suchen die inverse Matrix als Lösung von  $AX = E_3$ :

$$\begin{array}{l} \boxed{1} : \\ \boxed{2} : \\ \boxed{3} : \end{array} \begin{array}{ccc|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & & & \\ \hline 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}$$

$$\rightsquigarrow \begin{array}{l} \boxed{1} : \\ \boxed{2} - \boxed{1} =: \boxed{4} : \\ \boxed{3} - 2 \cdot \boxed{1} =: \boxed{5} : \end{array} \begin{array}{ccc|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & & & \\ \hline 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -3 & -1 & -2 & 0 & 1 \end{array}$$

$$\begin{array}{l}
\rightsquigarrow \begin{array}{l} \boxed{1} + 2 \cdot \boxed{4} =: \boxed{6} \\ \boxed{5} - 3 \cdot \boxed{4} =: \boxed{7} \end{array} : \begin{array}{c|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & \\ \hline 1 & 0 & 3 & -1 \quad 2 \quad 0 \\ 0 & -1 & 1 & -1 \quad 1 \quad 0 \\ 0 & 0 & -4 & 1 \quad -3 \quad 1 \end{array} \\
\rightsquigarrow \begin{array}{l} \boxed{6} \\ \boxed{4} \\ -\frac{1}{4} \cdot \boxed{7} =: \boxed{8} \end{array} : \begin{array}{c|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & \\ \hline 1 & 0 & 3 & -1 \quad 2 \quad 0 \\ 0 & -1 & 1 & -1 \quad 1 \quad 0 \\ 0 & 0 & 1 & -\frac{1}{4} \quad \frac{3}{4} \quad -\frac{1}{4} \end{array} \\
\rightsquigarrow \begin{array}{l} \boxed{6} - 3 \cdot \boxed{8} =: \boxed{9} \\ \boxed{4} - \boxed{8} =: \boxed{10} \\ \boxed{8} \end{array} : \begin{array}{c|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & \\ \hline 1 & 0 & 0 & -\frac{1}{4} \quad -\frac{1}{4} \quad \frac{3}{4} \\ 0 & -1 & 0 & -\frac{3}{4} \quad \frac{1}{4} \quad \frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{4} \quad \frac{3}{4} \quad -\frac{1}{4} \end{array} \\
\rightsquigarrow \begin{array}{l} -1 \cdot \boxed{10} =: \boxed{9} \\ \boxed{11} \\ \boxed{8} \end{array} : \begin{array}{c|ccc} \vec{x}_1^\top & \vec{x}_2^\top & \vec{x}_3^\top & \\ \hline 1 & 0 & 0 & -\frac{1}{4} \quad -\frac{1}{4} \quad \frac{3}{4} \\ 0 & 1 & 0 & \frac{3}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{4} \quad \frac{3}{4} \quad -\frac{1}{4} \end{array}
\end{array}$$

Somit haben wir ermittelt:

$$A^{-1} = \frac{1}{4} \begin{bmatrix} -1 & -1 & 3 \\ 3 & -1 & -1 \\ -1 & 3 & -1 \end{bmatrix}.$$

(b) Ist  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ , so gilt folgendes:

$A$  ist invertierbar  $\iff ad - bc \neq 0$  und in diesem Fall ist

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Das kann man durch Nachrechnen überprüfen.

### Satz 3.27 (Invertierbarkeitskriterium):

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix. Dann gilt:

$A$  ist invertierbar  $\iff \ker(A) = \{0\} \iff \text{rang}(A) = n$ .

## 4 Lineare Abbildungen

### Definition 4.1 (Lineare Abbildungen):

Eine Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  heißt *lineare Abbildung*, falls für alle  $\vec{u}, \vec{v} \in \mathbb{R}^n$  und alle  $\lambda \in \mathbb{R}$  gilt:

$$f(\lambda \vec{u}) = \lambda f(\vec{u}) \quad \text{und} \quad f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}).$$

Wir erinnern daran, dass es definitionsgemäß in jedem Vektorraum zwei typische Rechenoperationen gibt: „Vektor plus Vektor ergibt Vektor“ und „Zahl mal Vektor ergibt Vektor“, siehe Definition 1.2. Damit landen wir bei folgender Interpretation: eine Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist genau dann linear, wenn sie zu den beiden Vektorraumoperationen kompatibel ist.

**Beispiele 4.2:** (a) Viele elementare Transformationen der Geometrie und Mechanik sind linear:

- Drehungen
- Spiegelungen
- Projektionen

- Streckungen
- Scherungen

(b) Ist  $\vec{a} \in \mathbb{R}^n$  fest, so ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle$  linear, denn:

- $f(\lambda \vec{x}) = \langle \vec{a}, \lambda \vec{x} \rangle = \lambda \langle \vec{a}, \vec{x} \rangle = \lambda f(\vec{x})$ ,
- $f(\vec{x} + \vec{y}) = \langle \vec{a}, \vec{x} + \vec{y} \rangle = \langle \vec{a}, \vec{x} \rangle + \langle \vec{a}, \vec{y} \rangle = f(\vec{x}) + f(\vec{y})$ .

(c) Ist  $A \in \mathbb{R}^{m \times n}$ , so ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $f(\vec{x}) = A\vec{x}$  linear, denn:

- $f(\lambda \vec{x}) = A(\lambda \vec{x}) = \lambda A\vec{x} = \lambda f(\vec{x})$ ,
- $f(\vec{x} + \vec{y}) = A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = f(\vec{x}) + f(\vec{y})$ .

Siehe dazu auch Regeln 3.11.

**Bemerkung 4.3:**

Für jede lineare Abbildung  $f$  gilt  $f(\vec{0}) = \vec{0}$ , denn

$$f(\vec{0}) = f(\vec{0} + \vec{0}) = f(\vec{0}) + f(\vec{0}) \implies \vec{0} = f(\vec{0}).$$

Demzufolge sind Verschiebungen keine linearen Abbildungen (abgesehen von der Nullverschiebung).

Als nächstes zeigen wir, dass *jede* lineare Abbildung durch eine Matrix dargestellt werden kann. Es ist also Beispiel 4.2 (c) der Normalfall.

**Bemerkung 4.4 (Abbildungsmatrix):**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung. Sei  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  eine Basis des  $\mathbb{R}^n$ , und sei  $C = \{\vec{c}_1, \dots, \vec{c}_m\}$  eine Basis des  $\mathbb{R}^m$ .

Sei  $\vec{x} \in \mathbb{R}^n$  ein gegebener Vektor. Wir merken an, dass Basen und Koordinaten Werkzeuge sind, die man benutzt, um Vektoren handhabbar zu machen; und häufig wird durch die Einführung von Basen und Koordinaten das Rechnen überhaupt erst durchführbar. Wir entwickeln also den Vektor  $\vec{x}$  bezüglich der Basis  $B$ . Das heißt: es gibt eindeutig bestimmte Zahlen  $\xi_1, \dots, \xi_n \in \mathbb{R}$  mit

$$\vec{x} = \sum_{j=1}^n \xi_j \vec{b}_j.$$

Diese Zahlen  $\xi_j$  hatten wir *Koordinaten von  $\vec{x}$  bezüglich  $B$*  genannt (siehe Satz 1.17 und Bild 4.8).

Dann haben wir

$$\begin{aligned} f(\vec{x}) &= f\left(\sum_{j=1}^n \xi_j \vec{b}_j\right) && \left| \begin{array}{l} f \text{ ist linear} \end{array} \right. \\ &= \sum_{j=1}^n f\left(\xi_j \vec{b}_j\right) && \left| \begin{array}{l} f \text{ ist linear} \end{array} \right. \\ &= \sum_{j=1}^n \xi_j f(\vec{b}_j). \end{aligned}$$

Das bedeutet **zum Einigen**: wenn man die  $n$  Vektoren  $f(\vec{b}_1), \dots, f(\vec{b}_n)$  kennt, dann kann man für jedes  $\vec{x} \in \mathbb{R}^n$  den zugehörigen Ergebnisvektor  $f(\vec{x})$  sofort ausrechnen. Die  $n$  Vektoren  $f(\vec{b}_1), \dots, f(\vec{b}_n)$  legen die Abbildung  $f$  eindeutig fest.

**zum Anderen**: da die Ergebnisvektoren  $f(\vec{b}_1), \dots, f(\vec{b}_n)$  im  $\mathbb{R}^m$  liegen, besitzen sie alle eine Koordinatendarstellung bezüglich der Basis  $C$ . Das bedeutet, dass es Zahlen  $\alpha_{ij} \in \mathbb{R}$  gibt mit

$$f(\vec{b}_j) = \sum_{i=1}^m \alpha_{ij} \vec{c}_i, \quad j = 1, \dots, n.$$

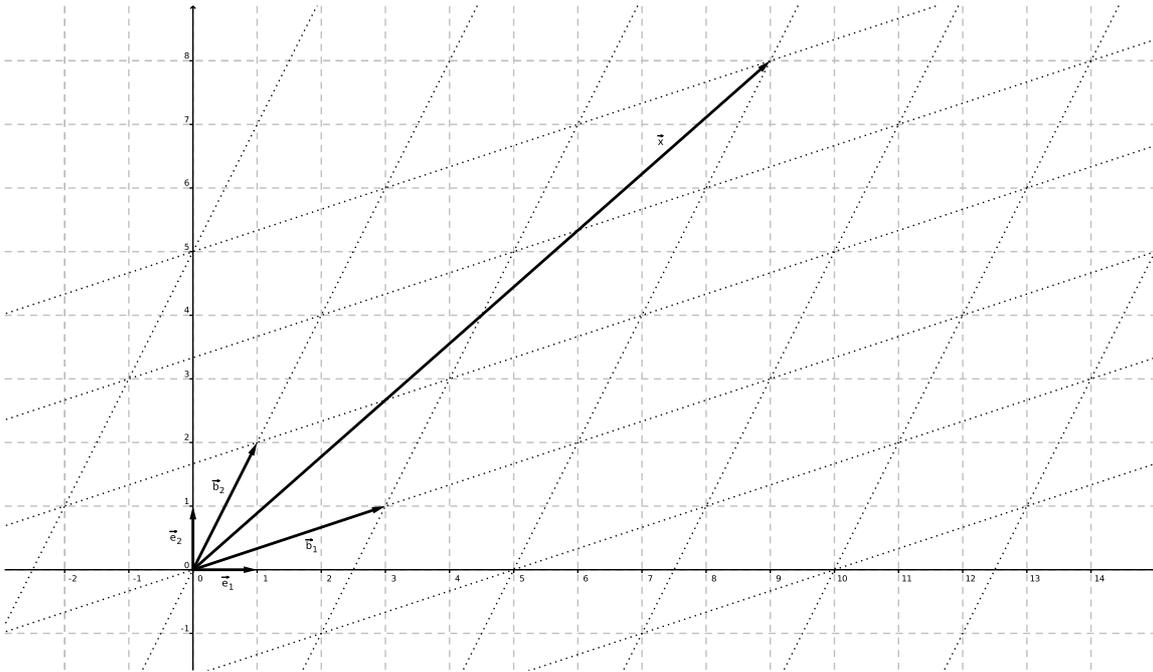


Abbildung 4.8: Wir haben die Standardbasis  $\{\vec{e}_1, \vec{e}_2\}$  mit  $\vec{e}_1 = [1, 0]^\top$  und  $\vec{e}_2 = [0, 1]^\top$ . Und wir haben eine weitere Basis  $B = \{\vec{b}_1, \vec{b}_2\}$  mit  $\vec{b}_1 = [3, 1]^\top$  und  $\vec{b}_2 = [1, 2]^\top$ . Das gestrichelte Raster gehört zur Standardbasis, und das gepunktete Raster gehört zur Basis  $B$ . Der Punkt  $\vec{x} = [9, 8]^\top$  hat in der Standardbasis die Koordinaten 9 und 8. Und gleichzeitig hat der Punkt  $\vec{x} = [9, 8]^\top$  in der Basis  $B$  die Koordinaten  $\xi_1 = 2$  und  $\xi_2 = 3$ . Dies kann man geometrisch erkennen, in dem man im gepunkteten Raster „zwei Schritte in Richtung  $\vec{b}_1$  geht und drei Schritte in Richtung  $\vec{b}_2$ “. Und man kann es rechnerisch sehen anhand der Identität  $[9, 8]^\top = [6, 2]^\top + [3, 6]^\top = 2 \cdot [3, 1]^\top + 3 \cdot [1, 2]^\top$ . Wir haben auch  $\begin{bmatrix} 9 \\ 8 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix}$  als Umrechnungsformel zwischen den Koordinaten  $[9, 8]^\top$  und  $[2, 3]^\top$ . In den Spalten der Basiswechsellmatrix stehen die Basisvektoren  $\vec{b}_1$  und  $\vec{b}_2$ .

Damit können wir unsere oben begonnene Rechnung fortsetzen:

$$f(\vec{x}) = \sum_{j=1}^n \xi_j \left( \sum_{i=1}^m \alpha_{ij} \vec{c}_i \right) = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \xi_j \vec{c}_i.$$

Jetzt führen wir zwei Schreibweisen ein. Die Zahlen  $\xi_j$  türmen wir zu einer Spalte  $\vec{\xi}$  auf:

$$\vec{\xi} := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Dies ist der Koordinatenvektor von  $\vec{x}$  bezüglich der Basis  $B$ .

Und die Zahlen  $\alpha_{ij}$  stapeln wir matrixförmig:

$$A := \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix}.$$

Damit kehren wir zu unserer Rechnung zurück:

$$f(\vec{x}) = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \xi_j \vec{c}_i = \sum_{i=1}^m \left( \underbrace{\sum_{j=1}^n \alpha_{ij} \xi_j}_{[A \cdot \vec{\xi}]_i} \right) \vec{c}_i.$$

Dies bedeutet:  $A \cdot \vec{\xi} \in \mathbb{R}^m$  ist der Koordinatenvektor von  $f(\vec{x})$  bezüglich der Basis  $C$ . Die Matrix  $A$  beschreibt also die Wirkung der Abbildung  $f$  und heißt *Abbildungsmatrix von  $f$  bezüglich der Basen  $B$  und  $C$* .

Die symbolische Schreibweise ist  $A = M_C^B(f)$ .

Wir haben somit gezeigt: jede lineare Abbildung von  $\mathbb{R}^n$  in den  $\mathbb{R}^m$  wird von einer Matrix erzeugt. Jetzt schauen wir uns an, wie man diese Matrix ermittelt.

**Bemerkung 4.5 (Bestimmung der Abbildungsmatrix):**

Wir haben eine lineare Abbildung  $f$ , die von einem Definitionsgebietsraum  $\mathbb{R}^n$  in einen Zielraum  $\mathbb{R}^m$  abbildet. Wir haben weiterhin eine Basis  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  für den  $\mathbb{R}^n$  und eine Basis  $C = \{\vec{c}_1, \dots, \vec{c}_m\}$  für den  $\mathbb{R}^m$ .

Wir wählen einen speziellen Vektor  $\vec{x}$ , nämlich  $\vec{x} = \vec{b}_j$ , und wir entwickeln ihn bezüglich der Basis  $B$ :

$$\vec{x} = 0 \cdot \vec{b}_1 + \dots + 0 \cdot \vec{b}_{j-1} + 1 \cdot \vec{b}_j + 0 \cdot \vec{b}_{j+1} + \dots + 0 \cdot \vec{b}_n,$$

also ergibt sich der Koordinatenvektor  $\vec{\xi}$  zum Vektor  $\vec{x}$  wie folgt:

$$\vec{\xi} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix},$$

mit einer 1 an der  $j$ -ten Position.

Wir machen uns klar, dass das Matrix-Vektor-Produkt  $A \cdot \vec{\xi}$  für diesen Vektor  $\xi$  genau die  $j$ -te Spalte  $\vec{a}^j$  von  $A$  zum Ergebnis hat. Damit folgt dann:

$$f(\vec{b}_j) = \sum_{i=1}^m [A \cdot \vec{\xi}]_i \vec{c}_i = \sum_{i=1}^m [\vec{a}^j]_i \vec{c}_i.$$

Das heißt: die  $j$ -te Spalte von  $M_C^B(f)$  hat als Einträge die Koordinaten von  $f(\vec{b}_j)$  bezüglich der Basis  $C$ .

Die Merkregel lautet:

*In den Spalten der Abbildungsmatrix stehen die Koordinaten der Bilder der Basisvektoren.*

**Zusammenfassung 4.6:**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear,  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  eine Basis des  $\mathbb{R}^n$ ,  $C = \{\vec{c}_1, \dots, \vec{c}_m\}$  eine Basis des  $\mathbb{R}^m$ . (Sehr häufig ist  $n = m$ , und man hat auch  $B = C$  gleich der Standardbasis.)

Dann gilt:

- Es ist  $M_C^B(f)$  diejenige Matrix, in deren Spalten die Koordinaten der Ergebnisvektoren  $f(\vec{b}_1), \dots, f(\vec{b}_n)$  bezüglich der Basis  $C$  stehen.
- Wenn man  $M_C^B(f)$  berechnet hat, dann erhält man für beliebiges  $\vec{x} \in \mathbb{R}^n$  den Ergebnisvektor  $f(\vec{x})$  wie folgt:
  - Man bestimme den Koordinatenvektor  $\vec{\xi}$  von  $\vec{x}$  bezüglich der Basis  $B$  (falls  $B$  die Standardbasis des  $\mathbb{R}^n$  ist, dann gilt  $\vec{\xi} = \vec{x}$ ).
  - Man berechne das Matrix-Vektor-Produkt  $A \cdot \vec{\xi} \in \mathbb{R}^m$ .
  - Dieses Produkt ist dann der Koordinatenvektor von  $f(\vec{x})$  bezüglich der Basis  $C$ , d.h.  $f(\vec{x}) = \sum_{i=1}^m [A\vec{\xi}]_i \vec{c}_i$  (falls  $C$  die Standardbasis des  $\mathbb{R}^m$  ist, gilt  $f(\vec{x}) = A \cdot \vec{\xi}$ ).

Wichtig ist dabei: der ganze Kalkül hängt von der Wahl der Basen  $B$  und  $C$  ab. Wenn man eine der Basen ändert, dann ändert sich meist auch die Abbildungsmatrix !

**Beispiele 4.7:** (a) Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Spiegelung an der  $y$ -Achse, siehe Bild 4.9.

Wir wählen  $B = C = \{\vec{e}_1, \vec{e}_2\}$ , also die Standardbasis.

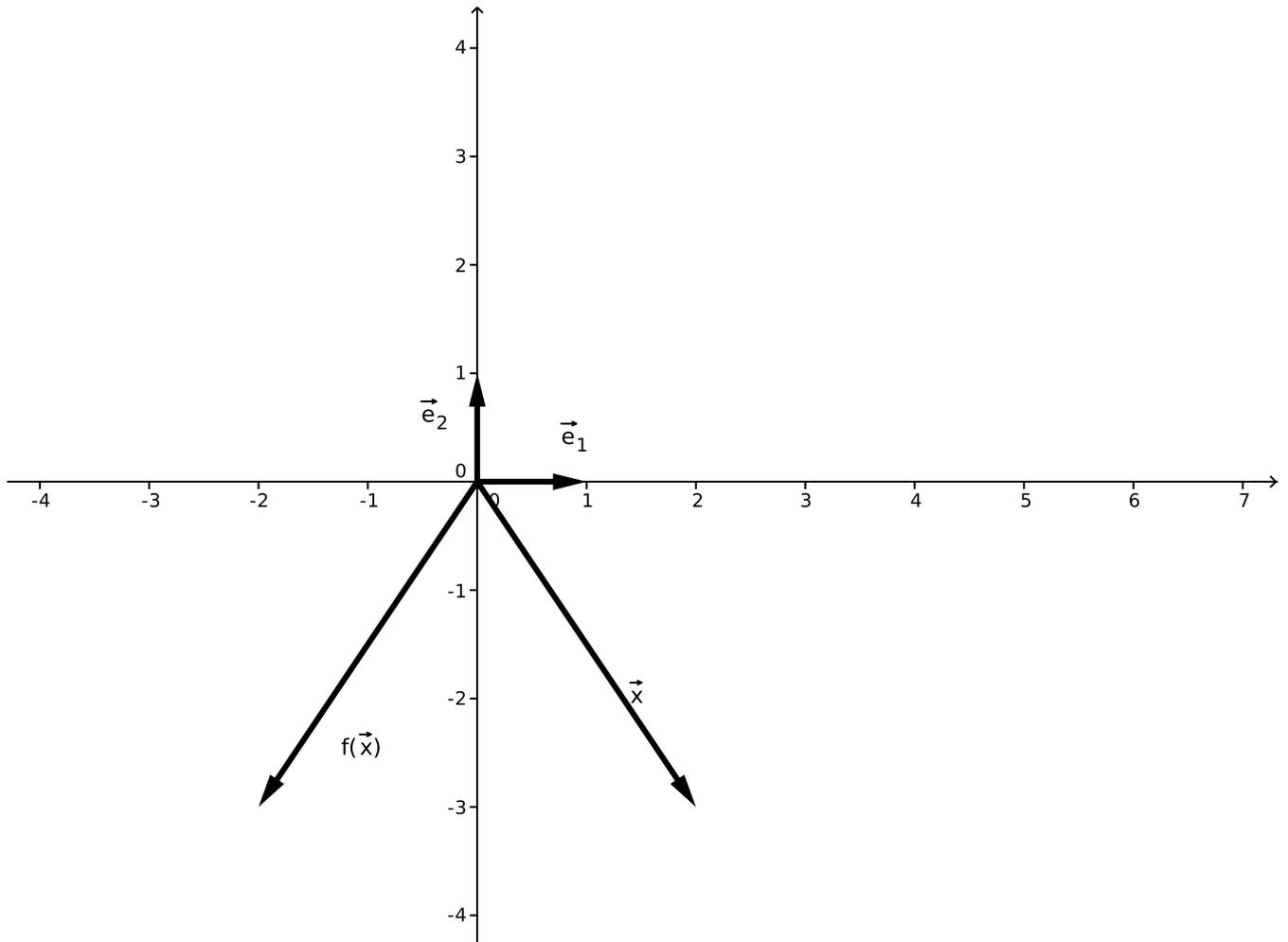


Abbildung 4.9: Die Spiegelung an der  $y$ -Achse.

Geometrisch ist klar, dass  $f(\vec{e}_1) = -\vec{e}_1$  und  $f(\vec{e}_2) = \vec{e}_2$  ist.

Der Koordinatenvektor von  $f(\vec{e}_1) = -\vec{e}_1$  bezüglich der Basis  $\{\vec{e}_1, \vec{e}_2\}$  ist  $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ .

Der Koordinatenvektor von  $f(\vec{e}_2) = \vec{e}_2$  bezüglich der Basis  $\{\vec{e}_1, \vec{e}_2\}$  ist  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ .

Also erhalten wir  $M_B^B(f)$  durch Nebeneinanderstellen dieser Koordinatenvektoren:

$$M_B^B(f) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Für  $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$  ist also (hierfür beachte man  $\vec{\xi} = \vec{x}$ )

$$f(\vec{x}) = A\vec{x} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}.$$

- (b) Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  eine Drehung um den Winkel  $\gamma$  (im Gegenuhrzeigersinn) um den Ursprung der Ebene, siehe Bild 4.10. Wir wählen  $B = C = \{\vec{e}_1, \vec{e}_2\}$ , also die Standardbasis. Dann machen wir uns geometrisch klar, dass die Bilder der Standardbasisvektoren die folgenden sind:

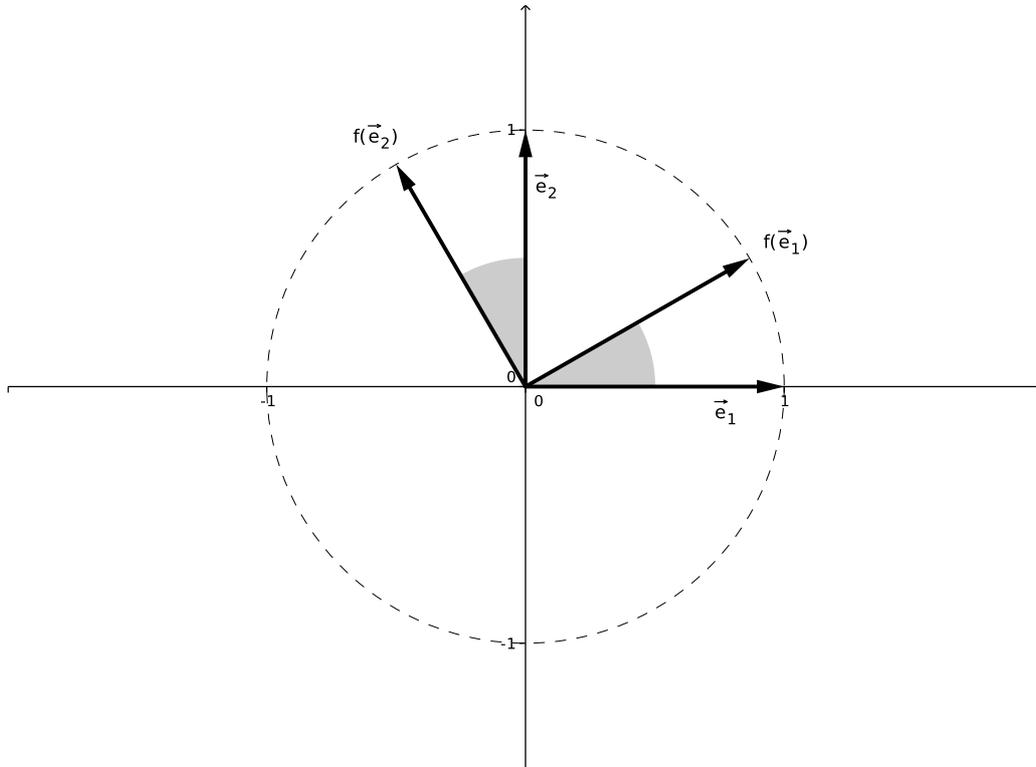


Abbildung 4.10: Die Drehung um einen Winkel  $\gamma$ .

$$f(\vec{e}_1) = \begin{bmatrix} \cos(\gamma) \\ \sin(\gamma) \end{bmatrix}, \quad f(\vec{e}_2) = \begin{bmatrix} -\sin(\gamma) \\ \cos(\gamma) \end{bmatrix}.$$

Also ist dann

$$M_B^B(f) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{bmatrix}.$$

Wenn wir in diesem Beispiel anstatt der Standardbasis übrigens irgendeine andere Orthonormalbasis für  $B = C$  genommen hätten, dann hätten wir die gleiche Matrix  $M_B^B(f)$  erhalten.

- (c) Wir betrachten  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , und  $f$  sei geometrisch definiert durch die Projektion auf den Untervektorraum

$$U := \text{Lin} \left( \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right)$$

in Richtung  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ . Wir stellen uns dies so vor, dass „von rechts oben“ Licht diagonal „nach links unten“ strahlt, und entlang  $U$  ist eine Leinwand angebracht, auf der der Schatten auftrifft. Dann ist  $f$  also die Schattenwurfabbildung, siehe Bild 4.11.

Wir suchen die Abbildungsmatrix zu  $f$ ; und dafür bietet es sich an, die Basis  $B$  dem Problem angepasst zu wählen. Wir nehmen

$$B = \left\{ \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \{\vec{b}_1, \vec{b}_2\}.$$

Dann haben wir naheliegenderweise  $f(\vec{b}_1) = \vec{b}_1$  und  $f(\vec{b}_2) = \vec{0}$ .

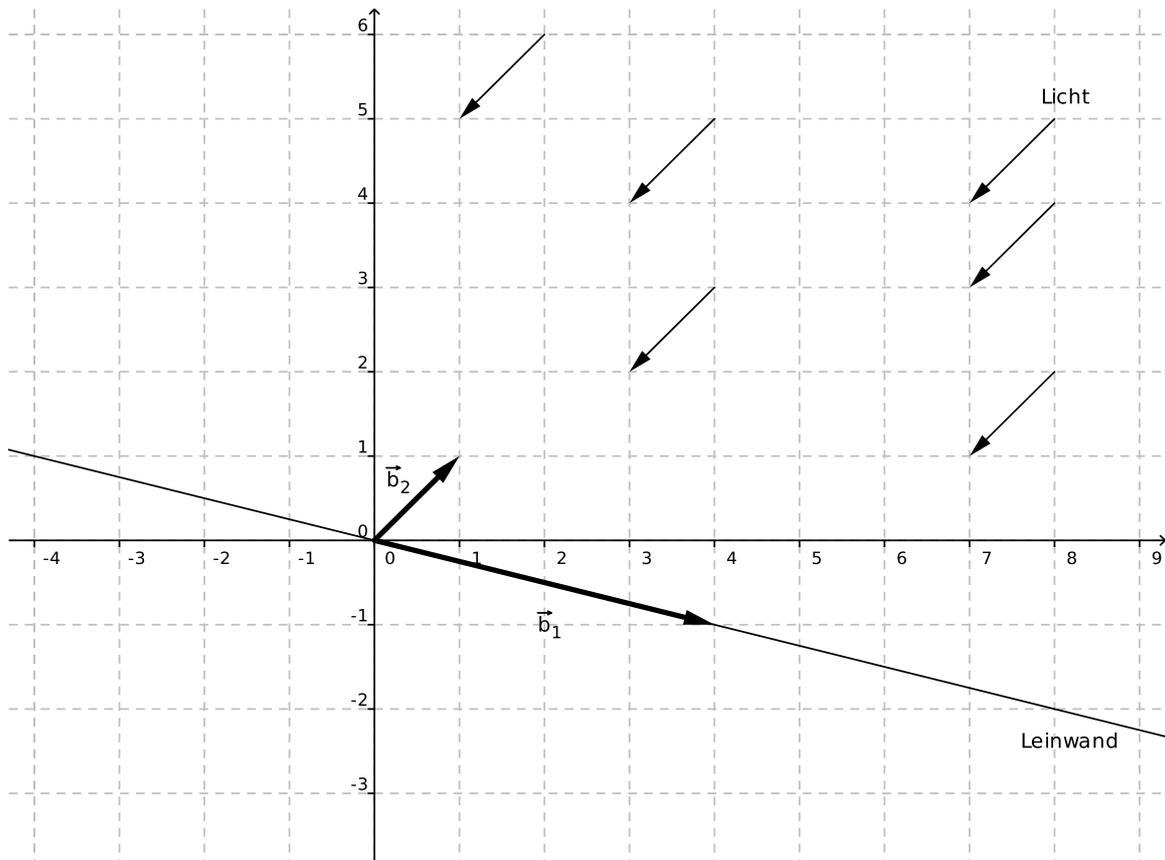


Abbildung 4.11: Licht fällt auf eine Leinwand.

Der Koordinatenvektor von  $f(\vec{b}_1) = \vec{b}_1$  bezüglich der Basis  $\{\vec{b}_1, \vec{b}_2\}$  ist  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ .

Der Koordinatenvektor von  $f(\vec{b}_2) = \vec{0}$  bezüglich der Basis  $\{\vec{b}_1, \vec{b}_2\}$  ist  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ .

Also erhalten wir  $M_B^B(f)$  durch Nebeneinanderstellen dieser Koordinatenvektoren:

$$M_B^B(f) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Beispielhaft betrachten wir den Vektor  $\vec{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ . Er wird von der Basis  $B$  erzeugt gemäß  $\vec{x} = \vec{b}_1 + \vec{b}_2$ , also lautet seine zugehörige Spalte  $\vec{\xi} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ . Dann folgt

$$M_B^B(f) \cdot \vec{\xi} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

und dies ist der Koordinatenvektor von  $f(\vec{x})$  bezüglich der Basis  $B$ , das heißt

$$f(\vec{x}) = 1 \cdot \vec{b}_1 + 0 \cdot \vec{b}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$$

Wir haben die Abbildungsmatrix von  $f$  bezüglich der Basis  $B$  ermittelt. Es wäre auch recht interessant zu erfahren, wie die Abbildungsmatrix von  $f$  bezüglich der Standardbasis  $E$  aussieht.

**Bemerkung 4.8 (Basiswechsel):**

Sei  $E = \{\vec{e}_1, \dots, \vec{e}_n\}$  die Standardbasis des  $\mathbb{R}^n$ , und sei  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  eine weitere Basis des  $\mathbb{R}^n$ . Gegeben sei eine lineare Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ .

**Gesucht** ist ein Zusammenhang zwischen  $M_B^B(f)$  und  $M_E^E(f)$ .

Wir formulieren die Frage etwas ausführlicher. Wir betrachten einen Vektor  $\vec{x} \in \mathbb{R}^n$  aus dem Originalraum

und seinen Bildvektor  $\vec{y} := f(\vec{x})$ . Beide Vektoren  $\vec{x}$  und  $\vec{y}$  entwickeln wir bezüglich der Standardbasis  $E$ :

$$\vec{x} = \sum_{j=1}^n x_j \vec{e}_j, \quad \vec{y} = \sum_{j=1}^n y_j \vec{e}_j,$$

und wir entwickeln sie bezüglich der Basis  $B$  (siehe auch den Begleittext zu Abbildung 4.8):

$$\vec{x} = \sum_{j=1}^n \xi_j \vec{b}_j, \quad \vec{y} = \sum_{j=1}^n \eta_j \vec{b}_j.$$

Diese Zeile definiert die Zahlen  $\xi_j$  und  $\eta_j$ , und wir stapeln diese Koordinaten zu Koordinatenvektoren  $\vec{\xi}$  und  $\vec{\eta}$  auf. Dann haben wir die Umrechnung zwischen den Koordinatenspalten gemäß

$$\vec{y} = M_E^E(f)\vec{x}, \quad \vec{\eta} = M_B^B(f)\vec{\xi},$$

denn so waren die Matrizen  $M_E^E(f)$  und  $M_B^B(f)$  definiert. Wie rechnet man zwischen diesen beiden Matrizen um?

**Wichtige Beobachtung:** Wenn wir die Spaltenvektoren  $\vec{b}_1, \dots, \vec{b}_n$  zu einer quadratischen Matrix  $S$  nebeneinanderstellen, also

$$S := \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \in \mathbb{R}^n$$

bilden, so gilt:

- $\text{rang}(S) = n$ , also ist  $S$  invertierbar;
- für jedes  $j \in \{1, \dots, n\}$  ist

$$S\vec{e}_j = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{b}_j,$$

also rechnet  $S$  von der Basis  $E$  in die Basis  $B$  um;

- Der Koordinatenvektor von  $\vec{x}$  bezüglich der Basis  $B$  war  $\vec{\xi}$  getauft worden, und es gilt

$$S\vec{\xi} = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = \sum_{j=1}^n \xi_j \vec{b}_j = \vec{x}.$$

Analog ist auch

$$S\vec{\eta} = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \sum_{j=1}^n \eta_j \vec{b}_j = \vec{y},$$

also rechnet  $S$  von den Koordinaten bezüglich  $B$  in die Koordinaten bezüglich  $E$  um;

- weil  $S^{-1}$  existiert, haben wir dann auch für jedes  $j \in \{1, \dots, n\}$ , dass

$$S^{-1}\vec{b}_j = \vec{e}_j, \quad \text{und auch} \quad S^{-1}\vec{x} = \vec{\xi}, \quad S^{-1}\vec{y} = \vec{\eta},$$

also rechnet  $S^{-1}$  von den Koordinaten bezüglich  $E$  in die Koordinaten bezüglich  $B$  um.

Jetzt bauen wir alles zusammen:

$$\text{einerseits ist } \vec{y} = M_E^E(f)\vec{x},$$

$$\text{andererseits ist } \vec{y} = S\vec{\eta} = S \cdot M_B^B(f)\vec{\xi} = S \cdot M_B^B(f) \cdot S^{-1}\vec{x},$$

und der Vergleich beider Zeilen bringt uns zur zentralen **Umrechnungsformel**:

$$M_E^E(f) = S \cdot M_B^B(f) \cdot S^{-1} \quad \text{mit} \quad S = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix}.$$

Wir haben auch folgende schematische Darstellung der Umrechnungsformel, die bildlich zeigen soll, dass es zwei Pfade gibt, um vom Koordinatenvektor  $\vec{x}$  zum Koordinatenvektor  $\vec{y}$  zu gelangen:

$$\begin{array}{ccc} \text{Vektoren:} & \left\{ \begin{array}{ccc} \text{Originalpunkt } \vec{x} & \xrightarrow{\text{Abbildung } f} & \text{Bildpunkt } \vec{y} \end{array} \right. \\ \\ \text{Koord. bzgl. } E : & \left\{ \begin{array}{ccc} \text{Koordinatenvektor } \vec{x} & \xrightarrow{\text{Matrix } M_E^E(f)} & \text{Koordinatenvektor } \vec{y} \\ & S^{-1} \downarrow & \uparrow S \end{array} \right. \\ \\ \text{Koord. bzgl. } B : & \left\{ \begin{array}{ccc} \text{Koordinatenvektor } \vec{\xi} & \xrightarrow{\text{Matrix } M_B^B(f)} & \text{Koordinatenvektor } \vec{\eta} \end{array} \right. \end{array}$$

Entsprechend gilt auch

$$M_B^B(f) = S^{-1} \cdot M_E^E(f) \cdot S.$$

Die Matrizen  $S$  und  $S^{-1}$  nennen wir *Basiswechselmatrizen*, siehe auch Abbildung 4.8 und den dazugehörigen Begleittext.

**Beispiele 4.9:** (a) Wir kehren zurück zu Beispiel 4.7 (c). Dort hatten wir

$$B = \left\{ \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}, \quad M_B^B(f) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

und somit entsteht dann (siehe auch Beispiel 3.26 (b))

$$\begin{aligned} M_E^E(f) &= S M_B^B(f) S^{-1} = \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 4 & 0 \\ -1 & 0 \end{bmatrix} \cdot \frac{1}{5} \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & -4 \\ -1 & 1 \end{bmatrix}. \end{aligned}$$

(b) Sei  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  die Spiegelung an der Ebene

$$\text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \right).$$

Wir suchen die Abbildungsmatrix von  $f$  in der Standardbasis. Wir betrachten zunächst eine geometrisch motivierte Hilfsbasis

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\} =: \{ \vec{b}_1, \vec{b}_2, \vec{b}_3 \}.$$

Dabei ist  $\vec{b}_3 := \vec{b}_1 \times \vec{b}_2$ , also steht  $\vec{b}_3$  senkrecht auf der Spiegelungsebene. Weil  $f$  eine Spiegelung sein soll, haben wir dann

$$f(\vec{b}_1) = \vec{b}_1, \quad f(\vec{b}_2) = \vec{b}_2, \quad f(\vec{b}_3) = -\vec{b}_3.$$

Die Merkregel zur Bestimmung der Abbildungsmatrix aus Bemerkung 4.5 bringt uns dann

$$M_B^B(f) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Solche übersichtlichen Matrizen nennt man *Diagonalmatrizen*. Wir wechseln jetzt von der Basis  $B$

zur Standardbasis  $E$ : Zunächst ist

$$\begin{aligned}
 S &= \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}, \quad \text{also ergibt sich dann} \\
 M_E^E(f) &= SM_B^B(f)S^{-1} \\
 &= \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \cdot S^{-1} \quad \left| \text{bestimme } S^{-1} \text{ mit Gauß-Algorithmus} \right. \\
 &= \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & -2 \\ 1 & 1 & -1 \end{bmatrix} \cdot \frac{1}{6} \begin{bmatrix} -1 & -4 & 7 \\ 2 & 2 & -2 \\ -1 & 2 & 1 \end{bmatrix} \\
 &= \frac{1}{6} \begin{bmatrix} 4 & 4 & 2 \\ 4 & -2 & -4 \\ 2 & -4 & 4 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix}.
 \end{aligned}$$

**Bemerkung 4.10 (Verknüpfung von linearen Abbildungen):**

Wenn  $f, g: \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $h: \mathbb{R}^m \rightarrow \mathbb{R}^p$  lineare Abbildungen sind, und  $B$  ist eine Basis von  $\mathbb{R}^n$ ,  $C$  eine Basis von  $\mathbb{R}^m$ ,  $D$  eine Basis von  $\mathbb{R}^p$ , dann gilt:

- (a)  $f + g: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist auch linear und es ist  $M_C^B(f + g) = M_C^B(f) + M_C^B(g)$ ;
- (b)  $\lambda f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist auch linear und es ist  $M_C^B(\lambda f) = \lambda M_C^B(f)$ ;
- (c)  $h \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^p$  ist auch linear und es ist  $M_D^B(h \circ f) = M_D^C(h) \cdot M_C^B(f)$  (Reihenfolge beachten!).

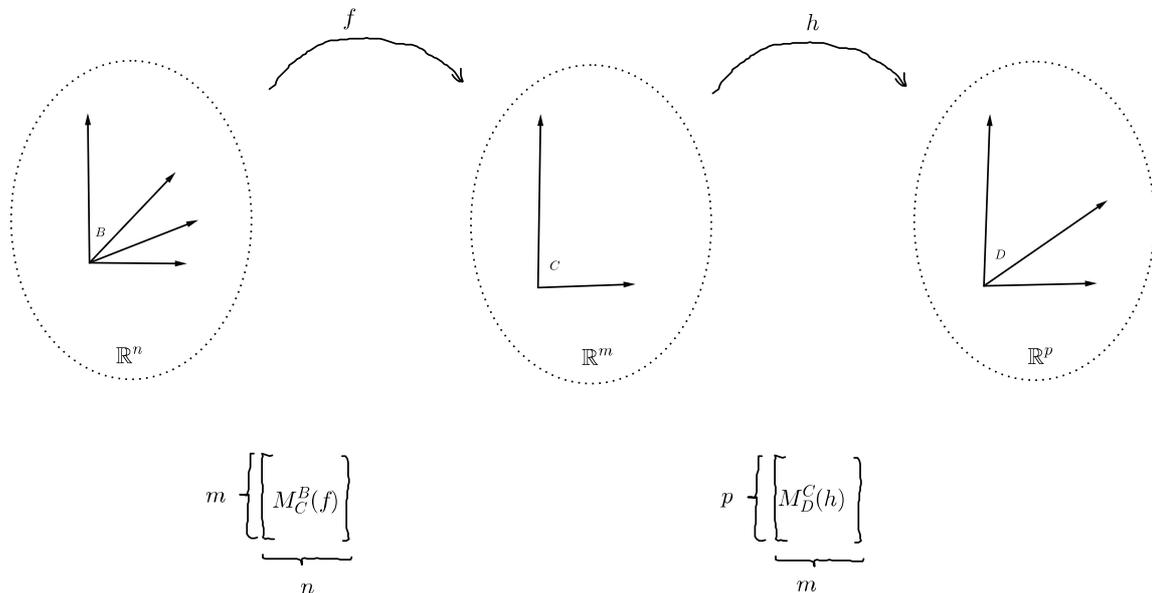


Abbildung 4.12: Zu Teil (c) von Bemerkung 4.10. Die Abbildung  $h \circ f$  bildet von  $\mathbb{R}^n$  via  $\mathbb{R}^m$  nach  $\mathbb{R}^p$  ab. Unten stehen schematisch die darstellenden Matrizen und deren Abmessungen. Die verknüpfte Abbildung  $h \circ f$  besitzt (in Bezug auf die eingezeichneten Basen  $B$  und  $D$ ) die Abbildungsmatrix  $M_D^C(h) \cdot M_C^B(f)$ . Beachte, dass die Matrizen  $M_C^B(f)$  und  $M_D^C(h)$  im Vergleich zur Graphik ihre Plätze getauscht haben.

**Satz 4.11:**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear und sei  $M_C^B(f) \in \mathbb{R}^{m \times n}$  die Abbildungsmatrix von  $f$  bezüglich der Basen  $B, C$ . Dann gilt:

- (a)  $f$  ist surjektiv  $\iff \text{rang}(A) = m$ ;
- (b)  $f$  ist injektiv  $\iff \ker(A) = \{\vec{0}\} \iff \text{rang}(A) = n$ ,  
(somit kann  $f$  bijektiv höchstens dann sein, wenn  $m = n$  sein sollte); und dann gilt:
- (c)  $f$  ist bijektiv  $\iff M_C^B(f)$  ist invertierbar. In diesem Fall ist  $M_B^C(f^{-1}) = (M_C^B(f))^{-1}$ , also ist dann die Abbildungsmatrix der Umkehrfunktion gleich der Inversen der Abbildungsmatrix.

Besonders wertvolle Basen sind Orthonormalbasen, weshalb wir die Frage stellen wollen, wie Basiswechselformen aussehen, wenn die beteiligten beiden Basen Orthonormalbasen sind.

**Definition 4.12 (Orthogonale Matrizen):**

Eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *orthogonal*, falls die Spalten von  $A$  eine Orthonormalbasis des  $\mathbb{R}^n$  bilden.

**Bemerkung 4.13 (Bedeutung von orthogonalen Matrizen):** (a) die Basiswechselformen zwischen zwei ONBen sind orthogonale Matrizen;

- (b) betrachtet man Drehungen und Spiegelungen im  $\mathbb{R}^n$ , so sind deren Abbildungsmatrizen bezüglich Orthonormalbasen stets orthogonale Matrizen. (Man probiere es aus mit Beispiel 4.7(a,b) !)

**Satz 4.14 (Kriterien für orthogonale Matrizen):**

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix. Die folgenden Bedingungen sind dann äquivalent:

- (a)  $A$  ist orthogonal;
- (b)  $A^\top$  ist orthogonal;
- (c)  $A^\top A = E_n$ , d.h.  $A$  ist invertierbar und  $A^{-1} = A^\top$ ;
- (d) die Zeilenvektoren von  $A$  bilden eine ONB des  $\mathbb{R}^n$ .

**Beispiel 4.15:**

Wir suchen die Abbildungsmatrix zur Abbildung

$$f(\vec{x}) = A\vec{x} = \frac{1}{9} \begin{bmatrix} 2 & 2 & 10 \\ 2 & 11 & -8 \\ 10 & -8 & 5 \end{bmatrix} \vec{x}$$

bezüglich der Basis

$$B = \left\{ \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

(Man rechnet nach, dass  $B$  tatsächlich eine ONB des  $\mathbb{R}^3$  ist.)

Es ist  $M_E^E(f) = A$ , also haben wir  $M_B^B(f) = S^{-1}AS$  zu bestimmen mit

$$S = \frac{1}{3} \begin{bmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{bmatrix}.$$

Es ist  $S$  orthogonal, weil  $B$  eine ONB ist. Demnach ist  $S^{-1} = S^\top$ , was unsere Rechnung deutlich verein-

facht (da wir den Gauß-Algorithmus zur Bestimmung von  $S^{-1}$  jetzt nicht benötigen):

$$\begin{aligned} M_B^B(f) &= S^T A S = \frac{1}{3} \begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ -2 & 1 & 2 \end{bmatrix} \cdot \frac{1}{9} \begin{bmatrix} 2 & 2 & 10 \\ 2 & 11 & -8 \\ 10 & -8 & 5 \end{bmatrix} \cdot \frac{1}{3} \begin{bmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{bmatrix} \\ &= \dots \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \end{aligned}$$

Die Abbildung  $f$  bekommt also eine besonders schöne darstellende Matrix, wenn wir die Basis  $B$  so wählen, wie hier in diesem Beispiel vorgeschlagen. Die tiefere geometrische Bedeutung darin werden wir später erkennen, wenn wir Eigenwerte und Eigenvektoren behandeln werden.

## 5 Determinanten

Für quadratische Matrizen werden wir jetzt einen mathematischen Begriff definieren, den wir *Determinante* nennen. Die Definition verläuft dabei rekursiv über die Abmessungen der Matrix.

### Definition 5.1 (Determinante):

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix.

- (a) Wir bezeichnen mit  $A_{ij} \in \mathbb{R}^{(n-1) \times (n-1)}$  diejenige Matrix, die entsteht, wenn man aus der Matrix  $A$  die  $i$ -te Zeile und die  $j$ -te Spalte herausstreicht.
- (b) Wenn  $n = 1$  ist, so definieren wir  $\det(A) = \det([a]) := a$ .
- (c) Wenn  $n \geq 2$  ist, so definieren wir

$$\det(A) := \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}).$$

Diese Formelzeile nennen wir *Entwicklung nach der ersten Zeile*.

Es gibt eine alternative (und kürzere) Schreibweise:

$$\text{statt } \det \left( \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \right) \quad \text{auch} \quad \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

**Bemerkung 5.2 (Spezialfälle):** (a) Für  $n = 2$  haben wir

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = (-1)^{1+1} a \cdot \det([d]) + (-1)^{1+2} b \cdot \det([c]) = ad - bc.$$

- (b) Für  $n = 3$  haben wir

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \cdot \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \cdot \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \cdot \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - fh) - b(di - fg) + c(dh - eg). \end{aligned}$$

Für größere  $n$  wird es sehr unübersichtlich !

(c) Im Falle einer unteren Dreiecksmatrix können wir so rechnen:

$$\begin{aligned}
 & \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} & \left| \begin{array}{l} \text{entwickle nach erster Zeile} \end{array} \right. \\
 & = a_{11} \cdot \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \underbrace{0 + 0 + \dots + 0}_{n-1 \text{ Summanden}} \\
 & = a_{11} \cdot \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} & \left| \begin{array}{l} \text{entwickle nach erster Zeile} \end{array} \right. \\
 & = a_{11} \cdot a_{22} \cdot \begin{vmatrix} a_{33} & \dots & 0 \\ \vdots & \ddots & \vdots \\ a_{n3} & \dots & a_{nn} \end{vmatrix} + \underbrace{a_{11} \cdot 0 + a_{11} \cdot 0 + \dots + a_{11} \cdot 0}_{n-2 \text{ Summanden}} \\
 & = a_{11} \cdot a_{22} \cdot \begin{vmatrix} a_{33} & \dots & 0 \\ \vdots & \ddots & \vdots \\ a_{n3} & \dots & a_{nn} \end{vmatrix} & \left| \begin{array}{l} \text{wiederhole diese Arbeitstechnik} \end{array} \right. \\
 & = \dots = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.
 \end{aligned}$$

**Bemerkung 5.3 (Anwendung in der Geometrie):** (a) Für  $n = 2$  ist der Betrag von  $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$  gleich dem Flächeninhalt des von den Vektoren  $\begin{bmatrix} a \\ c \end{bmatrix}$  und  $\begin{bmatrix} b \\ d \end{bmatrix}$  aufgespannten Parallelogramms, siehe Bild 5.13.

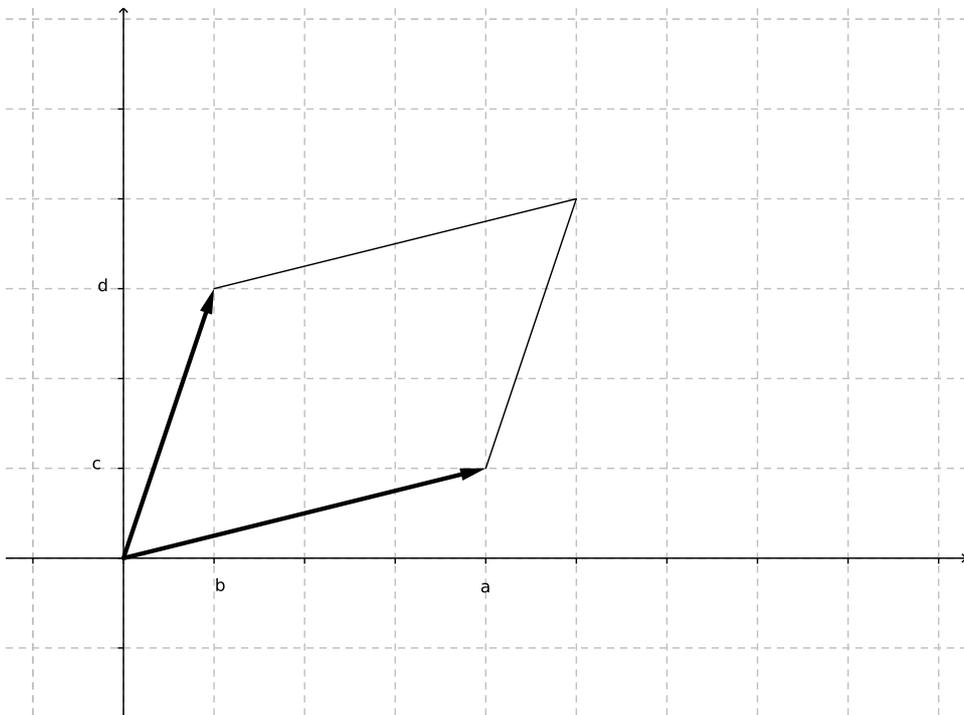


Abbildung 5.13: Die beiden Vektoren  $\begin{bmatrix} a \\ c \end{bmatrix}$  und  $\begin{bmatrix} b \\ d \end{bmatrix}$  leben in der Ebene und spannen dort ein Parallelogramm auf. Dessen Flächeninhalt ist der Betrag der Determinante  $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ . Wenn diese Determinante Null ist, dann ist das Parallelogramm „platt“, und die beiden Vektoren sind linear abhängig.

(b) Für  $\mathbf{n} = \mathbf{3}$  gilt: Der Betrag der Determinante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

ist das Volumen des *Parallelepipeds* (auch *Spat* genannt), das von den Vektoren  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}$  aufgespannt wird, siehe Bild 5.14. (Das Vorzeichen der Determinante gibt an, ob  $(\vec{a}, \vec{b}, \vec{c})$  ein Rechtssystem (positiv) oder ein Linkssystem (negativ) ist.)

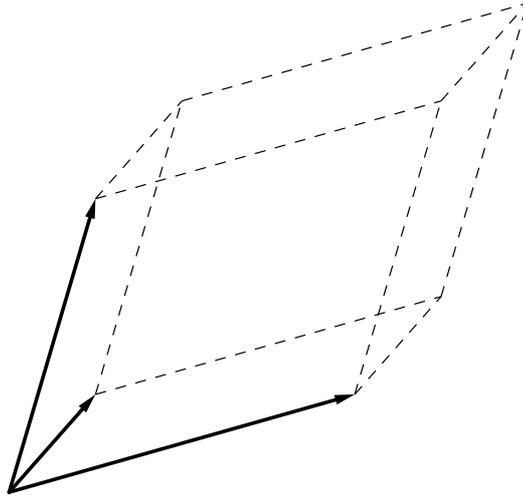


Abbildung 5.14: Die drei Vektoren leben im dreidimensionalen Raum und spannen dort einen Spat auf. Dessen Volumen ist der Betrag der im Text angegebenen Determinante. Wenn diese Determinante Null ist, dann ist der Spat „platt“, die drei Vektoren sind in einer Ebene, sind also linear abhängig.

(c) Ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear, so ist  $|\det(M_E^E(f))|$  der Faktor, um den  $f$  das Volumen verändert (z.B. gehören zu Drehungen immer Abbildungsmatrizen mit Determinante gleich Eins).

#### Rechenregeln 5.4:

Seien  $A, B \in \mathbb{R}^{n \times n}$  quadratische Matrizen. Dann gilt:

- (a)  $\det(A^\top) = \det(A)$ ;
- (b)  $\det(AB) = \det(A) \cdot \det(B)$ ;
- (c)  $\det(\lambda A) = \lambda^n \cdot \det(A)$  (*Exponent von  $\lambda$  beachten !*);
- (d)  $A$  ist invertierbar  $\iff \det(A) \neq 0$ . In diesem Fall gilt (auch weil  $E_n$  eine Dreiecksmatrix ist)

$$E_n = AA^{-1} \implies 1 = \det(E_n) = \det(AA^{-1}) = \det(A) \cdot \det(A^{-1}),$$

also auch

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

**Bemerkung 5.5:** (a) Die Rechenregel 5.4 (d) ist ein sehr wertvolles Invertierbarkeitskriterium.

(b) Man mache sich klar an einfachen Beispielen, dass im Allgemeinen

$$\det(A + B) \neq \det(A) + \det(B).$$

**Verfahren 5.6 (Praktische Berechnung der Determinante):** (a) Wenn eine Zeile (oder eine Spalte) von  $A$  der Nullvektor ist, dann gilt  $\det(A) = 0$ ;

(b) Multipliziert man einen Zeilenvektor (oder einen Spaltenvektor) von  $A$  mit der Zahl  $\lambda$ , so ändert sich die Determinante um den Faktor  $\lambda$  (siehe Bild 5.15);

(c) Vertauscht man zwei Zeilen (oder zwei Spalten) von  $A$ , so ändert die Determinante das Vorzeichen;

(d) Addition eines skalaren Vielfachen einer Zeile zu einer anderen Zeile (oder eines skalaren Vielfachen einer Spalte zu einer anderen Spalte) ändert die Determinante nicht (*dies ermöglicht die Verwendung des Gauß-Algorithmus sogar für Zeilen und Spalten !*). Siehe dazu auch Bild 5.16.

(e) Wir können allgemein nach der  $i$ -ten Zeile entwickeln:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij});$$

(f) Wir können auch nach der  $j$ -ten Spalte entwickeln:

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

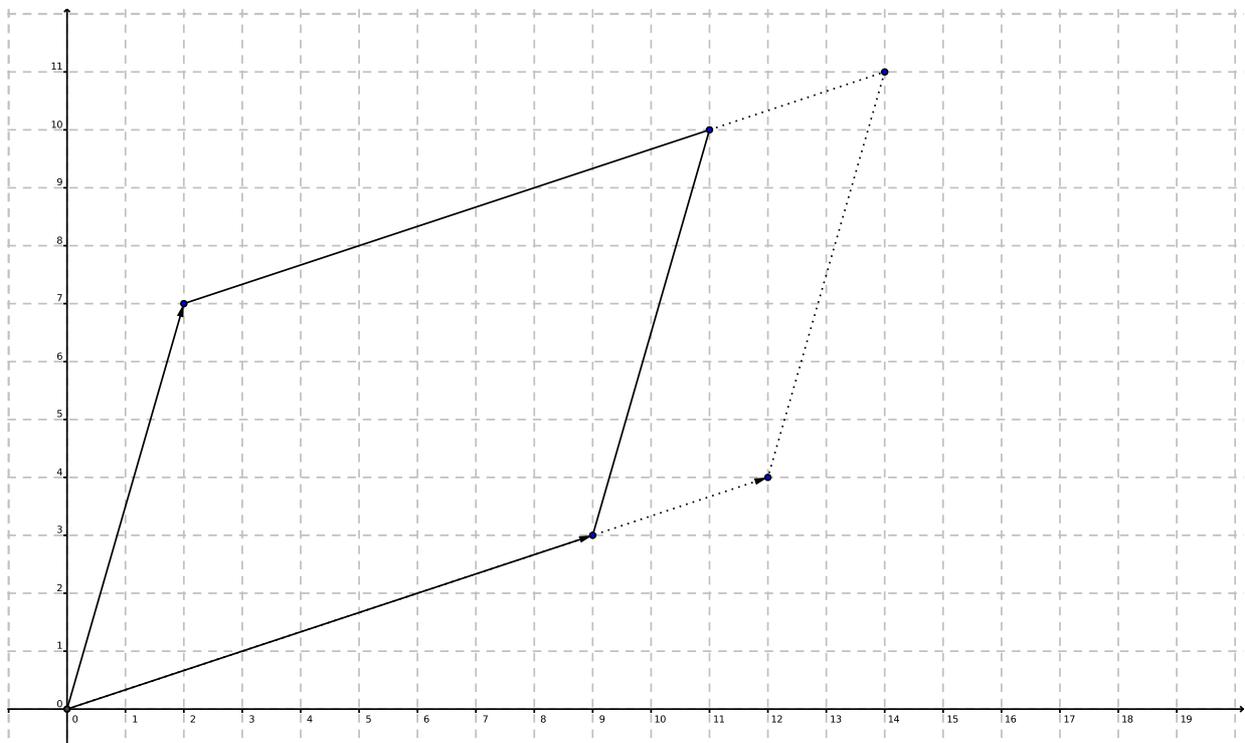


Abbildung 5.15: Es ist  $\det[\frac{4}{3}\vec{a}, \vec{b}] = \frac{4}{3} \det[\vec{a}, \vec{b}]$ . Siehe Bild 5.13 und Verfahren 5.6 (b).

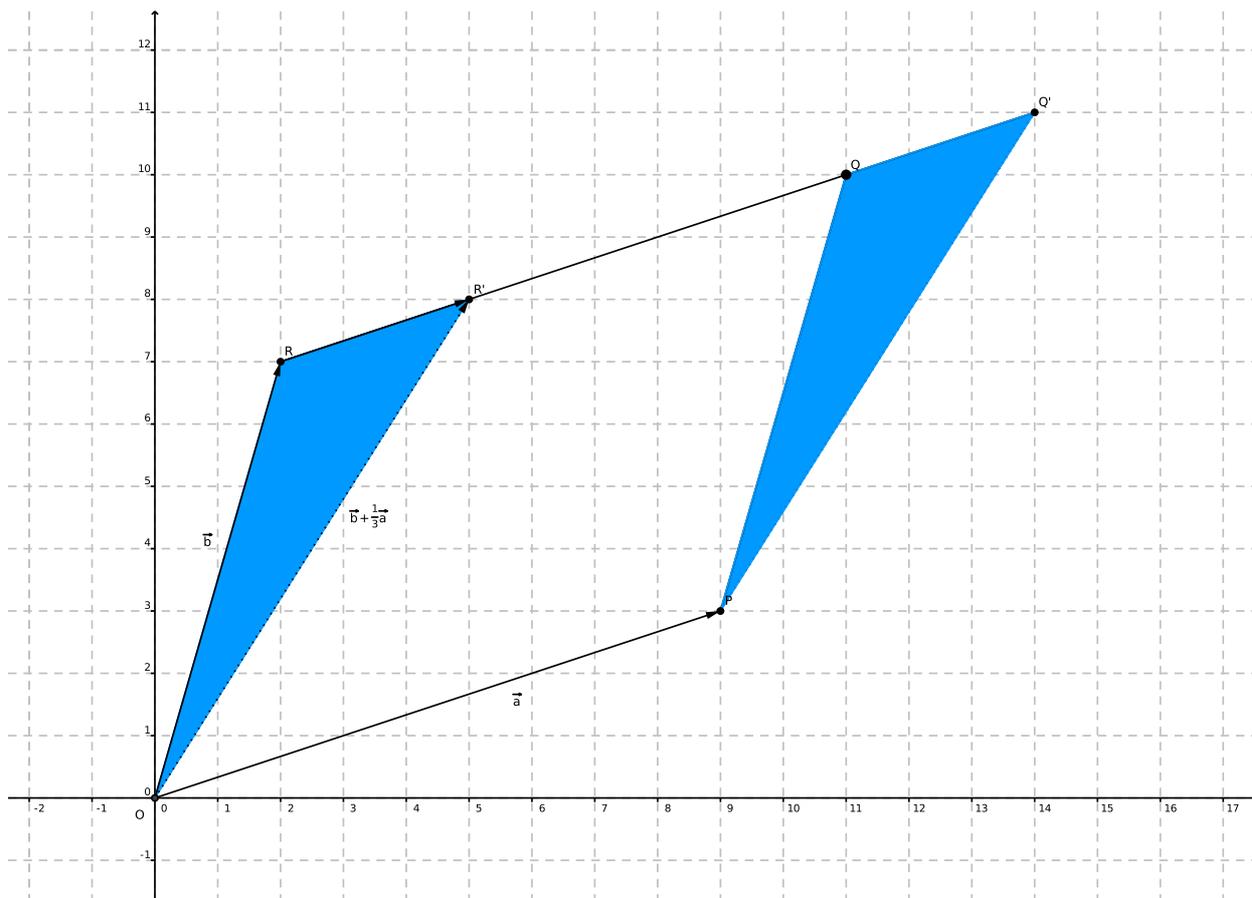


Abbildung 5.16: Die zweikomponentigen Spaltenvektoren  $\vec{a}$  und  $\vec{b}$  können zu einer  $2 \times 2$ -Matrix nebeneinandergestellt werden, und deren Determinante  $\det[\vec{a}, \vec{b}]$  ist der orientierte Flächeninhalt des Parallelogramms  $OPQR$ . Die blauen Dreiecke sind kongruent, also haben die Parallelogramme  $OPQR$  und  $OPQ'R'$  denselben Flächeninhalt, und deshalb ist  $\det[\vec{a}, \vec{b}] = \det[\vec{a}, \vec{b} + \frac{1}{3}\vec{a}]$ . Siehe Verfahren 5.6 (d).

**Beispiel 5.7:**

Wir rechnen eine Determinante aus und beginnen mit der Entwicklung nach der dritten Spalte:

$$\begin{aligned} \begin{vmatrix} 1 & 4 & 2 & 5 \\ 3 & 0 & 1 & 2 \\ 1 & 4 & 0 & 0 \\ 3 & 0 & 1 & 0 \end{vmatrix} &= +2 \cdot \begin{vmatrix} 3 & 0 & 2 \\ 1 & 4 & 0 \\ 3 & 0 & 0 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 4 & 5 \\ 1 & 4 & 0 \\ 3 & 0 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 4 & 5 \\ 3 & 0 & 2 \\ 1 & 4 & 0 \end{vmatrix} \\ &= -2 \cdot \begin{vmatrix} 3 & 0 & 0 \\ 1 & 4 & 0 \\ 3 & 0 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 3 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 4 & 5 \end{vmatrix} - 1 \cdot \begin{vmatrix} 0 & 0 & 5 \\ 3 & 0 & 2 \\ 1 & 4 & 0 \end{vmatrix} \\ &= -2 \cdot 3 \cdot 4 \cdot 2 + 1 \cdot 3 \cdot 4 \cdot 5 - 1 \cdot 5 \cdot \begin{vmatrix} 3 & 0 \\ 1 & 4 \end{vmatrix} \\ &= -48 + 60 - 5 \cdot 12 = -48. \end{aligned}$$

Die roten Vorzeichen entnehmen wir dem Schachbrettmuster

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix},$$

das eine bildliche Darstellung der Faktoren  $(-1)^{i+j}$  liefert.

**Bemerkung 5.8 (Nutzen der Determinante):**

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix und  $b \in \mathbb{R}^n$  ein Spaltenvektor. Dann gilt:

- (a)  $A$  ist invertierbar  $\iff \det(A) \neq 0 \iff \text{rang}(A) = n$ ;
- (b)  $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$  sind linear unabhängig und damit eine Basis des  $\mathbb{R}^n$   $\iff \det([\vec{v}_1, \dots, \vec{v}_n]) \neq 0$ ;
- (c) Das LGS  $A\vec{x} = \vec{b}$  ist eindeutig lösbar  $\iff \det(A) \neq 0$ ;
- (d) Wir haben die **Cramer'sche Regel**: ist  $\det(A) \neq 0$ , dann ist die (eindeutige) Lösung von  $A\vec{x} = \vec{b}$  gegeben durch

$$x_i = \frac{\det([\vec{a}_1, \dots, \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, \dots, \vec{a}_n])}{\det(A)}, \quad i = 1, \dots, n.$$

Aus Gründen des Rechenaufwandes ist diese Regel allerdings nur für kleine  $n$  zu empfehlen.

**Bemerkung 5.9:**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear,  $B$  sei eine Basis von  $\mathbb{R}^n$ ,  $E$  sei die Standardbasis. Wir betrachten die Abbildungsmatrizen  $M_B^B(f)$  und  $M_E^E(f)$ .

Wir wissen, dass es eine invertierbare Matrix  $S$  gibt mit  $M_E^E(f) = SM_B^B(f)S^{-1}$ , nämlich  $S = [\vec{b}_1, \dots, \vec{b}_n]$ . Also ist dann

$$\begin{aligned} \det(M_E^E(f)) &= \det(SM_B^B(f)S^{-1}) = \det(S) \cdot \det(M_B^B(f)) \cdot \det(S^{-1}) \\ &= \det(S) \cdot \det(M_B^B(f)) \cdot \frac{1}{\det(S)} = \det(M_B^B(f)). \end{aligned}$$

Das bedeutet:  $\det(M_B^B(f))$  ist für alle Basen  $B$  des  $\mathbb{R}^n$  dieselbe Zahl. Also ist die Determinante eine Kenngröße der Abbildung  $f$ , und nicht etwa eine Kenngröße der speziellen Abbildungsmatrix bezüglich  $B$  (siehe Bemerkung 5.3 (c)).

## 6 Eigenwerte und Eigenvektoren

Unsere **Ziele** sind:

- zu linearen Abbildungen des  $\mathbb{R}^n$  in sich selbst suchen wir „charakteristische Richtungen“ (wie etwa Drehachsen, Spiegelgeraden, Spiegelebenen, Projektionsebenen, Projektionsrichtungen (siehe Beispiel 4.7 (c)));
- zu einer linearen Abbildung suchen wir eine „angepasste Basis“, sodass die zugehörige Abbildungsmatrix möglichst simpel ist.

**Definition 6.1 (Eigenwert und Eigenvektor):** (a) Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine lineare Abbildung des  $\mathbb{R}^n$  in sich selbst. Dann heißt  $\lambda \in \mathbb{R}$  ein *Eigenwert von  $f$* , wenn es einen Vektor  $\vec{x} \in \mathbb{R}^n$  gibt mit

- $\vec{x} \neq \vec{0}$ ;
- $f(\vec{x}) = \lambda \cdot \vec{x}$ .

Der Vektor  $\vec{x}$  heißt dann *Eigenvektor zum Eigenwert  $\lambda$* .

- (b) Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix. Dann heißt  $\lambda \in \mathbb{R}$  ein *Eigenwert von  $A$* , wenn es einen Vektor  $\vec{x} \in \mathbb{R}^n$  gibt mit

- $\vec{x} \neq \vec{0}$ ;
- $A\vec{x} = \lambda \cdot \vec{x}$ .

**Bemerkung 6.2:**

Beide Teile der Definition sind kompatibel zueinander: Ist  $\lambda$  ein Eigenwert von  $f$ , so ist  $\lambda$  auch ein Eigenwert von jeder Darstellungsmatrix  $M_B^B(f)$ , und zwar unabhängig von der Basis  $B$ .

**Beispiel 6.3:**

Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Spiegelung an  $\text{Lin}(\begin{bmatrix} 1 \\ 1 \end{bmatrix})$ , siehe Bild 6.17. Dann erkennen wir

- $f(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ , also ist  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  ein Eigenvektor von  $f$  zum Eigenwert 1;
- $f(\begin{bmatrix} 1 \\ -1 \end{bmatrix}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ , also ist  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$  ein Eigenvektor von  $f$  zum Eigenwert  $-1$ .

Geometrisch ist glaubhaft, dass  $f$  keine weiteren Eigenwerte besitzt.

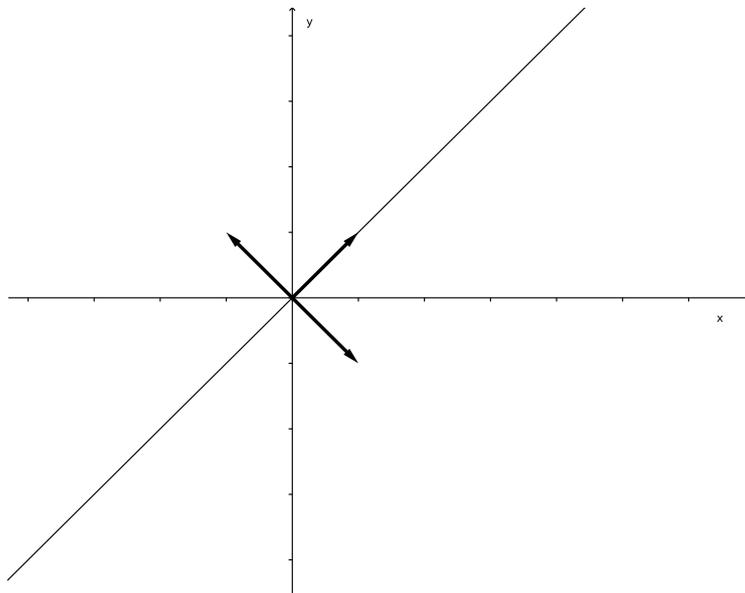


Abbildung 6.17: Die Spiegelung an der Winkelhalbierenden.

**Definition 6.4 (Eigenraum):**

Sei  $\lambda \in \mathbb{R}$  ein Eigenwert von  $A \in \mathbb{R}^{n \times n}$  (bzw. ein Eigenwert einer linearen Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ). Dann heißt

$$E_\lambda := \{ \vec{x} \in \mathbb{R}^n : \vec{x} \text{ ist Eigenvektor von } A \} \cup \{ \vec{0} \}$$

*Eigenraum zum Eigenwert  $\lambda$ .*

Es ist  $E_\lambda$  ein Untervektorraum des  $\mathbb{R}^n$ , denn seien  $\vec{x}, \vec{y} \in E_\lambda$  und  $\alpha, \beta \in \mathbb{R}$ . Dann gilt tatsächlich

$$A(\alpha\vec{x} + \beta\vec{y}) = \alpha A\vec{x} + \beta A\vec{y} = \alpha\lambda\vec{x} + \beta\lambda\vec{y} = \lambda(\alpha\vec{x} + \beta\vec{y}).$$

**Definition / Satz 6.5 (Diagonalisierbarkeit):**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear. Wir setzen voraus, dass es  $n$  Eigenvektoren  $\vec{b}_1, \dots, \vec{b}_n$  von  $f$  gibt, die eine Basis  $B = \{ \vec{b}_1, \dots, \vec{b}_n \}$  bilden. Insbesondere ist dann  $f(\vec{b}_j) = \lambda_j \vec{b}_j$ , für  $j = 1, \dots, n$ , mit Eigenwerten  $\lambda_j$ .

Dann ist  $M_B^B(f)$  eine Diagonalmatrix,

$$M_B^B(f) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

und die Abbildung  $f$  heißt *diagonalisierbar*.

Entsprechend definieren wir den Begriff der Diagonalisierbarkeit für quadratische Matrizen  $A \in \mathbb{R}^{n \times n}$ .

**Achtung:** leider ist nicht jede Matrix diagonalisierbar (siehe Beispiel 6.12) !

**Verfahren 6.6 (Bestimmung von Eigenwerten):**

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix. Es gilt dann

$$A\vec{x} = \lambda\vec{x} \iff A\vec{x} - \lambda\vec{x} = \vec{0} \iff A\vec{x} - \lambda E_n \vec{x} = \vec{0} \iff (A - \lambda E_n)\vec{x} = \vec{0}.$$

Definitionsgemäß ist  $\lambda$  genau dann ein Eigenwert von  $A$ ,

- falls ein Vektor  $\vec{x} \neq \vec{0}$  existiert, der dieses LGS löst;
- also falls das LGS  $(A - \lambda E_n)\vec{x} = 0$  **nicht eindeutig** lösbar ist;
- also falls  $\det(A - \lambda E_n) = 0$  ist (siehe Bemerkung 5.8 (c)).

Insgesamt haben wir also folgendes Ergebnis:

$$\lambda \text{ ist Eigenwert von } A \iff \det(A - \lambda E_n) = 0.$$

Es ist  $\det(A - \lambda E_n)$  ein Polynom von Grad  $n$  in der Variablen  $\lambda$ . Dieses Polynom nennen wir *charakteristisches Polynom*.

Also gilt:

*Die Eigenwerte von  $A$  sind genau die Nullstellen des charakteristischen Polynoms von  $A$ .*

Insbesondere gilt: keine Matrix  $A \in \mathbb{R}^{n \times n}$  kann mehr als  $n$  verschiedene Eigenwerte haben.

**Verfahren 6.7 (Bestimmung von Eigenvektoren):**

Sei  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix, und  $\lambda$  sei ein Eigenwert von  $A$ .

Dann sind die Eigenvektoren zum Eigenwert  $\lambda$  alle diejenigen Vektoren  $\vec{x} \neq \vec{0}$  mit  $A\vec{x} = \lambda\vec{x}$ ; also suchen wir alle Lösungen  $\vec{x}$  des LGS  $(A - \lambda E_n)\vec{x} = \vec{0}$  mit der Einschränkung  $\vec{x} \neq \vec{0}$ .

**Beispiel 6.8:**

Wir betrachten die Matrix

$$A = \begin{bmatrix} 5 & -9 & -3 \\ 4 & -9 & -4 \\ -6 & 15 & 8 \end{bmatrix}.$$

Wir stellen das charakteristische Polynom auf

$$\begin{aligned} \det(A - \lambda E_3) &= \begin{vmatrix} 5 - \lambda & -9 & -3 \\ 4 & -9 - \lambda & -4 \\ -6 & 15 & 8 - \lambda \end{vmatrix} = \begin{vmatrix} 5 - \lambda & -9 & 2 - \lambda \\ 4 & -9 - \lambda & 0 \\ -6 & 15 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda) \begin{vmatrix} 5 - \lambda & -9 & 1 \\ 4 & -9 - \lambda & 0 \\ -6 & 15 & 1 \end{vmatrix} = (2 - \lambda) \begin{vmatrix} 11 - \lambda & -24 & 0 \\ 4 & -9 - \lambda & 0 \\ -6 & 15 & 1 \end{vmatrix} \\ &= (2 - \lambda) \begin{vmatrix} 11 - \lambda & -24 \\ 4 & -9 - \lambda \end{vmatrix} = (2 - \lambda) \cdot (\lambda^2 - 2\lambda - 99 + 96) \\ &= (2 - \lambda) \cdot (\lambda^2 - 2\lambda - 3) = (2 - \lambda) \cdot (\lambda - 3) \cdot (\lambda + 1), \end{aligned}$$

also besitzt die quadratische Matrix  $A$  die Eigenwerte 2, 3,  $-1$ .

Beispielhaft bestimmen wir die Eigenvektoren zum Eigenwert  $\lambda_1 = -2$ . Dazu bestimmen wir die Lösungen des LGS  $(A - 2E_3)\vec{x} = \vec{0}$ :

$$\begin{array}{l}
\begin{array}{c} \boxed{1} \\ \boxed{2} \\ \boxed{3} \end{array} : \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 3 & -9 & -3 & 0 \\ 4 & -11 & -4 & 0 \\ -6 & 15 & 6 & 0 \end{array} \rightsquigarrow \begin{array}{c} \frac{1}{3} \cdot \boxed{1} \\ \frac{1}{3} \cdot \boxed{3} \end{array} =: \begin{array}{c} \boxed{4} \\ \boxed{5} \end{array} : \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & -3 & -1 & 0 \\ 4 & -11 & -4 & 0 \\ -2 & 5 & 2 & 0 \end{array} \\
\rightsquigarrow \begin{array}{c} \boxed{2} - 4 \cdot \boxed{4} \\ \boxed{5} + 2 \cdot \boxed{4} \end{array} =: \begin{array}{c} \boxed{6} \\ \boxed{7} \end{array} : \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & -3 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \\
\rightsquigarrow \begin{array}{c} \boxed{4} + 3 \cdot \boxed{6} \\ \boxed{6} + \boxed{7} \end{array} =: \begin{array}{c} \boxed{9} \\ \boxed{8} \end{array} : \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}
\end{array}$$

Die zugehörigen Lösungen sind

$$t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Also sind  $[t, 0, t]^\top$  mit  $t \in \mathbb{R}$  und  $t \neq 0$  alle Eigenvektoren zum Eigenwert  $\lambda_1 = 2$ .

Die Matrix  $A$  ist diagonalisierbar, denn es gibt drei verschiedene Eigenwerte und es gilt Satz 6.9 (a).

**Satz 6.9 (Regeln für Eigenwerte und Eigenvektoren):** (a) Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

(b) Die Matrizen  $A$  und  $A^\top$  haben dieselben Eigenwerte, denn

$$\det(A^\top - \lambda E_n) = \det(A^\top - \lambda E_n^\top) = \det(A^\top - (\lambda E_n)^\top) = \det((A - \lambda E_n)^\top) = \det(A - \lambda E_n).$$

Achtung: die zugehörigen Eigenvektoren sind im Allgemeinen unterschiedlich!

(c) Sei  $A$  eine invertierbare Matrix und  $A\vec{x} = \lambda\vec{x}$ . Dann ist  $A^{-1}\vec{x} = \frac{1}{\lambda}\vec{x}$ . Also gilt: falls  $\vec{x}$  ein Eigenvektor von  $A$  zum Eigenwert  $\lambda \neq 0$  ist, dann ist  $\vec{x}$  auch ein Eigenvektor von  $A^{-1}$  zum Eigenwert  $\frac{1}{\lambda}$ .

**Beispiel 6.10:**

Wir bestimmen die Eigenwerte und Eigenvektoren zu

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

Dazu stellen wir das charakteristische Polynom auf:

$$\det(A - \lambda E_3) = \begin{vmatrix} 2 - \lambda & 0 & 0 \\ -1 & 2 - \lambda & -1 \\ -1 & 0 & 1 - \lambda \end{vmatrix} = (2 - \lambda) \begin{vmatrix} 2 - \lambda & -1 \\ 0 & 1 - \lambda \end{vmatrix} = (2 - \lambda)^1 \cdot (1 - \lambda).$$

Es ergeben sich die Eigenwerte  $\lambda_1 = \lambda_2 = 2$  und  $\lambda_3 = 1$ . Man rechnet nach, dass ein Eigenvektor zum Eigenwert  $\lambda_3 = 1$  der Vektor  $[0, 1, 1]^\top$  ist. Wir bestimmen jetzt die Eigenvektoren zu  $\lambda_1 = \lambda_2 = 2$ :

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & -1 & 0 \end{array} \rightsquigarrow \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Dieses LGS ist schnell gelöst, und der Eigenraum zum Eigenwert 2 ist demnach

$$E_2 = \text{Lin} \left( \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right).$$

Wir können also für den  $\mathbb{R}^3$  eine Basis aus Eigenvektoren angeben, nämlich

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Also ist die Matrix  $A$  diagonalisierbar, und nach einem Basiswechsel auf die Eigenvektorbasis gilt dann

$$S^{-1}AS = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{mit} \quad S = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

**Bemerkung 6.11:**

Diese Diagonalform ist sehr praktisch: zum Beispiel wollen wir uns überlegen, wie man mit wenig Aufwand die Potenz  $A^{10} = A \cdot A \cdot \dots \cdot A$  (10 Faktoren) bestimmen kann, wobei  $A$  aus dem vorigen Beispiel stammt.

Dazu bemerken wir, dass  $A = SDS^{-1}$  mit

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Also ist dann

$$\begin{aligned} A^{10} &= (SDS^{-1})^{10} = \cancel{SDS^{-1}}\cancel{SDS^{-1}}\cancel{SD} \dots \cancel{S^{-1}}\cancel{SDS} = SD^{10}S^{-1} \\ &= S \begin{bmatrix} 1024 & 0 & 0 \\ 0 & 1024 & 0 \\ 0 & 0 & 1 \end{bmatrix} S^{-1} \quad \left| \text{bestimme } S^{-1} \text{ mit Gauß-Algorithmus} \right. \\ &= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1024 & 0 & 0 \\ 0 & 1024 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & -1 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1024 & 0 \\ 1024 & 0 & 1 \\ 0 & 1024 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & -1 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1024 & 0 & 0 \\ -1023 & 1024 & -1023 \\ -1023 & 0 & 1 \end{bmatrix}. \end{aligned}$$

**Beispiel 6.12:**

Wir untersuchen die Matrix

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

Deren charakteristisches Polynom lautet

$$\det(A - \lambda E_3) = \begin{vmatrix} 1 - \lambda & 3 & 0 \\ 0 & 1 - \lambda & 2 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^3,$$

weil hier eine obere Dreiecksmatrix vorliegt. Also ist  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ . Jetzt bestimmen wir Eigenvektoren, indem wir das Gleichungssystem  $(A - E_3)\vec{x} = \vec{0}$  lösen:

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Der Eigenraum zum Eigenwert 1 ermittelt sich dann zu

$$E_1 = \text{Lin} \left( \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right).$$

Andere Eigenvektoren gibt es nicht. Wir schaffen es demnach nicht, drei linear unabhängige Eigenvektoren für diese Matrix  $A$  zu finden. Leider ist  $A$  also nicht diagonalisierbar.

**Satz 6.13 (Symmetrische Matrizen sind diagonalisierbar):**

Sei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix. Dann gilt: es ist immer möglich,  $n$  linear unabhängige Eigenvektoren von  $A$  zu finden, die sogar eine Orthonormalbasis des  $\mathbb{R}^n$  sind. Das bedeutet:

- $A$  ist diagonalisierbar;
- es existiert eine orthogonale Matrix  $S \in \mathbb{R}^{n \times n}$  mit

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Beachte, dass wir die Vereinfachung  $S^{-1} = S^T$  haben, weil  $S$  eine orthogonale Matrix ist (siehe Satz 4.14).

## 7 Quadratische Formen und Kegelschnitte

Unser **Ziel** ist: wir wollen geometrische Figuren beschreiben wie etwa die folgenden:

- Kreise, Kugeln, Ellipsen, Ellipsoide,
- Paraboloid, Hyperboloid, ...

**bisher:** Geraden / Ebenen / ... sind Lösungsmengen *linearer* Gleichungen;

**jetzt:** die uns interessierenden Figuren sind Lösungsmengen *quadratischer* Gleichungen.

Die *allgemeine quadratische Gleichung in  $n$  Variablen* lautet

$$\vec{x}^T \cdot A \cdot \vec{x} + \vec{a}^T \cdot \vec{x} + \alpha = 0, \quad (\text{Q})$$

und hierbei haben wir folgende gegebene Koeffizienten: eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$ , ein Spaltenvektor  $\vec{a} \in \mathbb{R}^n$ , und eine reelle Zahl  $\alpha$ .

Gesucht sind alle Spaltenvektoren  $\vec{x} \in \mathbb{R}^n$ , die (Q) lösen.

**Beispiel 7.1:**

Im  $\mathbb{R}^2$  betrachten wir

$$Q(x, y) = x^2 - 2y^2 + 4xy + \sqrt{5}x + 2\sqrt{5}y - \frac{3}{4} = 0.$$

Wir führen die Schreibweise  $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$  ein und können dann die Gleichung  $Q(x, y) = 0$  auf folgende Form bringen:

$$\begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \sqrt{5} & 2\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} - \frac{3}{4} = 0,$$

also bekommen wir die Gleichung (Q) mit

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}, \quad \vec{a} = \begin{bmatrix} \sqrt{5} \\ 2\sqrt{5} \end{bmatrix}, \quad \alpha = -\frac{3}{4}.$$

Die uns interessierende **Frage** ist jetzt: was beschreibt diese Gleichung geometrisch ?

**Definition 7.2 (Quadriken und Kegelschnitte):**

Die Lösungsmenge einer Gleichung der Form (Q) heißt *Quadrik* oder *Kegelschnitt*.

Um die geometrische Gestalt einer Quadrik besser verstehen zu können, werden wir zu einem besser geeigneten Koordinatensystem übergehen. Das Verfahren dazu besteht aus mehreren Schritten und verläuft wie folgt:

**Verfahren 7.3 (Transformation auf Normalform):**

**Schritt 1: diagonalisiere  $A$  (Hauptachsentransformation).** Laut Voraussetzung ist die Matrix  $A$  symmetrisch. Gemäß Satz 6.13 existiert eine ONB  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  aus Eigenvektoren, die beim Nebeneinanderstellen eine Matrix  $S$  ergeben,  $S = [\vec{b}_1, \dots, \vec{b}_n]$ , und diese Matrix  $S$  ist sogar eine orthogonale Matrix, und es gilt

$$S^T A S = S^{-1} A S = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} =: A',$$

wobei  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$  sind.

**Schritt 2: substituiere in (Q).** Wir führen neue Koordinaten  $\vec{x}'$  ein gemäß  $\vec{x} = S\vec{x}'$ . Wenn wir dies in (Q) einsetzen, dann folgt

$$\begin{aligned} (S\vec{x}')^T A S\vec{x}' + \vec{a}^T (S\vec{x}') + \alpha &= 0 \\ \iff \vec{x}'^T S^T A S\vec{x}' + (S^T \vec{a})^T \vec{x}' + \alpha &= 0 \\ \iff \vec{x}'^T A' \vec{x}' + \vec{a}'^T \vec{x}' + \alpha &= 0 \quad \text{mit } \vec{x}' := S^T \vec{x}. \end{aligned}$$

Der Nutzen dieses Schrittes besteht darin, dass in dieser quadratischen Gleichung jetzt keine gemischten Produkte mehr enthalten sind, denn  $A'$  ist eine Diagonalmatrix.

**Schritt 3: verschiebe den Ursprung.** Durch quadratische Ergänzung kann nun für alle  $i \in \{1, \dots, n\}$  mit  $\lambda_i \neq 0$  der Summand  $a'_i x_i$  zum Verschwinden gebracht werden. Das liefert einen erneuten Koordinatenwechsel von  $\vec{x}'$  zu  $\vec{x}''$ . Hierbei wird der Ursprung des Koordinatensystems verschoben.

Das Ergebnis ist dann die *Normalform*

$$\vec{x}''^T A' \vec{x}'' + \vec{a}''^T \vec{x}'' + \alpha' = 0, \quad \text{wobei } a''_i = 0 \quad \text{wenn immer } \lambda_i \neq 0.$$

**Beispiel 7.4:**

Wir setzen Beispiel 7.1 fort und führen die Transformation auf Normalform durch. Die zu untersuchende Quadrik ist die Lösungsmenge der Gleichung

$$Q(x, y) := \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \sqrt{5} & 2\sqrt{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \frac{3}{4} = 0.$$

**Schritt 1:** Das charakteristische Polynom von  $A$  ist

$$\begin{vmatrix} 1 - \lambda & 2 \\ 2 & -2 - \lambda \end{vmatrix} = \lambda^2 + \lambda - 2 - 4 = \lambda^2 + \lambda - 6 = (\lambda - 2)(\lambda + 3),$$

und es ergeben sich die Eigenwerte  $\lambda_1 = 2$  und  $\lambda_2 = -3$ . Die zugehörigen Eigenvektoren (auf Länge Eins normiert) sind dann

$$\begin{aligned} \text{zu } \lambda_1 = 2: \quad \vec{b}_1 &= \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \\ \text{zu } \lambda_2 = -3: \quad \vec{b}_2 &= \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \end{bmatrix}. \end{aligned}$$

Es ist  $\{\vec{b}_1, \vec{b}_2\}$  eine ONB. Also ist

$$S := \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix},$$

und diese Matrix beschreibt eine Drehung um  $26.56\dots^\circ$  im Gegen Uhrzeigersinn. Mit dieser Matrix  $S$  gilt

$$S^T A S = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} =: A'.$$

**Schritt 2:** Es ist

$$\vec{a}' = S^T \vec{a} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{5} \\ 2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix},$$

und deshalb gilt: ein Koordinatenvektor  $\begin{bmatrix} x \\ y \end{bmatrix}$  löst die Gleichung  $Q(x, y) = 0$  genau dann, wenn der neue Koordinatenvektor  $\begin{bmatrix} x' \\ y' \end{bmatrix} := S^T \begin{bmatrix} x \\ y \end{bmatrix}$  folgende Gleichung erfüllt:

$$Q'(x', y') = \begin{bmatrix} x' & y' \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} - \frac{3}{4} = 0.$$

**Schritt 3:** Wir führen die quadratische Ergänzung durch:

$$\begin{aligned} Q'(x', y') &= 0 \\ \iff 2x'^2 - 3y'^2 + 4x' + 3y' - \frac{3}{4} &= 0 \\ \iff 2(x' + 1)^2 - 2 - 3(y' - \frac{1}{2}) - \frac{3}{4} &= 0 \\ \iff 2(x' + 1)^2 - 3 \left( (y' - \frac{1}{2})^2 - \frac{1}{4} \right) - \frac{11}{4} &= 0 \\ \iff 2(x' + 1)^2 - 3 \left( y' - \frac{1}{2} \right)^2 + \frac{3}{4} - \frac{11}{4} &= 0 \\ \iff 2(x' + 1)^2 - 3 \left( y' - \frac{1}{2} \right)^2 - 2 &= 0 \\ \iff (x' + 1)^2 - \frac{(y' - 1/2)^2}{2/3} - 1 &= 0 \\ \iff x''^2 - \frac{y''^2}{2/3} - 1 &= 0, \end{aligned}$$

und dies ist die Normalform. Wir fassen die Umrechnungen zusammen:

$$\begin{aligned} x'' &= x' + 1, & y'' &= y' - \frac{1}{2}, \\ x' &= \frac{2}{\sqrt{5}}x - \frac{1}{\sqrt{5}}y, & y' &= \frac{2}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y. \end{aligned}$$

**Bemerkung 7.5:**

Im Falle  $n = 2$  haben wir folgende Normalformen:

**Fall A:** beide Eigenwerte von  $A$  sind  $\neq 0$ :

$$\begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 &= 0 && \text{Ellipse (evtl. ein Kreis)} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 &= 0 && \text{leere Menge} \\ \frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 &= 0 && \text{Hyperbel} \\ x^2 + a^2 y^2 &= 0 && \text{ein Punkt} \\ x^2 - a^2 y^2 &= 0 && \text{ein Geradenpaar} \end{aligned}$$

**Fall B:** ein Eigenwert von  $A$  ist Null:

$$\begin{aligned} x^2 - 2py &= 0 && \text{Parabel} \\ x^2 - a^2 &= 0 && \text{ein Paar paralleler Geraden} \\ x^2 + a^2 &= 0 && \text{leere Menge} \\ x^2 &= 0 && \text{eine Gerade.} \end{aligned}$$

**Bemerkung 7.6:**

Im Falle  $n = 3$  gibt es bereits 17 verschiedene Normalformen, auf deren Auflistung wir hier verzichten wollen.

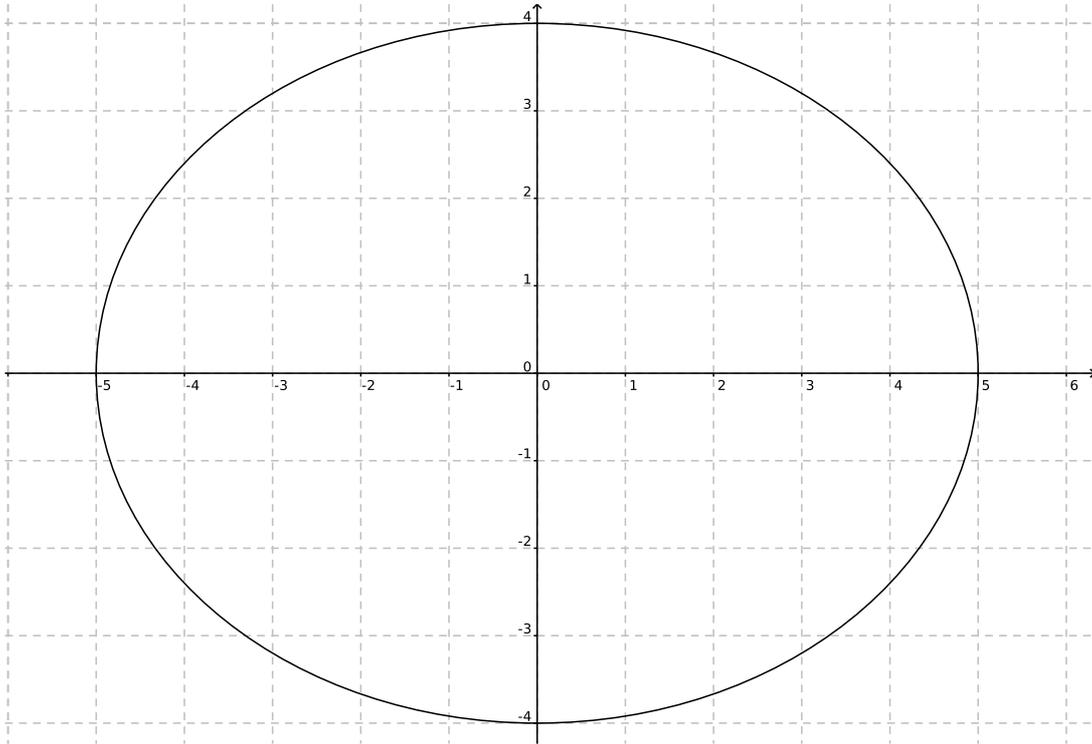


Abbildung 7.18: Eine Normalform-Ellipse mit Parametern  $a = 5$  und  $b = 4$ .

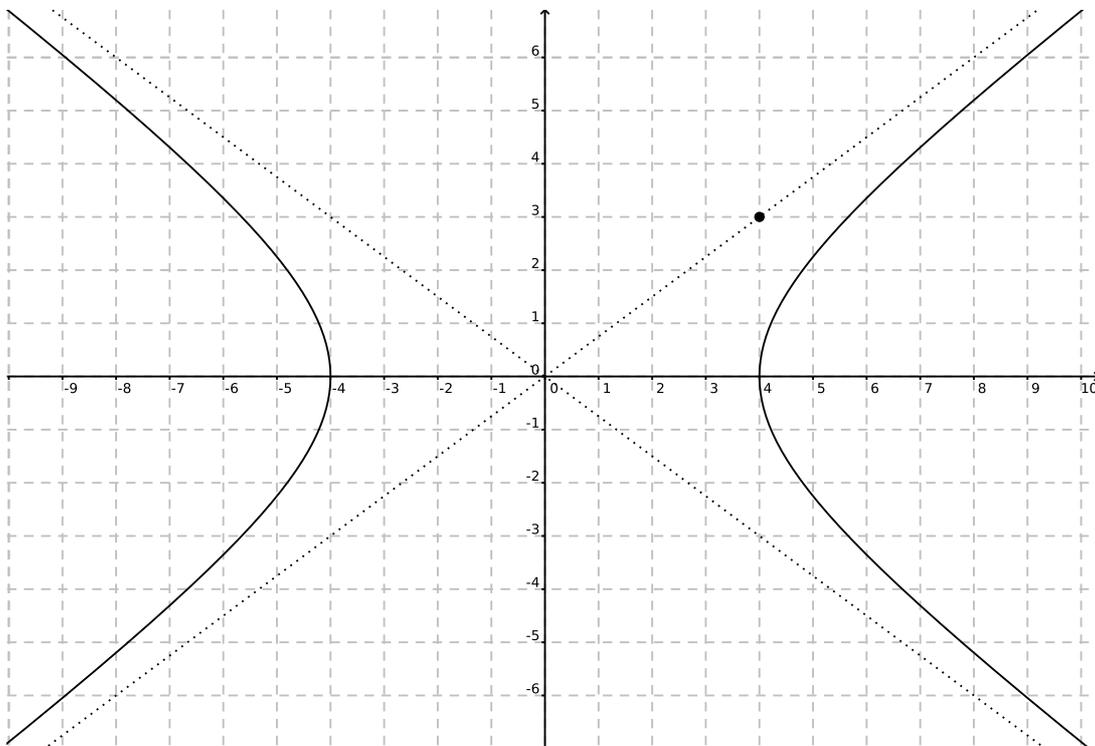
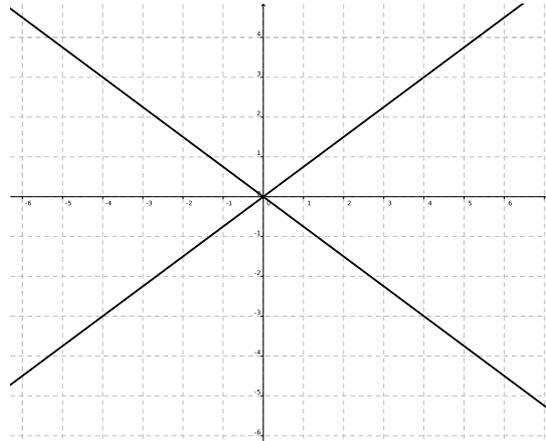
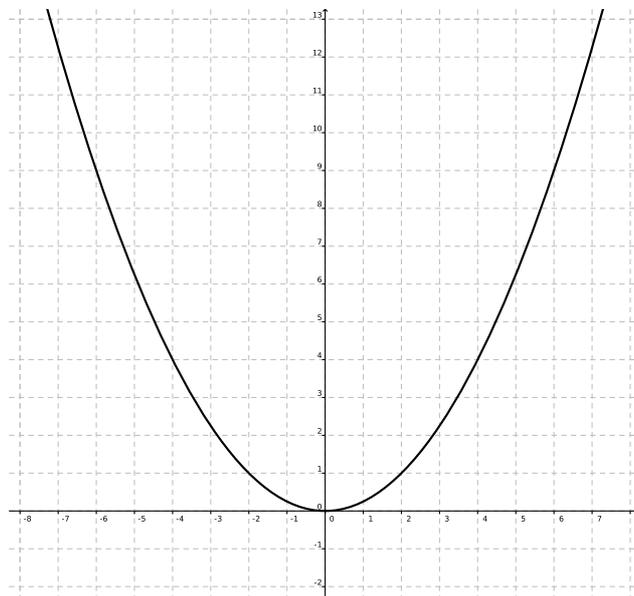


Abbildung 7.19: Eine Normalform-Hyperbel mit Parametern  $a = 4$  und  $b = 3$ .

Abbildung 7.20: Ein Geradenpaar als Lösung der Gleichung  $x^2 - a^2y^2 = 1$  mit  $a = \frac{4}{3}$ .Abbildung 7.21: Eine Normalform-Parabel als Lösung der Gleichung  $x^2 - 2py = 0$  mit  $p = 2$ .**Beispiel 7.7:**

Wir vollenden Beispiel 7.1 und 7.4. Die erhaltene Normalform ist

$$x''^2 - \frac{y''^2}{3} - 1 = 0,$$

also liegt eine Hyperbel vor. Für diese können wir weitere geometrische Informationen angeben: die Halbachsenparameter  $a$  und  $b$  erfüllen  $a^2 = 1$  und  $b^2 = \frac{2}{3}$ , also ist  $a = 1$  und  $b = \sqrt{\frac{2}{3}}$ .

Die Lage des *Mittelpunkts* der Hyperbel ergibt sich durch die Bedingung  $\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ . Wir hatten die Umrechnungen  $x'' = x' + 1$  und  $y'' = y' - \frac{1}{2}$ , also gilt für den Mittelpunkt  $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -1 \\ 1/2 \end{bmatrix}$ . Im  $x - y$ -System bekommen wir wegen  $S^{-1} = S^\top$  dann für die Koordinaten des Zentrums:

$$\begin{bmatrix} x \\ y \end{bmatrix} = S \begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1/2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -5/2 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{5}}{2} \\ 0 \end{bmatrix}.$$

Die Matrix  $S$  beschreibt eine Drehung um  $26.56^\circ$ , also ist die  $x''$ -Achse gegen die  $x$ -Achse um  $26.56^\circ$  gedreht.

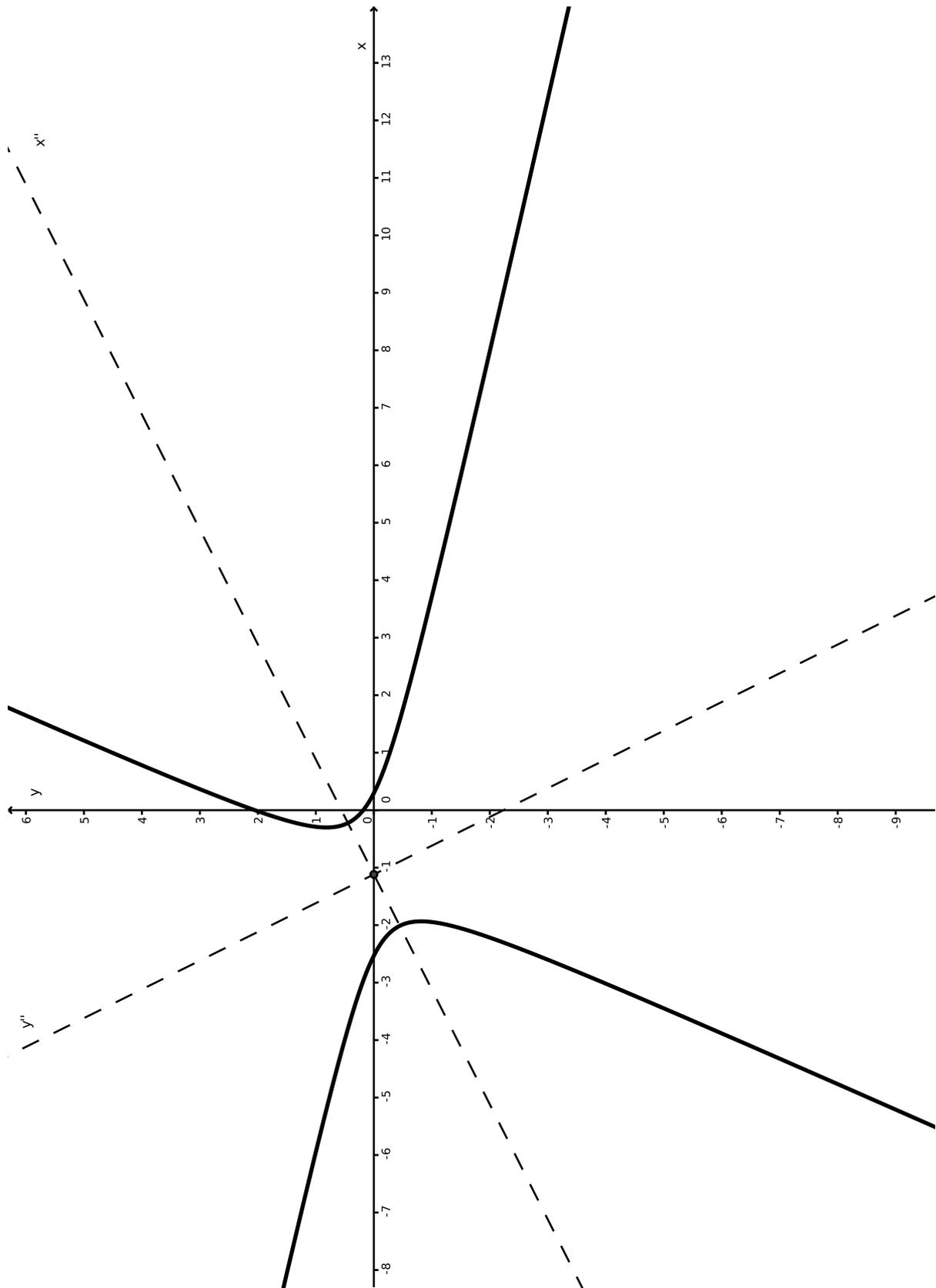


Abbildung 7.22: Die Quadrik aus Beispiel 7.1 und Beispiel 7.4 (aus Papierformatsgründen ist das Bild um  $90^\circ$  gedreht worden).

# Kapitel 2

## Differentialrechnung in mehreren Variablen

### 1 Funktionen mehrerer Veränderlicher

**Definition 1.1 (Funktionen mehrerer Veränderlicher):**

Sei  $M \subseteq \mathbb{R}^n$ . Eine Abbildung  $f: M \rightarrow \mathbb{R}$  heißt *skalare Funktion in  $n$  Variablen* bzw. *skalare Funktion von  $n$  Veränderlichen*. Die Menge  $M$  heißt *Definitionsbereich* (oder *Definitionsbereich*) von  $f$ .

**Beispiele 1.2:** (a)  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(\vec{x}) = c$  (wobei  $c \in \mathbb{R}$  eine feste Konstante ist). Das ist also eine konstante Funktion.

(b)  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(\vec{x}) = \vec{x}^\top A \vec{x}$  (wobei  $A \in \mathbb{R}^{n \times n}$  eine konstante Matrix ist). Wir reden hier von einer allgemeinen quadratischen Funktion.

(c) Wir wählen

$$M := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x \neq 0 \right\}$$

und betrachten die Funktion  $f: M \rightarrow \mathbb{R}$ , die beschrieben wird durch die Vorschrift

$$f \left( \begin{bmatrix} x \\ y \end{bmatrix} \right) := \arctan \left( \frac{y}{x} \right).$$

In Zukunft schreiben wir  $f(x, y)$  anstatt des sperrigen  $f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ .

**Definition 1.3 (Vektorfelder):**

Sei  $M \subseteq \mathbb{R}^n$ , und  $f_1, \dots, f_m: M \rightarrow \mathbb{R}$  seien skalare Funktionen. Dann heißt die Funktion  $\vec{f}: M \rightarrow \mathbb{R}^m$  mit

$$\vec{f}(\vec{x}) := \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$$

*vektorwertige Funktion von  $n$  Veränderlichen* oder kurz *Vektorfeld*. Die Funktionen  $f_1, \dots, f_m$  heißen *Koordinatenfunktionen*.

**Beispiele 1.4:** (a) die Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $\vec{f}(\vec{x}) = \vec{c}$  (wobei  $\vec{c} \in \mathbb{R}^m$  fest ist) ist eine konstante Funktion.

(b) Für jede konstante Matrix  $A \in \mathbb{R}^{m \times n}$  ist  $\vec{f}: \mathbb{R}^n \times m$  mit  $\vec{f}(\vec{x}) := A\vec{x}$  eine lineare Funktion (lineare Abbildung).

- (c) Sei  $M \subseteq \mathbb{R}^n$ . Dann ist  $\text{id}_M: M \rightarrow \mathbb{R}^n$ , definiert durch  $\text{id}_M(\vec{x}) := \vec{x}$ , die identische Abbildung auf  $M$ , die auch *Identität auf  $M$*  genannt wird.
- (d) Die Umrechnungsformel von Polarkoordinaten auf kartesische Koordinaten ist auch eine Vektorfunktion: wir haben dann  $\vec{f}: [0, \infty) \times (-\pi, \pi]$ , wobei

$$\vec{f}(r, \varphi) := \begin{bmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{bmatrix}.$$

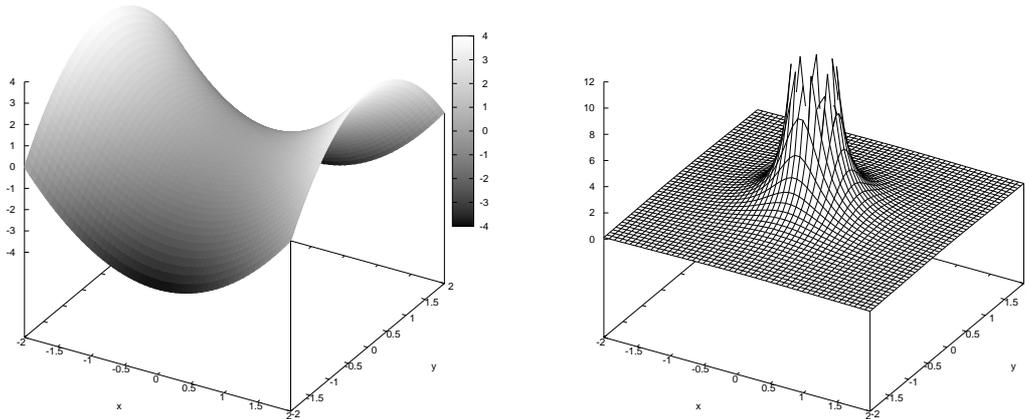


Abbildung 1.1: Links der Graph zur Funktion  $f(x, y) = x^2 - y^2$ , rechts der Graph zur Funktion  $f(x, y) = 1/(x^2 + y^2)$  mit  $M = \{[x, y] \in \mathbb{R}^2: [x, y] \neq [0, 0]\}$ .

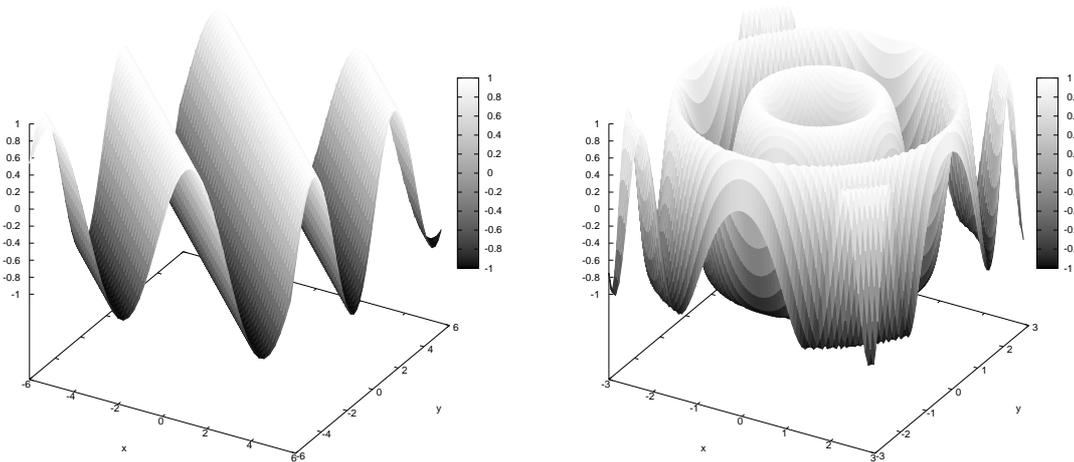


Abbildung 1.2: Links der Graph zur Funktion  $f(x, y) = \sin(x + y)$ , rechts der Graph zur Funktion  $f(x, y) = \sin(x^2 + y^2)$ . Die Höhe ist zusätzlich farblich kodiert worden: je heller, desto größer ist dort der Funktionswert.

**Bemerkung 1.5 (Veranschaulichung solcher Funktionen):** (a) Für Funktionen  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  kann man sich den Graph anschauen wie in den Bildern 1.1 und 1.2. Oder man kann einen Höhenlinienplot anfertigen, vgl. Bild 1.3.

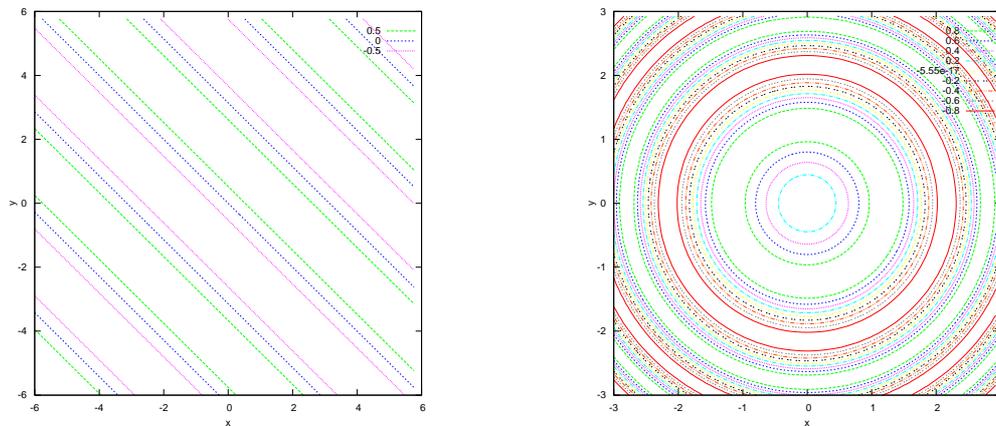


Abbildung 1.3: Links der Höhenlinienplot zur Funktion  $f(x, y) = \sin(x + y)$ , rechts zur Funktion  $f(x, y) = \sin(x^2 + y^2)$ .

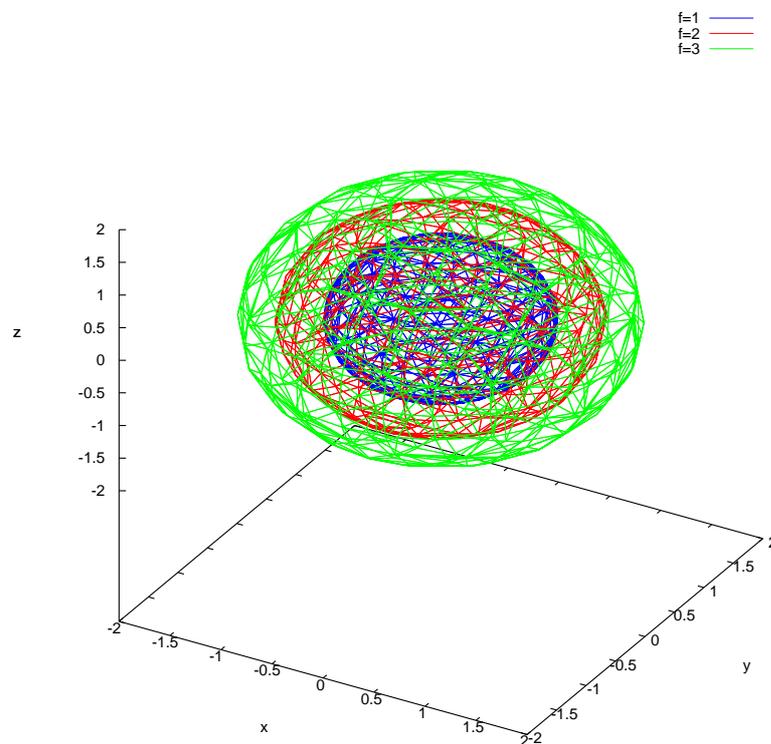


Abbildung 1.4: Drei Niveauflächen zur Funktion  $f(x, y, z) = x^2 + y^2 + z^2$ .

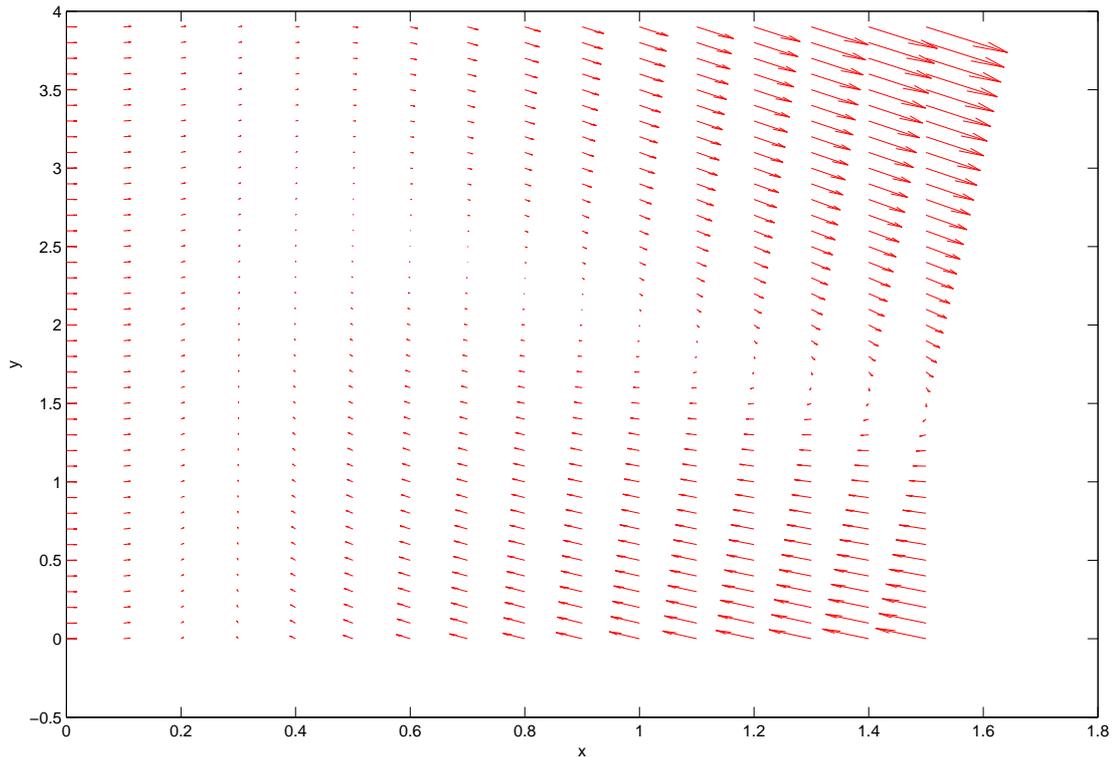


Abbildung 1.5: Ein Vektorfeld in der Ebene.

- (b) Für Funktionen  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$  bietet sich eine Darstellung der Niveauflächen an, vgl. Bild 1.4.
- (c) Für Funktionen  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  verwendet man oft ein Pfeilbild: am Punkt  $[x, y]^T$  wird ein Pfeil  $\vec{f}(x, y)$  angeheftet, vgl. Bild 1.5

**Definition 1.6 (Niveaulinie / Niveaufläche):**

Sei  $M \subseteq \mathbb{R}^n$  und  $f: M \rightarrow \mathbb{R}$  sowie  $c \in \mathbb{R}$  eine feste Konstante. Dann heißt

$$N_c := \{\vec{x} \in M: f(\vec{x}) = c\}$$

*Niveaufläche* bzw. *Niveaulinie* oder *Höhenlinie* von  $f$  zur Höhe  $c$ .

Unser nächstes **Ziel** soll es sein, für solche Funktionen Begriffe wie *Stetigkeit* und *Differenzierbarkeit* zu behandeln.

Dafür benötigen wir

- Grenzwerte im  $\mathbb{R}^n$ ,
- Eigenschaften der Menge  $M$  (offen, abgeschlossen, kompakt, ...).

**Erinnerung:**

- (a) Die Konvergenz von (Zahlen-)Folgen war im ersten Semester behandelt worden.
- (b) Die Konvergenz einer Folge zu einem Grenzwert bedeutet, dass die Glieder der Folge dem Grenzwert beliebig nahe kommen (ab einem bestimmten Folgenindex). Für eine Definition dafür, was „beliebig nahe“ bedeutet, brauchen wir im  $\mathbb{R}^n$  eine Methode, Abstände zu messen. Für genau diesen Zweck hatten wir die Norm  $\|\vec{x}\|$  eines Vektors  $\vec{x} \in \mathbb{R}^n$  definiert:

$$\|\vec{x}\| := \sqrt{\sum_{j=1}^n x_j^2}.$$

Die Norm ist eine nichtnegative reelle Zahl und entspricht für  $n = 2$  und  $n = 3$  der üblichen geometrischen Länge des Pfeiles. Für  $\vec{x}, \vec{y} \in \mathbb{R}^n$  ist dann  $\|\vec{x} - \vec{y}\|$  der Abstand von  $\vec{x}$  zu  $\vec{y}$ .

**Definition 1.7 (Konvergenz von Folgen im  $\mathbb{R}^n$ ):**

Eine Folge

$$(\vec{x}_k)_{k \in \mathbb{N}} = \left( \begin{array}{c} x_{1,k} \\ \vdots \\ x_{n,k} \end{array} \right)_{k \in \mathbb{N}}$$

im  $\mathbb{R}^n$  heißt *konvergent*, falls alle Koordinatenfolgen  $(x_{1,k})_{k \in \mathbb{N}}, \dots, (x_{n,k})_{k \in \mathbb{N}}$  konvergent sind. Dann gilt

$$\lim_{k \rightarrow \infty} \vec{x}_k = \lim_{k \rightarrow \infty} \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{n,k} \end{bmatrix} = \begin{bmatrix} \lim_{k \rightarrow \infty} x_{1,k} \\ \vdots \\ \lim_{k \rightarrow \infty} x_{n,k} \end{bmatrix},$$

was wir auch *koordinatenweise Konvergenz* nennen.

Wenn die Folge  $(\vec{x}_k)_{k \in \mathbb{N}}$  nicht konvergent ist, dann nennt man sie *divergent*.

**Beispiele 1.8:** (a) Betrachte die Vektoren  $\vec{x}_k := [\frac{1}{k}, \sqrt[k]{k}]^\top \in \mathbb{R}^2$ . Dann ist

$$\lim_{k \rightarrow \infty} \vec{x}_k = \begin{bmatrix} \lim_{k \rightarrow \infty} \frac{1}{k} \\ \lim_{k \rightarrow \infty} \sqrt[k]{k} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

(b) Betrachte die Vektoren  $\vec{x}_k := [k, \frac{1}{k}]^\top \in \mathbb{R}^2$ . Die von diesen Vektoren gebildete Folge  $(\vec{x}_k)_{k \in \mathbb{N}}$  ist divergent, weil  $\lim_{k \rightarrow \infty} k$  nicht existiert.

Wir erinnern uns, dass es im Vektorraum  $\mathbb{R}^n$  einige Rechenoperationen gibt:

- Vektor plus Vektor ergibt Vektor,
- Zahl mal Vektor ergibt Vektor,
- Vektor mal Vektor ergibt Zahl (Skalarprodukt),
- die Norm eines Vektors ist eine Zahl.

Und für diese Rechenoperationen gelten ihrerseits Regeln (siehe zum Beispiel Regeln 1.1). Die Rechenoperationen und deren Regeln erzeugen zusammen eine gewisse Struktur, und die Teile (b) und (c) des folgenden Satzes besagen, dass der Grenzwertbegriff zu dieser Struktur kompatibel ist.

**Satz 1.9 (Grenzwertsätze):**

Seien  $(\vec{x}_k)_{k \in \mathbb{N}}$  und  $(\vec{y}_k)_{k \in \mathbb{N}}$  Folgen im  $\mathbb{R}^n$ . Sei  $c \in \mathbb{R}$ . Dann gelten folgende Aussagen:

- (a)  $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0 \iff \lim_{k \rightarrow \infty} \|\vec{x}_k - \vec{x}_0\| = 0$  (beachte, dass der rechte Limes sich auf eine Zahlenfolge bezieht, keine Vektorfolge),
- (b) wenn  $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0$ , dann ist  $\lim_{k \rightarrow \infty} \|\vec{x}_k\| = \|\vec{x}_0\|$ ,
- (c) wenn  $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0$  und  $\lim_{k \rightarrow \infty} \vec{y}_k = \vec{y}_0$ , dann ist
  - $\lim_{k \rightarrow \infty} (\vec{x}_k \pm \vec{y}_k) = \vec{x}_0 \pm \vec{y}_0$ ,
  - $\lim_{k \rightarrow \infty} (c\vec{x}_k) = c\vec{x}_0$ ,
  - $\lim_{k \rightarrow \infty} \langle \vec{x}_k, \vec{y}_k \rangle = \langle \vec{x}_0, \vec{y}_0 \rangle$ .

Wir haben jetzt also den Grenzwertbegriff für Vektorfolgen definiert und einige seiner Eigenschaften erwähnt. Als nächstes benutzen wir diesen Grenzwertbegriff für Vektorfolgen, um Stetigkeit für Funktionen mehrerer Veränderlicher zu definieren.

**Definition 1.10 (Stetigkeit):**

Sei  $M \subseteq \mathbb{R}^n$ ,  $f: M \rightarrow \mathbb{R}$  und  $\vec{x}_0 \in M$ .

- (a) Die Funktion  $f$  heißt *stetig in  $x_0$* , wenn für **jede** Folge  $(\vec{x}_k)_{k \in \mathbb{N}}$  in  $M$  mit  $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0$  folgendes gilt:

$$\lim_{k \rightarrow \infty} f(\vec{x}_k) = f\left(\lim_{k \rightarrow \infty} \vec{x}_k\right) = f(\vec{x}_0).$$

Ansonsten heißt  $f$  *unstetig in  $\vec{x}_0$* .

- (b) Die Funktion  $f$  heißt *stetig auf  $M' \subseteq M$* , falls  $f$  stetig ist in jedem Punkt  $\vec{x}_0 \in M'$ .
- (c) Eine Vektorfunktion  $f: M \rightarrow \mathbb{R}^m$  mit  $f = [f_1, f_2, \dots, f_m]^\top$  ist genau dann stetig in einem Punkt  $\vec{x}_0 \in M$  (bzw. stetig auf  $M' \subseteq M$ ), falls alle Koordinatenfunktionen  $f_1, \dots, f_m$  das sind.

Als nächstes schauen wir uns an, wie man aus einfachen Funktionen, denen man schnell ihre Stetigkeit ansieht, kompliziertere stetige Funktionen zusammenbauen kann.

**Satz 1.11 (Baukastenprinzip):** (a) Seien  $M \subseteq \mathbb{R}^n$  und  $f, g: M \rightarrow \mathbb{R}$  in  $\vec{x}_0 \in M$  stetig. Sei  $c \in \mathbb{R}$ . Dann sind auch die Funktionen  $f + g$ ,  $f - g$ ,  $f \cdot g$  und  $c \cdot f$  stetig in  $\vec{x}_0$ . Wenn zusätzlich  $g(\vec{x}_0) \neq 0$  ist, dann ist auch  $\frac{f}{g}$  stetig in  $\vec{x}_0$ .

- (b) Wenn  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  stetig in  $\vec{x}_0 \in \mathbb{R}^n$  ist, und wenn  $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$  stetig in  $f(\vec{x}_0) \in \mathbb{R}^m$  ist, so ist  $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^p$  stetig in  $\vec{x}_0$ . Das ist das Verkettungsprinzip. Hierbei sind natürlich  $n = 1$ ,  $m = 1$  und  $p = 1$  zugelassen.

**Beispiele 1.12:** (a) konstante, lineare und quadratische Funktionen sind auf ganz  $\mathbb{R}^n$  stetig.

- (b) Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^2 \sin(x - y^3) + xy$  ist stetig auf  $\mathbb{R}^2$ , denn sie ergibt sich über die Summe/Produkt/Verkettung von stetigen Funktionen.
- (c) Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , die gegeben wird durch

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & : \text{wenn } (x, y) \neq (0, 0), \\ 0 & : \text{wenn } (x, y) = (0, 0), \end{cases}$$

ist stetig auf  $\mathbb{R}^2 \setminus \{[0, 0]^\top\}$  als Quotient von stetigen Funktionen mit Nenner ungleich Null.

Allerdings ist sie unstetig im Punkt  $[0, 0]^\top$ , denn für die Folge

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} := \begin{bmatrix} \frac{1}{k} \\ \frac{1}{k} \end{bmatrix}, \quad k \in \mathbb{N}, \quad \text{gilt zwar } \lim_{k \rightarrow \infty} \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

aber es ist

$$\lim_{k \rightarrow \infty} f(x_k, y_k) = \lim_{k \rightarrow \infty} f\left(\frac{1}{k}, \frac{1}{k}\right) = \lim_{k \rightarrow \infty} \frac{\frac{1}{k^2}}{\frac{1}{k^2} + \frac{1}{k^2}} = \lim_{k \rightarrow \infty} \frac{1}{2} = \frac{1}{2} \neq 0 = f(0, 0).$$

Siehe auch Bild 1.6 und Bild 1.7 (linke Hälfte).

- (d) Die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , die gegeben wird durch

$$f(x, y) = \begin{cases} \frac{x^2 y^2}{x^2 + y^2} & : \text{wenn } (x, y) \neq (0, 0), \\ 0 & : \text{wenn } (x, y) = (0, 0), \end{cases}$$

ist stetig auf  $\mathbb{R}^2 \setminus \{[0, 0]^\top\}$ , mit derselben Begründung wie in (c).

Und sie ist stetig im Punkt  $[0, 0]^\top$ , denn

$$\text{sei } \left( \begin{bmatrix} x_k \\ y_k \end{bmatrix} \right)_{k \in \mathbb{N}} \text{ eine Folge mit } \lim_{k \rightarrow \infty} \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Dann haben wir

$$0 \leq |f(x_k, y_k)| = \frac{x_k^2 y_k^2}{x_k^2 + y_k^2} \leq \frac{x_k^2 y_k^2}{x_k^2} = y_k^2,$$

und wenn wir jetzt in dieser Ungleichungskette  $k \rightarrow \infty$  schicken, dann konvergiert der linke Term nach Null, der rechte auch, und  $|f(x_k, y_k)|$  ist dazwischen eingesperrt. Es ist also  $\lim_{k \rightarrow \infty} f(x_k, y_k) = 0 = f(0, 0)$ . Siehe Bild 1.7 (rechte Hälfte).

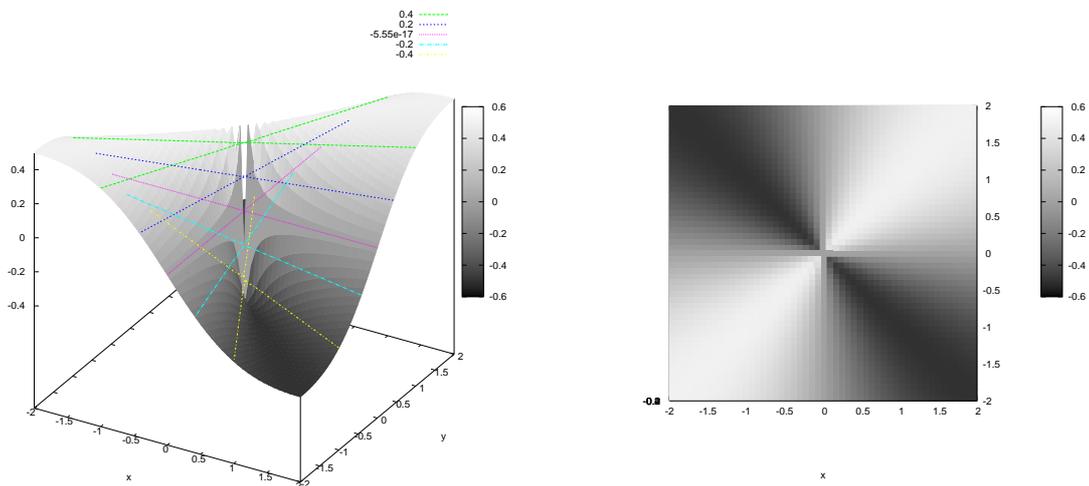


Abbildung 1.6: Links: die skalare Funktion  $f(x, y) = \frac{xy}{x^2 + y^2}$  über dem Quadrat  $[-2, 2] \times [-2, 2]$ . Die farbigen Linien sind Geraden, die in den Graphen eingezeichnet sind, und dort hat  $f$  jeweils den gleichen Wert. Rechts sieht man den Graphen derselben Funktion in der Draufsicht. Diese Funktion ist unstetig im Nullpunkt. Siehe auch die nächste Abbildung und Beispiel 1.12 (c).

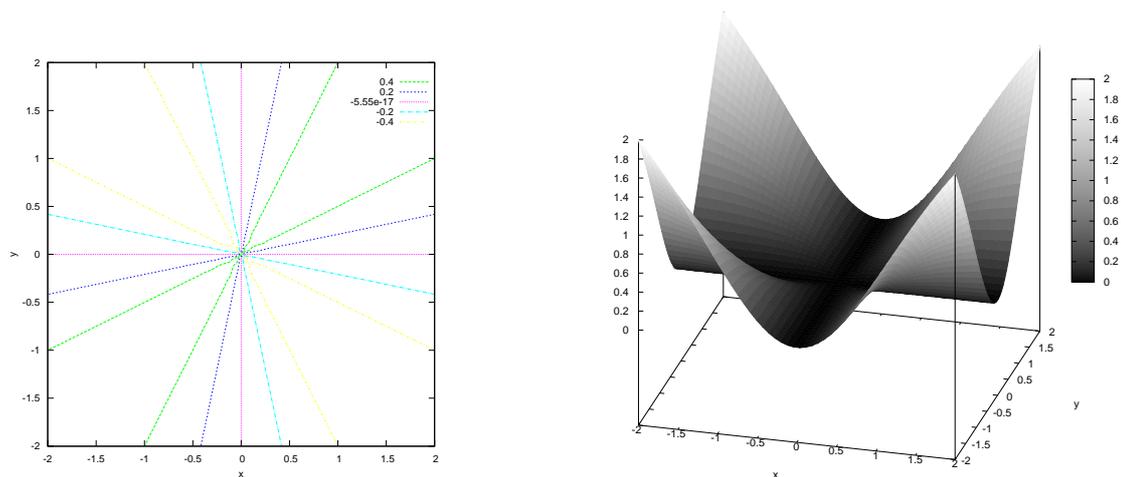


Abbildung 1.7: Links der Höhenlinienplot zur Funktion  $f(x, y) = \frac{xy}{x^2 + y^2}$ , und rechts die Funktion  $f(x, y) = \frac{x^2 y^2}{x^2 + y^2}$ .

**Definition 1.13 (Topologische Begriffe):**

Sei  $M \subseteq \mathbb{R}^n$ .

- (a) Für  $\varepsilon > 0$  und  $\vec{x}_0 \in \mathbb{R}^n$  heißt

$$U_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| < \varepsilon\}$$

$\varepsilon$ -Umgebung um  $\vec{x}_0$  oder auch offene  $\varepsilon$ -Kugel um  $\vec{x}_0$ .

- (b) Ein Punkt  $\vec{x}_0 \in M$  heißt *innerer Punkt* von  $M$ , falls es einen Radius  $\varepsilon > 0$  gibt mit  $U_\varepsilon(\vec{x}_0) \subseteq M$ .
- (c) Die Menge  $M$  heißt *offen*, falls alle Punkte von  $M$  innere Punkte von  $M$  sind.
- (d) Der Punkt  $\vec{x}_0$  heißt *Häufungspunkt* von  $M$ , wenn es eine Folge  $(\vec{x}_k)_{k \in \mathbb{N}}$  gibt, die in  $M$  liegt und gegen  $\vec{x}_0$  konvergiert. (Achtung: wir machen keine Aussage darüber, ob  $\vec{x}_0$  ein Element von  $M$  ist oder nicht.)
- (e) Die Menge  $M$  heißt *abgeschlossen*, wenn jeder Häufungspunkt von  $M$  auch ein Element von  $M$  ist.
- (f) Die Menge  $M$  heißt *beschränkt*, wenn eine Zahl  $C \geq 0$  existiert mit der Eigenschaft, dass  $\|\vec{x}\| \leq C$  gilt für jeden Punkt  $\vec{x} \in M$ .
- (g) Die Menge  $M \subseteq \mathbb{R}^n$  heißt *kompakt*, wenn  $M$  beschränkt und abgeschlossen ist.

Die anschaulichen Interpretationen sind: die Menge  $U_\varepsilon(\vec{x}_0)$  ist das Innere der  $n$ -dimensionalen Kugel mit dem Mittelpunkt  $\vec{x}_0$  und dem Radius  $\varepsilon$ . Die Haut dieser Kugel gehört nicht zu  $U_\varepsilon(\vec{x}_0)$ . Ab jetzt sagen wir oft *Rand* anstatt *Haut*. Eine offene Menge enthält keinen ihrer Randpunkte. Eine abgeschlossene Menge enthält jeden ihrer Randpunkte. Eine Menge ist beschränkt, wenn man sie in einer Kugel von endlichem Radius  $C$  unterbringen kann.

**Warnung 1.14:** (a) Mengen sind keine Türen: die meisten Mengen sind weder offen noch abgeschlossen. Man nehme zum Beispiel das Intervall  $(2, 3]$ .

- (b) Auch Punkte außerhalb von  $M$  können Häufungspunkte von  $M$  sein. Zum Beispiel hat das Intervall  $M = (1, 2) \subset \mathbb{R}^1$  die Häufungspunkte 1, 2, 1.74 (und noch jede weitere Zahl auf der Zahlengeraden, die echt zwischen 1 und 2 liegt). Die beiden Punkte 1 und 2 sind aber kein Element von  $M$ .

**Beispiele 1.15:** (a) Die Menge  $U_\varepsilon(\vec{x}_0)$  ist für beliebige  $\vec{x}_0$  und beliebiges positives  $\varepsilon$  offen und beschränkt, aber nicht abgeschlossen.

- (b) Die Menge

$$K_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| \leq \varepsilon\}$$

ist kompakt. Es handelt sich hier um die abgeschlossene Kugel um den Mittelpunkt  $\vec{x}_0$  mit Radius  $\varepsilon$ .

- (c) Die Menge  $M = \mathbb{R}^n$  ist offen *und* abgeschlossen, aber nicht beschränkt (und deshalb auch nicht kompakt).
- (d) Die „gelochte Kugel mit Haut“  $K_\varepsilon(\vec{0}) \setminus \{\vec{0}\}$  ist nicht abgeschlossen, nicht offen, aber beschränkt.

**Satz 1.16 (Regeln für offene und abgeschlossene Mengen):** (a) Eine Menge  $M \subseteq \mathbb{R}^n$  ist genau dann offen, wenn ihre Komplementmenge  $\mathbb{R}^n \setminus M$  abgeschlossen ist.

- (b) Wenn zwei Mengen  $M_1$  und  $M_2$  offen sind, dann sind ihre Vereinigung  $M_1 \cup M_2$  und ihr Durchschnitt  $M_1 \cap M_2$  auch offen.
- (c) Wenn zwei Mengen  $M_1$  und  $M_2$  abgeschlossen sind, dann sind ihre Vereinigung  $M_1 \cup M_2$  und ihr Durchschnitt  $M_1 \cap M_2$  auch abgeschlossen.

**Satz 1.17 (Maximumprinzip):**

Wenn  $M \subseteq \mathbb{R}^n$  eine kompakte Menge ist und  $f: M \rightarrow \mathbb{R}$  eine stetige Funktion, so nimmt  $f$  auf  $M$  ihr Minimum und Maximum an, d.h. es gibt ein  $\vec{x}_* \in M$  und ein  $\vec{x}^* \in M$  mit

$$f(\vec{x}_*) \leq f(\vec{x}) \leq f(\vec{x}^*) \quad \text{für jedes } \vec{x} \in M.$$

## 2 Differentialrechnung in mehreren Variablen

### Bemerkung 2.1:

Die Einführung der Differentialrechnung in mehreren Variablen birgt ein **Problem**:

Und zwar sei  $M \subseteq \mathbb{R}^n$  gegeben, sei  $f: M \rightarrow \mathbb{R}$  eine Funktion, also  $f(\vec{x}) = f(x_1, x_2, \dots, x_n)$ . Diese Funktion  $f$  soll abgeleitet werden.

Folgendes funktioniert nicht:  $f'(\vec{x}_0) = \lim_{\vec{x} \rightarrow \vec{x}_0} \frac{f(\vec{x}) - f(\vec{x}_0)}{\vec{x} - \vec{x}_0}$ , denn im Nenner steht der Vektor  $\vec{x} - \vec{x}_0$ , und eine Division durch Vektoren gibt es nicht (und wird es auch nie geben).

Der anschauliche Hintergrund ist: das Steigungsverhalten von  $f$  im Punkt  $\vec{x}_0$  kann nicht durch eine einzelne Zahl beschrieben werden. Im  $\mathbb{R}^n$  gibt es (unendlich) viele Richtungen, und in jeder einzelnen Richtung hat die Funktion  $f$  ein gewisses Verhalten. All diese (womöglich unterschiedlichen) Verhaltensweisen müssen sich im zu definierenden Ableitungsbegriff widerspiegeln.

Ab jetzt entwickeln wir zwei Konzepte zur Lösung:

1. **totale Ableitung**
  - diese entspricht der bekannten eindimensionalen Ableitung;
  - es ist ein starkes Konzept, welches eine schöne Theorie ermöglicht;
  - leider ist der Begriff der totalen Ableitung zum konkreten Rechnen nicht direkt geeignet.
2. **partielle Ableitung**
  - es ist ein schwaches Konzept, das nur eine hässliche Theorie gestattet;
  - aber die Berechnung verläuft recht einfach wie in Mathematik I.

Und entscheidend ist natürlich

### 3. der Zusammenhang zwischen beiden Ableitungsbegriffen.

Wir beginnen mit der partiellen Ableitung und betrachten die totale Ableitung danach.

### Definition 2.2 (Richtungsableitung):

Sei  $M \subseteq \mathbb{R}^n$  eine offene Menge,  $f: M \rightarrow \mathbb{R}$  eine Funktion,  $\vec{x}_0 \in M$  ein Punkt, und sei  $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$  eine Richtung. Wenn der Grenzwert

$$(\partial_{\vec{v}}f)(\vec{x}_0) := \lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h}$$

existiert, dann heißt  $f$  im Punkt  $\vec{x}_0$  in Richtung  $\vec{v}$  differenzierbar. Die Zahl  $(\partial_{\vec{v}}f)(\vec{x}_0)$  heißt *Richtungsableitung von  $f$  im Punkt  $\vec{x}_0$  in Richtung  $\vec{v}$* .

Anschaulich bedeutet dies: in der Menge  $M$  befinden wir uns am Punkt  $\vec{x}_0$ . Wir laufen von dort in Richtung  $\vec{v}$ , und der Anstieg von  $f$ , den wir bei diesem Weg spüren, ist die Richtungsableitung von  $f$  im Punkt  $\vec{x}_0$  in Richtung  $\vec{v}$ . Siehe Bild 2.8.

**Definition 2.3 (Partielle Ableitung):** (a) Wenn  $\{\vec{e}_1, \dots, \vec{e}_n\}$  die Standardbasis des  $\mathbb{R}^n$  bezeichnet, dann nennen wir

$$\partial_j f(\vec{x}_0) := (\partial_{\vec{e}_j} f)(\vec{x}_0), \quad j = 1, 2, \dots, n,$$

die *partielle Ableitung von  $f$  in  $\vec{x}_0$  nach der  $j$ -ten Koordinate* (es ist also die Richtungsableitung in Richtung der Standardbasisvektoren). Wenn alle partiellen Ableitungen im Punkt  $\vec{x}_0$  existieren, dann nennt man  $f$  *partiell differenzierbar* in  $\vec{x}_0$ . Synonyme Schreibweisen sind

$$\partial_j f(\vec{x}_0) = \partial_{x_j} f(\vec{x}_0) = \frac{\partial f}{\partial x_j}(\vec{x}_0) = f_{x_j}(\vec{x}_0).$$

- (b) Wenn  $f: M \rightarrow \mathbb{R}$  in jedem Punkt  $\vec{x}_0 \in M$  partiell differenzierbar ist, und wenn die  $n$  Funktionen  $\vec{x}_0 \mapsto \partial_j f(\vec{x}_0)$  (wobei  $j = 1, \dots, n$ ) stetig in  $M$  sind, so heißt die Funktion  $f$  *auf  $M$  stetig partiell differenzierbar*.

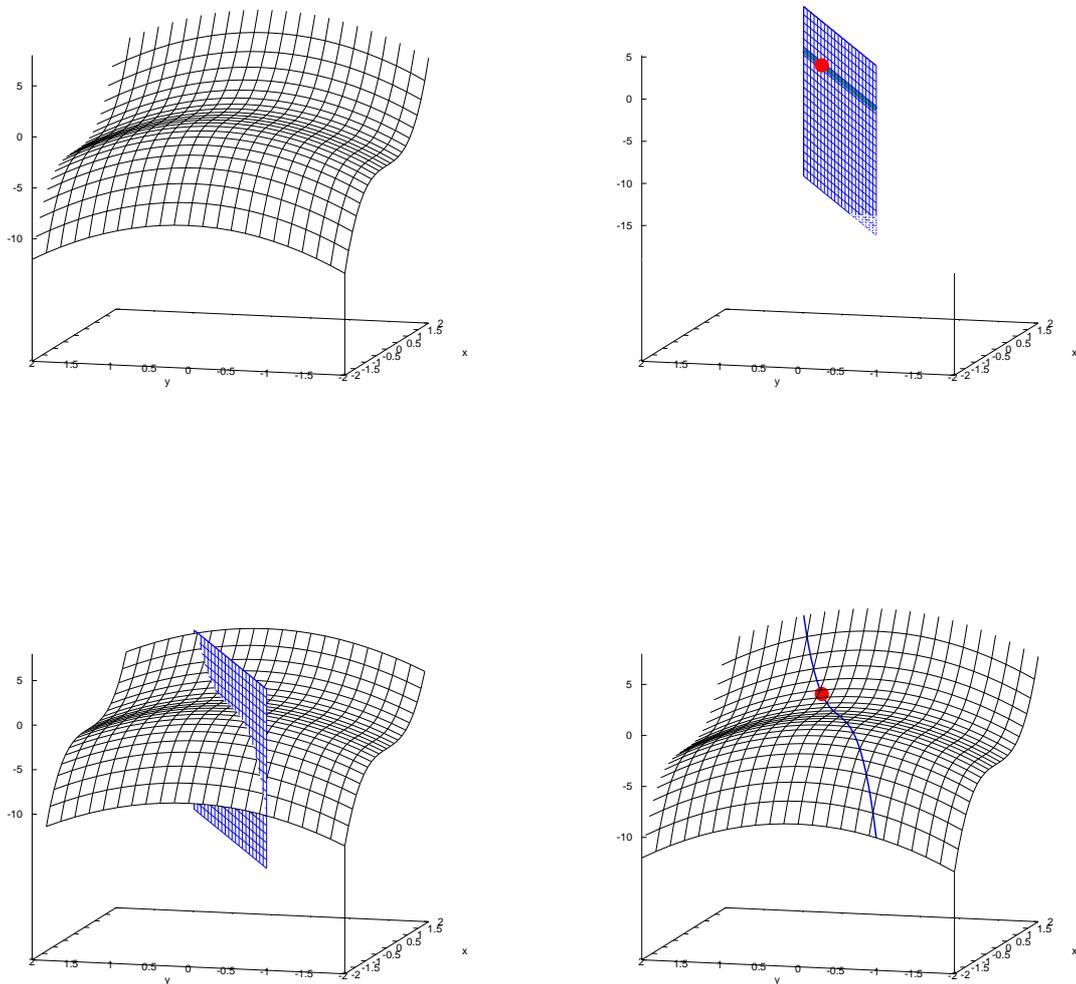


Abbildung 2.8: Links oben: die durch die Funktion  $z = f(x, y) = x^3 - y^2$  beschriebene Fläche. Wir wollen im Punkt  $[x_0, y_0]^T = [1, 0.5]^T$  die Richtungsableitung in Richtung  $\vec{v} = \frac{1}{\sqrt{5}}[2, 1]^T$  bestimmen. Rechts oben: die Gerade durch den roten Punkt  $[x_0, y_0]^T$  in Richtung  $\vec{v}$  ist eine grüne Gerade in der  $xy$ -Ebene. Der Vektor  $\vec{v}$  zeigt entlang der grünen Geraden „in das Papier hinein“. Wenn wir diese Gerade entlang der  $z$ -Achse parallel verschieben, entsteht eine blaue Ebene. Links unten: wir schneiden diese blaue Ebene mit der Fläche zur Funktion  $z = f(x, y) = x^3 - y^2$ . Rechts unten: als Schnittfigur entsteht eine blaue Kurve. Von dieser Kurve interessieren wir uns für das Steigungsverhalten am roten Punkt  $[1, 0.5, 0.75]^T$ . Beachte, dass der Vektor  $\vec{v}$  in der  $xy$ -Ebene nach „links hinten“ zeigt, also wird die Richtungsableitung positiv sein. Weil wir  $\vec{v}$  auf Länge Eins normiert haben, ist die Richtungsableitung von  $f$  gleich der schultypischen Ableitung, wenn wir die blaue Ebene als Papier-Ebene wählen. Bei anderen Längen von  $\vec{v}$  käme noch ein entsprechender Korrekturfaktor hinzu.

**Verfahren 2.4 (Bestimmung der partiellen Ableitungen):**

Wir wählen die Richtung  $\vec{v} = \vec{e}_1$ . Dann ist

$$\partial_1 f(x_1, x_2, \dots, x_n) = \lim_{h \rightarrow 0} \frac{f(\vec{x} + h\vec{e}_1) - f(\vec{x})}{h} = \lim_{h \rightarrow 0} \frac{f(x_1 + h, x_2, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{h}.$$

Das bedeutet: wir betrachten die restlichen Variablen  $x_2, \dots, x_n$  als festgefrorene Parameter, und lediglich  $x_1$  ist beweglich. Wir leiten diese Funktion dann „wie gewohnt“ nach der Variablen  $x_1$  ab.

Die Ableitungen nach den anderen Variablen  $x_2, x_3, \dots, x_n$  werden entsprechend gebildet.

**Beispiel 2.5:**

Wir betrachten die Funktion  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$  mit  $f(x, y, z) = xe^{xz+y^2}$ . Dann rechnen wir wie folgt:

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y, z) &= \partial_1 f(x, y, z) = e^{xz+y^2} + xe^{xz+y^2} \cdot z, \\ \frac{\partial f}{\partial y}(x, y, z) &= \partial_2 f(x, y, z) = xe^{xz+y^2} \cdot 2y, \\ \frac{\partial f}{\partial z}(x, y, z) &= \partial_3 f(x, y, z) = xe^{xz+y^2} \cdot x = x^2 e^{xz+y^2}. \end{aligned}$$

**Definition 2.6 (Gradient):**

Sei die Funktion  $f$  im Punkt  $\vec{x}_0$  partiell differenzierbar. Dann heißt der Ausdruck

$$\nabla f(\vec{x}_0) := [\partial_1 f(\vec{x}_0), \partial_2 f(\vec{x}_0), \dots, \partial_n f(\vec{x}_0)] = [\partial_1 f(\vec{x}_0) \quad \partial_2 f(\vec{x}_0) \quad \dots \quad \partial_n f(\vec{x}_0)]$$

Gradient von  $f$ . Das Symbol  $\nabla$  spricht man „nabla“.

**Bemerkung 2.7:** (a) Beachte: der Gradient ist ein Zeilenvektor, kein Spaltenvektor.

(b) Der Gradient besitzt folgende anschauliche Bedeutung: wenn wir innerhalb des  $\mathbb{R}^n$  an dem Punkt  $\vec{x}_0$  den Vektor  $\nabla f(\vec{x}_0)$  anheften, dann liegt dieser Vektor wie folgt:

- der Vektor  $\nabla f(\vec{x}_0)$  zeigt die Richtung an, in der der Graph von  $f$  am steilsten ansteigt,
- die Länge von  $\nabla f(\vec{x}_0)$ , also  $\|\nabla f(\vec{x}_0)\|$ , ist die Steigung dieses steilsten Anstiegs,
- der Vektor  $\nabla f(\vec{x}_0)$  steht senkrecht auf der Höhenlinie durch den Punkt  $\vec{x}_0$ .

Dazu ein Beispiel: wir wählen im  $\mathbb{R}^2$  die offene Einheitskreisscheibe  $M$ ,

$$M := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 < 1 \right\},$$

und dann betrachten wir die Funktion  $f: M \rightarrow \mathbb{R}$ , gegeben durch

$$f(x, y) := \sqrt{1 - x^2 - y^2}.$$

Anschaulich beschreibt diese Funktion eine obere Halbkugel. Dann haben wir den Gradienten

$$\begin{aligned} \nabla f(x, y) &= \left[ \frac{\partial f}{\partial x}(x, y), \quad \frac{\partial f}{\partial y}(x, y) \right] \\ &= \left[ \frac{-2x}{2\sqrt{1-x^2-y^2}}, \quad \frac{-2y}{2\sqrt{1-x^2-y^2}} \right] \\ &= \frac{-1}{\sqrt{1-x^2-y^2}} [x, \quad y]. \end{aligned}$$

Siehe Bild 2.9.

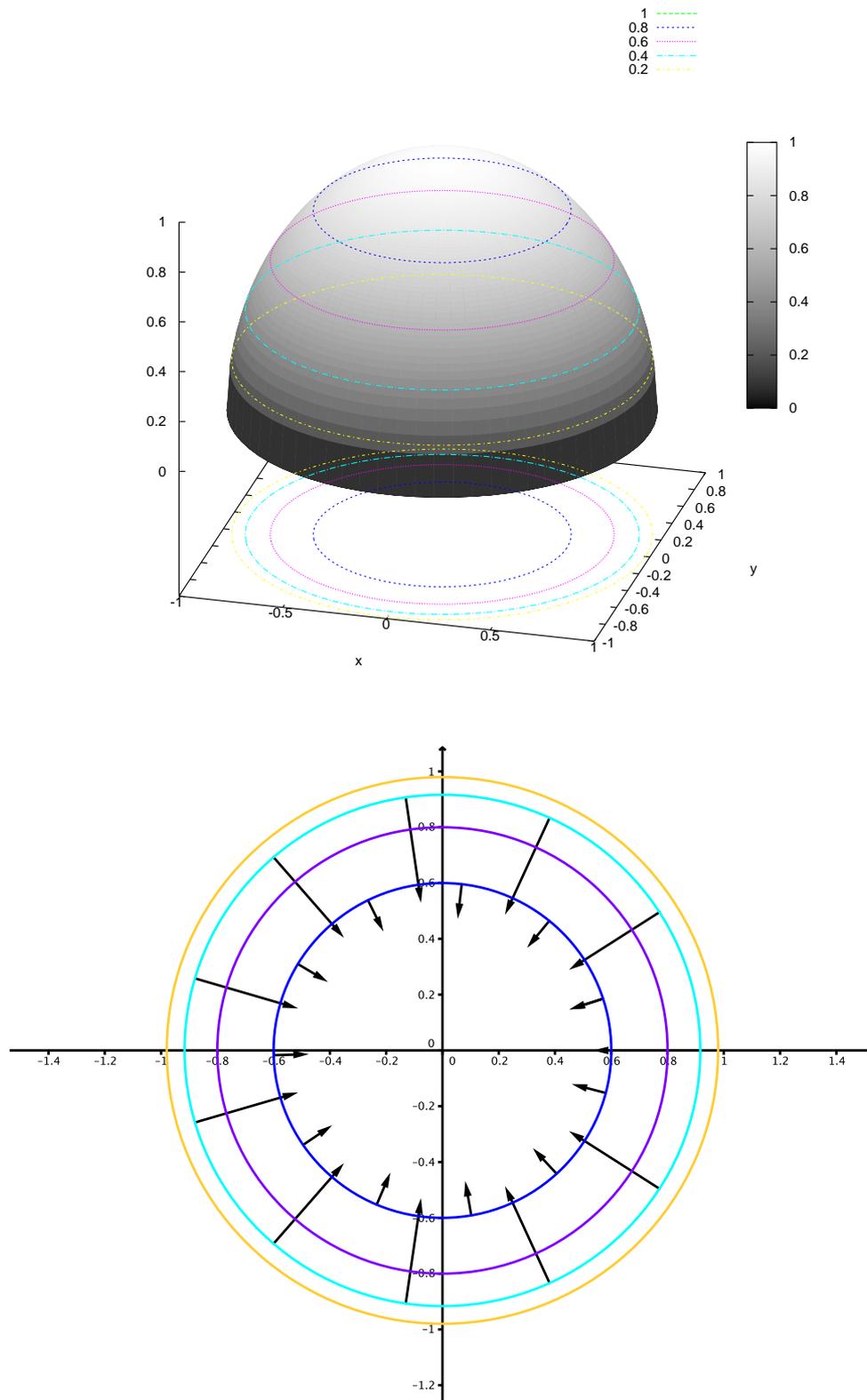


Abbildung 2.9: Oben: die skalare Funktion  $f(x, y) = \sqrt{1 - x^2 - y^2}$ . Die Höhenlinien sind zweifach eingezeichnet worden: einmal auf der Halbkugelfläche, und ein weiteres mal im Definitionsbereich  $M \subset \mathbb{R}^2$ . Unten sieht man die Höhenlinien in  $M$  und einige Gradientenvektoren. Die Gradientenvektoren sind um den Faktor 6 verkürzt eingezeichnet, damit die Graphik nicht überquillt. In der Nähe des Äquators ist die Halbkugelfläche steil, und deshalb sind dort die Gradientenvektoren lang. In der Nähe des Nordpols ist die Halbkugelfläche vergleichsweise „unsteil“, und deshalb sind dort die Gradientenvektoren kurz.

**Definition 2.8 (Partielle Ableitungen von Vektorfeldern):**

Sei  $M \subseteq \mathbb{R}^n$  und  $\vec{x}_0 \in M$ . Ein Vektorfeld  $f = [f_1, \dots, f_m]^\top : M \rightarrow \mathbb{R}^m$  heißt *partiell differenzierbar* in  $\vec{x}_0$ , falls die Komponenten  $f_1, f_2, \dots, f_m$  dies sind. Wir schreiben dann

$$J_f(\vec{x}_0) := \begin{bmatrix} \nabla f_1(\vec{x}_0) \\ \nabla f_2(\vec{x}_0) \\ \vdots \\ \nabla f_m(\vec{x}_0) \end{bmatrix} = \begin{bmatrix} \partial_1 f_1(\vec{x}_0) & \partial_2 f_1(\vec{x}_0) & \dots & \partial_n f_1(\vec{x}_0) \\ \partial_1 f_2(\vec{x}_0) & \partial_2 f_2(\vec{x}_0) & \dots & \partial_n f_2(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m(\vec{x}_0) & \partial_2 f_m(\vec{x}_0) & \dots & \partial_n f_m(\vec{x}_0) \end{bmatrix} \in \mathbb{R}^{m \times n}$$

und nennen diese Matrix *Jacobi-Matrix von  $f$  in  $\vec{x}_0$* . Wir führen auch partielle Ableitungen des Vektorfelds  $f$  ein:

$$\partial_j f(\vec{x}_0) = \begin{bmatrix} \partial_j f_1(\vec{x}_0) \\ \partial_j f_2(\vec{x}_0) \\ \vdots \\ \partial_j f_m(\vec{x}_0) \end{bmatrix}, \quad j = 1, \dots, n.$$

**Beispiel 2.9:**

Wir betrachten noch einmal dasjenige  $f$ , das von Polarkoordinaten zu den üblichen kartesischen Koordinaten umrechnet (siehe Beispiel 1.4 (d)). Dabei ist  $f : (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2$  mit

$$f(r, \varphi) = \begin{bmatrix} r \cos \varphi \\ r \sin \varphi \end{bmatrix} = \begin{bmatrix} f_1(r, \varphi) \\ f_2(r, \varphi) \end{bmatrix}.$$

Die Jacobi-Matrix berechnen wir dann so:

$$J_f(r, \varphi) = \begin{bmatrix} \partial_r f_1(r, \varphi) & \partial_\varphi f_1(r, \varphi) \\ \partial_r f_2(r, \varphi) & \partial_\varphi f_2(r, \varphi) \end{bmatrix} = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix}.$$

**Satz 2.10 (Kettenregel):**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$  partiell differenzierbar. Dann ist auch die zusammengesetzte Funktion

$$h := g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$$

partiell differenzierbar, und es ist

$$J_h(\vec{x}_0) = J_g(f(\vec{x}_0)) \cdot J_f(\vec{x}_0).$$

Man beobachte, dass die Matrix-Matrix-Multiplikation auf der rechten Seite tatsächlich durchführbar ist, weil die Abmessungen der beiden Matrizen wirklich kompatibel sind !

**Beispiel 2.11:**

Wir betrachten eine „äußere“ Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , gegeben durch

$$g(p, q) = \begin{bmatrix} e^p \cos(q) \\ e^p \sin(q) \end{bmatrix},$$

und eine „innere“ Funktion  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , gegeben durch

$$f(x, y, z) = \begin{bmatrix} x + y \\ xz \end{bmatrix}.$$

Dann ergibt sich die zusammengesetzte („verkettete“) Funktion

$$(g \circ f)(x, y, z) = g(f(x, y, z)) = g(x + y, xz) = \begin{bmatrix} e^{x+y} \cos(xz) \\ e^{x+y} \sin(xz) \end{bmatrix}.$$

Die Jacobi-Matrizen der drei Funktionen  $g \circ f$ ,  $g$  und  $f$  stehen zueinander in Beziehung wie folgt:

$$J_{g \circ f}(x, y, z) = J_g(f(x, y, z)) \cdot J_f(x, y, z).$$

Wir rechnen die linke Seite aus, indem wir die beiden Faktoren auf der rechten Seite ermitteln: Zunächst ist

$$J_g(p, q) = \begin{bmatrix} e^p \cos(q) & -e^p \sin(q) \\ e^p \sin(q) & e^p \cos(q) \end{bmatrix}.$$

Für  $(p, q)$  setzen wir jetzt  $f(x, y, z)$  ein:

$$J_g(f(x, y, z)) = J_g(x + y, xz) = \begin{bmatrix} e^{x+y} \cos(xz) & -e^{x+y} \sin(xz) \\ e^{x+y} \sin(xz) & e^{x+y} \cos(xz) \end{bmatrix} = e^{x+y} \begin{bmatrix} \cos(xz) & -\sin(xz) \\ \sin(xz) & \cos(xz) \end{bmatrix}.$$

Damit erhalten wir also, wie angekündigt:

$$\begin{aligned} J_{g \circ f}(x, y, z) &= e^{x+y} \begin{bmatrix} \cos(xz) & -\sin(xz) \\ \sin(xz) & \cos(xz) \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ z & 0 & x \end{bmatrix} \\ &= e^{x+y} \begin{bmatrix} \cos(xz) - z \sin(xz) & \cos(xz) & -x \sin(xz) \\ \sin(xz) + z \cos(xz) & \sin(xz) & x \cos(xz) \end{bmatrix}. \end{aligned}$$

**Definition 2.12 (Partielle Ableitungen zweiter Ordnung):**

Sei  $M \subseteq \mathbb{R}^n$ , und sei  $f: M \rightarrow \mathbb{R}$  stetig partiell differenzierbar auf  $M$ . Dann setzen wir für  $\vec{x}_0 \in M$

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) := \frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) (\vec{x}_0).$$

Die folgenden alternativen Schreibweisen sind gebräuchlich:

$$\partial_i \partial_j f(\vec{x}_0), \quad \partial_{x_i} \partial_{x_j} f(\vec{x}_0), \quad f_{x_i x_j}(\vec{x}_0).$$

**Bemerkung 2.13:**

Analog führt man höhere Ableitungen (und deren Schreibweisenvarianten) ein, zum Beispiel

$$\partial_1 \partial_3 \partial_1 f, \quad \frac{\partial^3 f}{\partial x_1 \partial x_1^2}, \quad f_{xzy}, \quad f_{xxxxxy}, \quad \dots$$

Die partiellen Ableitungen zweiter Ordnung ordnen wir jetzt etwas systematischer an:

**Definition 2.14 (Hesse-Matrix):**

Wenn die Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal im Punkt  $\vec{x}_0$  differenzierbar ist, dann nennen wir die Matrix

$$H_f(\vec{x}_0) := \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\vec{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\vec{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\vec{x}_0) \end{bmatrix}$$

Hesse-Matrix von  $f$  in  $\vec{x}_0$ .

**Beispiel 2.15:**

Wir betrachten die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , gegeben durch die Formel  $f(x, y) = x^3 y + x e^y$ . Die partiellen Ableitungen erster Ordnung bilden den Gradienten von  $f$ :

$$\nabla f(x, y) = [f_x(x, y), f_y(x, y)] = [3x^2 y + e^y, x^3 + x e^y].$$

Die vier partiellen Ableitungen zweiter Ordnung stehen in der Hesse-Matrix:

$$H_f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix} = \begin{bmatrix} 6xy & 3x^2 + e^y \\ 3x^2 + e^y & x e^y \end{bmatrix}.$$

Und dann gibt es noch die Ableitungen dritter Ordnung (insgesamt 8 Stück):

$$\begin{aligned} f_{xxx}(x, y) &= 6y, & f_{xxy}(x, y) &= 6x, & f_{xyx}(x, y) &= 6x, & f_{yxx}(x, y) &= 6x, \\ f_{xyy}(x, y) &= e^y, & f_{yyx}(x, y) &= e^y, & f_{xyy}(x, y) &= e^y, & f_{xyy}(x, y) &= e^y, \\ f_{yyy}(x, y) &= x e^y. \end{aligned}$$

Wir beobachten, dass die Hesse-Matrix symmetrisch ist. Und wir beobachten auch, dass einige der Ableitungen dritter Ordnung gleich sind. Das ist kein Zufall:

**Satz 2.16 (Satz von Schwarz):**

Sei  $M \subseteq \mathbb{R}^d$  offen, und die Funktion  $f: M \rightarrow \mathbb{R}$  sei  $k$  mal differenzierbar, und *alle Ableitungen seien stetig*. Dann sind alle partiellen Ableitungen bis zur Ordnung  $k$  vertauschbar (also  $\partial_1 \partial_2 f = \partial_2 \partial_1 f$  usw.).

In der Einleitung wurde bereits erwähnt, dass die partiellen Ableitungen zwar recht einfach ausgerechnet werden können, aber nur eine hässliche Theorie zulassen. Jetzt kommen wir zu diesen unangenehmen Eigenschaften der partiellen Ableitungen, wobei wir die schreckliche Funktion aus Beispiel 1.12 (c) noch einmal aufgreifen. Siehe die Bilder 1.6 und 1.7 (links) für bildliche Darstellungen dieser Funktion, die im Nullpunkt eben nicht stetig ist.

**Beispiel 2.17:** (a) Wir betrachten die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , die gegeben ist durch

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & : \text{wenn } (x, y) \neq (0, 0), \\ 0 & : \text{wenn } (x, y) = (0, 0). \end{cases}$$

Für jeden Punkt  $(x, y) \neq (0, 0)$  gilt

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^2 + y^2) - 2x^2 y}{(x^2 + y^2)^2} = \frac{y^3 - x^2 y}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = (\text{analog}) = \frac{x^3 - 2xy^2}{(x^2 + y^2)^2}.$$

Im Ursprung gilt (gemäß Definition der partiellen Ableitungen)

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = (\text{analog}) = 0.$$

Das bedeutet: diese Funktion  $f$  ist im Ursprung partiell differenzierbar (denn dort existieren die partiellen Ableitungen tatsächlich), aber im Nullpunkt unstetig. Das ist seltsam, denn eigentlich würde man sich wünschen, dass Differenzierbarkeit eine stärkere Eigenschaft ist als Stetigkeit.

(b) Noch wilder ist folgende Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ :

$$f(x, y) = \begin{cases} 0 & : \text{wenn } 0 < x^3 < y < x^2, \\ 1 & : \text{sonst.} \end{cases}$$

Bei dieser Funktion existieren im Ursprung sämtliche Richtungsableitungen, und alle haben den Wert Null. Trotzdem ist die Funktion  $f$  nicht stetig im Ursprung. Siehe Bild 2.10.

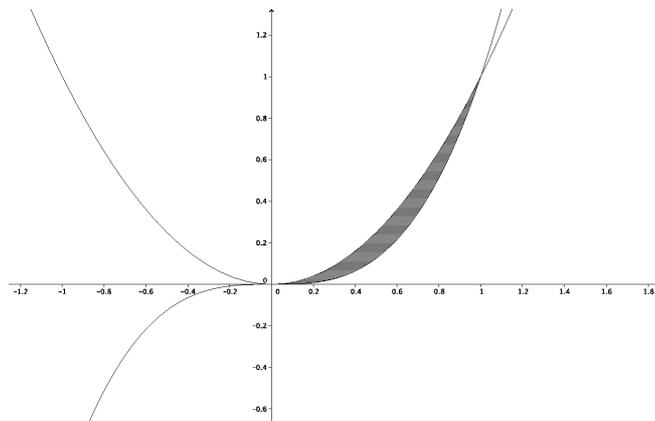


Abbildung 2.10: Die Funktion aus Beispiel 2.17 (b) hat den Wert 0 im schraffierten Bereich, und überall sonst (insbesondere im Nullpunkt) hat sie den Wert 1.

Die Angelegenheit wird viel schöner, wenn wir zum Konzept der **Totalen Differenzierbarkeit** übergehen.

**Erinnerung:** In der Vorlesung Mathematik I haben wir gelernt: eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  ist genau dann differenzierbar im Punkt  $x_0$ , wenn man die Werte  $f(x)$  für  $x$  aus einer Umgebung von  $x_0$  wie folgt in drei Anteile zerlegen kann:

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + r(x),$$

siehe auch Bild 2.11, wobei der Restsummand  $r(x)$  folgende Bedingung erfüllt:

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

Dies bedeutet, dass  $r(x)$  für  $x \rightarrow x_0$  schneller als  $x - x_0$  nach Null strebt. Wir können die angegebene Bedingung auch umformulieren zu

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

Im Nenner steht jetzt der Abstand der Zahlen  $x$  und  $x_0$  auf der Zahlengeraden. Der Abstand ist eine reelle Zahl, und durch Zahlen ( $\neq 0$ ) darf man dividieren. Daraus gewinnen wir eine Idee, wie wir einen Differenzierbarkeitsbegriff im  $\mathbb{R}^n$  definieren könnten, siehe Definition 2.18.

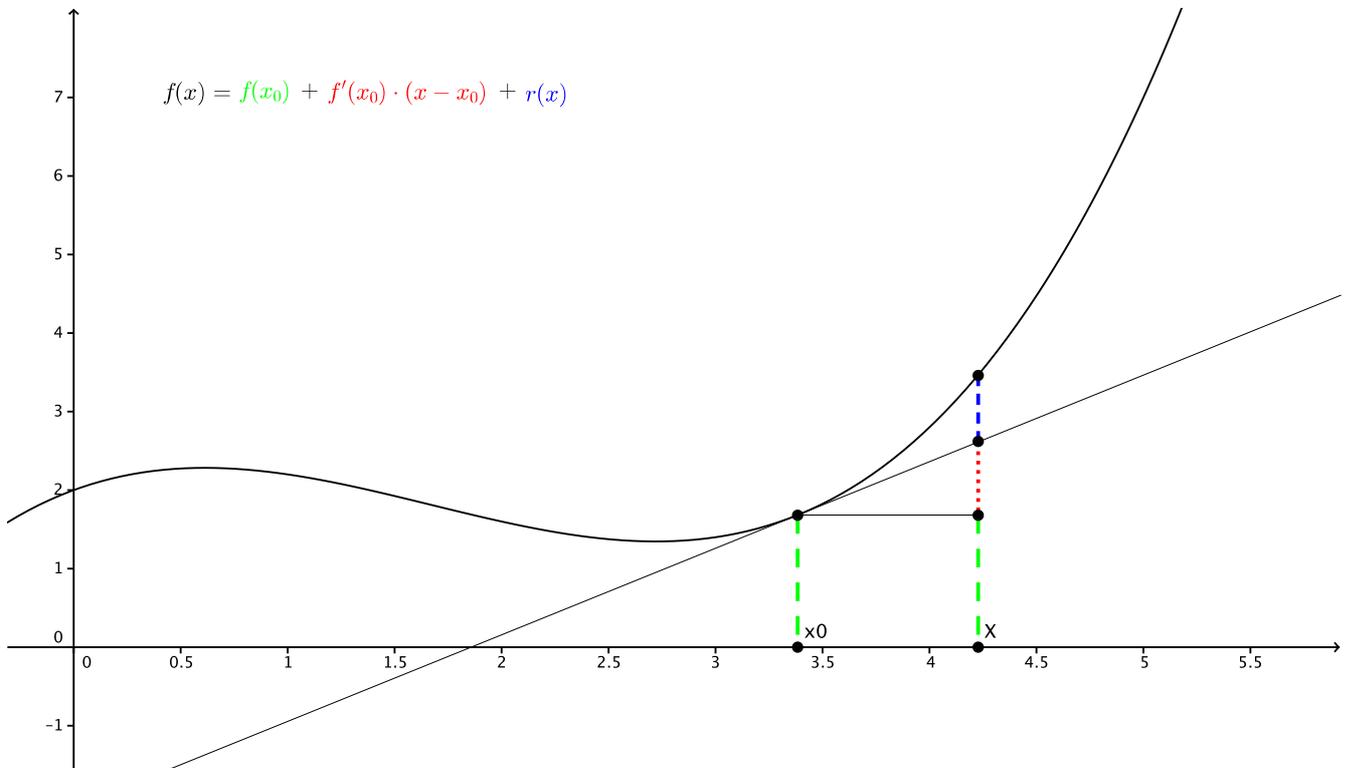


Abbildung 2.11: Wir zerlegen den Wert  $f(x)$  in drei Teile.

**Definition 2.18 (Totale Differenzierbarkeit):**

Sei  $M \subseteq \mathbb{R}^n$  eine offene Menge,  $\vec{x}_0 \in M$ . Eine Funktion  $f: M \rightarrow \mathbb{R}^m$  heißt *total differenzierbar* in  $\vec{x}_0$ , wenn es eine Matrix  $A \in \mathbb{R}^{m \times n}$  und eine Funktion  $r: M \rightarrow \mathbb{R}^m$  gibt mit der Eigenschaft

$$f(\vec{x}) = f(\vec{x}_0) + A \cdot (\vec{x} - \vec{x}_0) + r(\vec{x}), \quad \vec{x} \in M,$$

wobei der Summand  $r(\vec{x})$  folgende Bedingung erfüllt:

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{r(\vec{x})}{\|\vec{x} - \vec{x}_0\|} = \vec{0}.$$

Die Matrix  $A$  heißt *totale Ableitung von  $f$  in  $\vec{x}_0$* . Wir schreiben  $Df(\vec{x}_0) := A$ .

Man würde sich wünschen, dass vernünftigerweise die Differenzierbarkeit eine stärkere Eigenschaft ist als die Stetigkeit. Bei der totalen Differenzierbarkeit geht dieser Wunsch in Erfüllung:

**Satz 2.19:**

Wenn eine Funktion  $f$  im Punkt  $\vec{x}_0$  total differenzierbar ist, dann ist  $f$  in diesem Punkt stetig.

*Beweis.* Wir wissen, dass wir den Wert  $f(\vec{x})$  in die drei Anteile zerlegen können. Wir wollen zeigen, dass die Funktion  $f$  im Punkt  $\vec{x}_0$  stetig ist, also dass  $\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = f(\vec{x}_0)$  gilt. Dazu rechnen wir den Limes einfach aus:

$$\begin{aligned}
 \lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) & \quad \left| \text{zerlege } f(\vec{x}) \text{ in die drei Teile} \right. \\
 &= \lim_{\vec{x} \rightarrow \vec{x}_0} \left( f(\vec{x}_0) + A \cdot (\vec{x} - \vec{x}_0) + r(\vec{x}) \right) \quad \left| \text{ziehe den Limes auf jeden Summanden} \right. \\
 &= \lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}_0) + \lim_{\vec{x} \rightarrow \vec{x}_0} A \cdot (\vec{x} - \vec{x}_0) + \lim_{\vec{x} \rightarrow \vec{x}_0} r(\vec{x}) \quad \left| \text{vereinfache ersten Summanden} \right. \\
 &= f(\vec{x}_0) + \lim_{\vec{x} \rightarrow \vec{x}_0} A \cdot (\vec{x} - \vec{x}_0) + \lim_{\vec{x} \rightarrow \vec{x}_0} r(\vec{x}) \quad \left| \text{lineare Abbildungen sind stetig (Bsp. 1.12 (a))} \right. \\
 &= f(\vec{x}_0) + A \cdot \vec{0} + \lim_{\vec{x} \rightarrow \vec{x}_0} r(\vec{x}) \quad \left| \text{erweitere clever} \right. \\
 &= f(\vec{x}_0) + \lim_{\vec{x} \rightarrow \vec{x}_0} \left( \frac{r(\vec{x})}{\|\vec{x} - \vec{x}_0\|} \cdot \|\vec{x} - \vec{x}_0\| \right) \quad \left| \text{ziehe Limes auf jeden Faktor} \right. \\
 &= f(\vec{x}_0) + \left( \lim_{\vec{x} \rightarrow \vec{x}_0} \frac{r(\vec{x})}{\|\vec{x} - \vec{x}_0\|} \right) \cdot \lim_{\vec{x} \rightarrow \vec{x}_0} \|\vec{x} - \vec{x}_0\| \quad \left| \text{nutze Wissen über } r(x) \right. \\
 &= f(\vec{x}_0) + \vec{0} \cdot 0 = f(\vec{x}_0).
 \end{aligned}$$

Das wollten wir haben. □

**Satz 2.20 (Zusammenhang zwischen partiellen und totalen Ableitungen):**

Sei  $M \subseteq \mathbb{R}^n$  offen,  $\vec{x}_0 \in M$  und  $f: M \rightarrow \mathbb{R}^m$ .

- (a) Wenn  $f$  total differenzierbar in  $\vec{x}_0$  ist, dann existieren alle Richtungsableitungen  $\partial_{\vec{v}} f(\vec{x}_0)$  im Punkt  $\vec{x}_0$ , für alle  $\vec{v} \in \mathbb{R}^n \setminus \{0\}$ , und es gilt

$$\partial_{\vec{v}} f(\vec{x}_0) = Df(\vec{x}_0) \cdot \vec{v},$$

wobei rechts ein Produkt von Matrix und Vektor steht.

Zusätzlich gilt: die Ableitung  $Df(\vec{x}_0)$  ist gleich der Jacobi-Matrix  $J_f(\vec{x}_0)$  (bzw. gleich dem Gradienten von  $f$ , wenn  $m = 1$ ).

- (b) Wenn  $f$  partiell differenzierbar ist, dann braucht  $f$  nicht total differenzierbar sein. Jedoch gilt: wenn die partiellen Ableitungen stetig sind (also wenn die Funktion  $f$  stetig partiell differenzierbar ist), dann ist die Funktion  $f$  total differenzierbar.

*Beweis.* (a) Weil  $f$  total differenzierbar ist, können wir  $f(\vec{x})$  in die drei Teile zerlegen. Das benutzen wir jetzt, um  $\partial_{\vec{v}} f(\vec{x}_0)$  auszurechnen. Sei also eine beliebige Richtung  $\vec{v}$  gegeben. Dann haben wir:

$$\begin{aligned}
 (\partial_{\vec{v}} f)(\vec{x}_0) &= \lim_{h \rightarrow 0} \frac{1}{h} \left( f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0) \right) \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \left( \cancel{f(\vec{x}_0)} + Df(\vec{x}_0) \cdot h\vec{v} + r(\vec{x}_0 + h\vec{v}) - \cancel{f(\vec{x}_0)} \right) \\
 &= \lim_{h \rightarrow 0} \frac{Df(\vec{x}_0) \cdot h\vec{v}}{h} + \lim_{h \rightarrow 0} \frac{r(\vec{x}_0 + h\vec{v})}{h} \\
 &= Df(\vec{x}_0) \cdot \vec{v} + \lim_{h \rightarrow 0} \left( \frac{r(\vec{x}_0 + h\vec{v})}{\|(\vec{x}_0 + h\vec{v}) - \vec{x}_0\|} \cdot \|\vec{v}\| \right) \\
 &= Df(\vec{x}_0) \cdot \vec{v} + \left( \lim_{h \rightarrow 0} \frac{r(\vec{x}_0 + h\vec{v})}{\|(\vec{x}_0 + h\vec{v}) - \vec{x}_0\|} \right) \cdot \|\vec{v}\| \\
 &= Df(\vec{x}_0) \cdot \vec{v} + \vec{0} \cdot \|\vec{v}\| = Df(\vec{x}_0) \cdot \vec{v}.
 \end{aligned}$$

Das wollten wir zeigen.

Insbesondere gilt für  $j = 1, \dots, n$ , dass

$$\partial_j f(\vec{x}_0) = \partial_{\vec{e}_j} f(\vec{x}_0) = Df(\vec{x}_0) \cdot \vec{e}_j.$$

Gemäß Definition 2.8 steht links die  $j$ -te Spalte der Jacobi-Matrix  $J_f(\vec{x}_0)$ , und gemäß der Merkregel auf Seite 31 steht rechts die  $j$ -te Spalte der Matrix  $Df(\vec{x}_0)$ . Diese Argumentation klappt für jedes  $j = 1, \dots, n$ . Also ist  $Df(\vec{x}_0) = J_f(\vec{x}_0)$ , wenn  $f$  in  $\vec{x}_0$  total differenzierbar ist.

- (b) Die Funktion  $f$  aus Beispiel 2.17 (b) ist partiell differenzierbar im Nullpunkt, aber sie kann dort nicht total differenzierbar sein, denn dann wäre sie ja stetig (wegen Satz 2.19). Auf den Beweis der zweiten Behauptung verzichten wir.

□

Insgesamt bekommen wir folgendes **Verfahren zur Bestimmung der totalen Ableitung** von  $f$ :

- wir bestimmen die partiellen Ableitungen  $\frac{\partial f_i}{\partial x_k}$  mit Rechenmethoden ähnlich wie in der Schule (indem wir die anderen Variablen festhalten und lediglich  $x_k$  laufen lassen),
- all diese partiellen Ableitungen bilden die Einträge der Jacobi-Matrix  $J_f(\vec{x}_0)$ ,
- wir prüfen nach, ob diese partiellen Ableitungen  $\frac{\partial f_i}{\partial x_k}$  stetig sind,
- wenn ja: dann ist die totale Ableitung gleich der Jacobi-Matrix  $J_f(\vec{x}_0)$ .

Die Glattheitseigenschaften von Funktionen stehen wie folgt in Beziehung zueinander:

$$\begin{array}{ccccc} \text{stetig partiell differenzierbar} & \implies & \text{total differenzierbar} & \implies & \text{stetig} \\ \downarrow & & \downarrow & & \\ \text{partiell differenzierbar} & \iff & \text{alle Richtungsableitungen existieren} & & \end{array}$$

Alle Umkehrungen sind falsch !

### Bemerkung 2.21:

Eine totale Ableitung nach einer einzelnen Variablen (z.B. nach einer Zeitvariablen  $t$ ) wird oft als  $\frac{d}{dt}$  geschrieben. Wir müssen sorgfältig unterscheiden:

$$\begin{array}{ll} \frac{\partial}{\partial t} & \text{gibt uns die partielle Ableitung nach der **Koordinate } t, \\ \frac{d}{dt} & \text{gibt uns die totale Ableitung nach der **Variablen } t. \end{array}****$$

Als Beispiel betrachten wir die Funktion  $f$  mit der Gleichung  $f(x, y) = x^3 y + x e^y$  und suchen  $\frac{d}{dt} f(t^2, t^3)$ . Dann ist

$$f(t^2, t^3) = (t^2)^3 t^3 + t^2 e^{(t^3)} = t^9 + t^2 e^{(t^3)},$$

also bekommen wir das Ergebnis

$$\frac{d}{dt} f(t^2, t^3) = 9t^8 + 2te^{(t^3)} + t^2 e^{(t^3)} \cdot 3t^2 = 9t^8 + e^{(t^3)}(2t + 3t^4).$$

Anstatt  $e^{(t^3)}$  schreibt man gewöhnlich  $e^{t^3}$ , weil der Exponent  $t^3$  sowieso Rechenpriorität hat.

### 3 Anwendungen

Unser **Ziel** ist der Satz von Taylor für Funktionen von mehreren Variablen.

Wir **erinnern** an die Vorlesung Mathematik I: dazu betrachten wir eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  und eine Entwicklungsstelle  $x_0 \in \mathbb{R}$ . Dann haben wir die Zerlegung

$$f(x) = T_k(x) + R_k(x),$$

wobei  $T_k$  das *Taylorpolynom* vom Grade  $k$  ist,

$$T_k(x) = \sum_{j=0}^k \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j,$$

und für das *Restglied*  $R_k(x)$  gibt es die Formel

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!} (x - x_0)^{k+1},$$

wobei von der Zahl  $\xi$  lediglich bekannt ist, dass sie irgendwo zwischen  $x$  und  $x_0$  liegt.

In den praktischen Anwendungen möchte man das Restglied  $R_k$  oft vernachlässigen (weil man ein Polynom  $T_k$  im Allgemeinen besser versteht als eine allgemeine Funktion  $f$ ), was typischerweise nur dann angemessen ist, wenn  $|R_k(x)|$  „klein genug“ ist. Das ist meist dann durchführbar, wenn  $x$  „nahe genug“ am  $x_0$  liegt, weil dann  $|(x - x_0)^{k+1}|$  klein ist, und wenn zusätzlich  $|f^{(k+1)}(\xi)|$  nicht riesig ist.

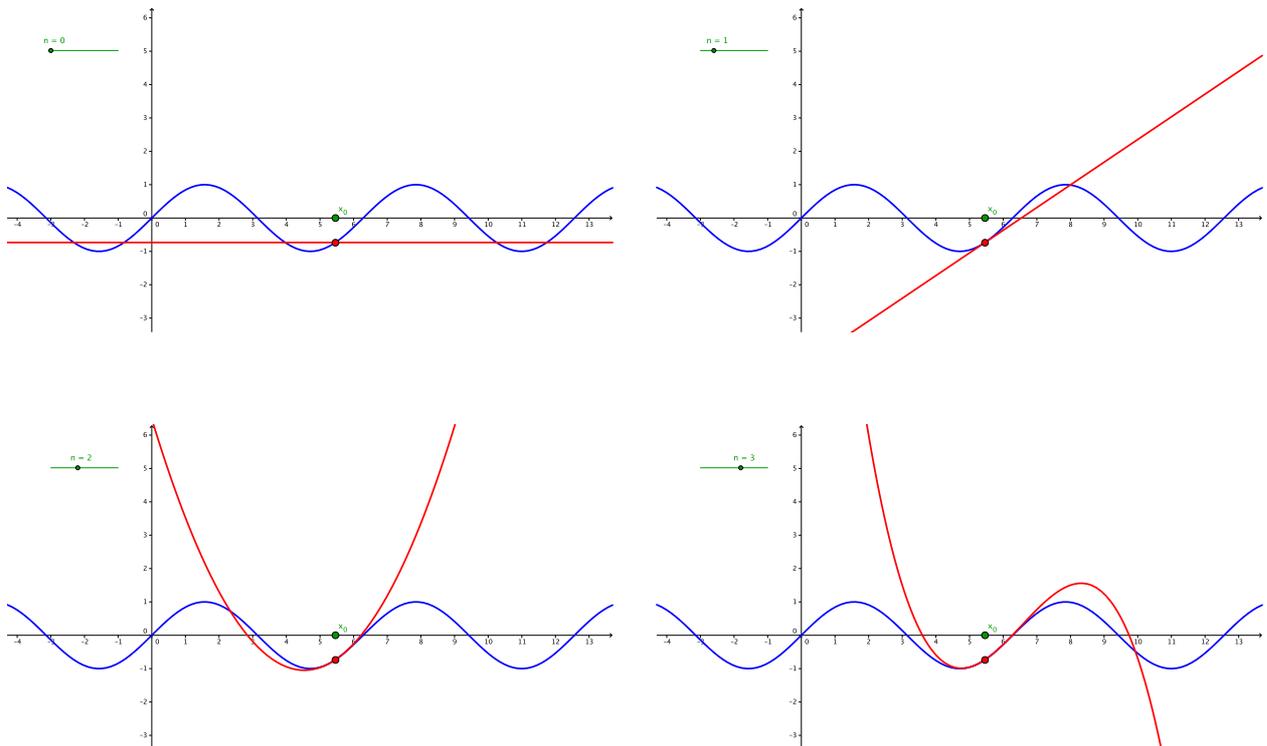


Abbildung 3.12: Die Taylorpolynome  $T_0, T_1, T_2, T_3$  (rot) für die Sinusfunktion. Wir beobachten, dass die Näherung  $T_k(x) \approx \sin(x)$  tauglich ist für  $x \approx x_0$ . Wenn  $x \approx x_0$ , dann wird diese Näherung mit steigendem  $k$  typischerweise besser, aber dann steigt auch der Aufwand für die Auswertung des Polynoms  $T_k$ , weil dieses dann mehr Summanden besitzt. Falls allerdings der Abstand  $|x - x_0|$  zu groß wird, dann wird die Näherung  $T_k(x) \approx \sin(x)$  offenkundig untauglich.

Jetzt gehen wir mit dem  $x$  vom  $\mathbb{R}^1$  in den  $\mathbb{R}^n$  und betrachten die Taylorentwicklung mit quadratischem Restglied:

**Satz 3.1 (Taylorscher Satz mit Taylorpolynom ersten Grades):**

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig partiell differenzierbar und  $\vec{x}_0 \in \mathbb{R}^n$ . Dann gibt es für jedes  $\vec{x} \in \mathbb{R}^n$  ein  $\vec{\xi}$  auf der Verbindungsstrecke von  $\vec{x}$  nach  $\vec{x}_0$  mit

$$f(\vec{x}) = \underbrace{f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)}_{\text{Taylorpolynom ersten Grades}} + \underbrace{\frac{1}{2}(\vec{x} - \vec{x}_0)^\top \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0)}_{\text{Restglied}}.$$

Beachte, dass  $\nabla f(\vec{x}_0)$  ein Zeilenvektor ist und  $(\vec{x} - \vec{x}_0)$  ein Spaltenvektor, sodass im Produkt  $\nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$  die Abmessungen kompatibel sind. Entsprechendes gilt für das Drei-Faktor-Produkt im Restglied.

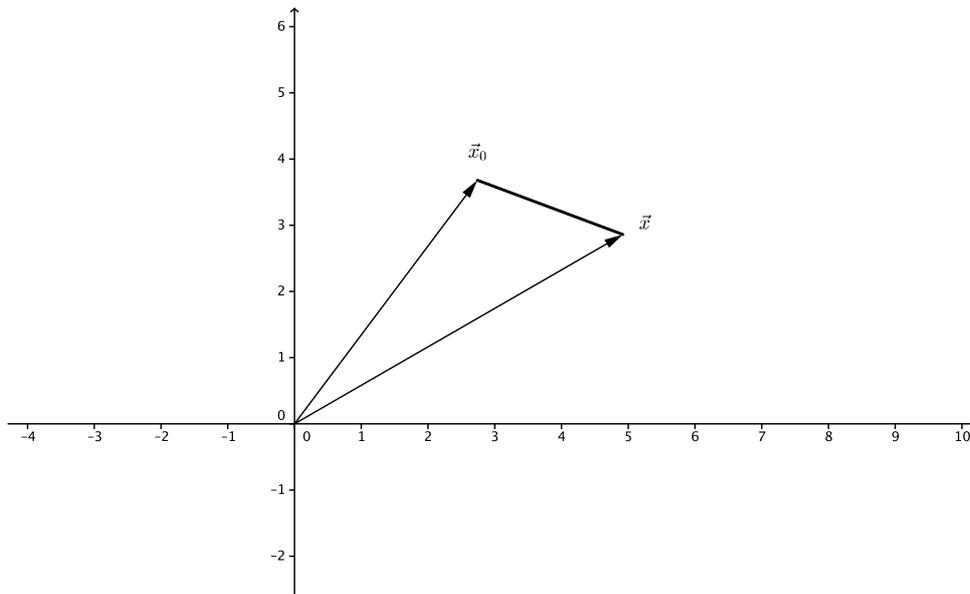


Abbildung 3.13: Die Verbindungsstrecke zwischen  $\vec{x}$  und  $\vec{x}_0$ .

**Bemerkung 3.2 (Verbindungsstrecke):**

Die Verbindungsstrecke  $S$  zwischen  $\vec{x}$  und  $\vec{x}_0$  kann man mathematisch beschreiben wie folgt:

$$S = \{\vec{x}_0 + \lambda(\vec{x} - \vec{x}_0) : 0 \leq \lambda \leq 1\} = \{\lambda\vec{x} + (1 - \lambda)\vec{x}_0 : 0 \leq \lambda \leq 1\}.$$

Irgendwo auf der Strecke  $S$  liegt der Punkt  $\vec{\xi}$  (mehr wissen wir vom  $\vec{\xi}$  nicht).

**Bemerkung 3.3 (Taylorpolynom zweiten Grades):**

Man kann auch quadratische Taylorpolynome betrachten, die dann aussehen wie folgt:

$$f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^\top \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0),$$

und das dazugehörige Restglied wäre dann „kubisch in Bezug auf die Differenz  $\vec{x} - \vec{x}_0$ “. Allerdings ist die Formel für das Restglied dann sehr kompliziert, weshalb wir darauf verzichten, es hinzuschreiben.

Wir wollen exemplarisch für den Fall  $n = 2$  (also für zwei Variablen  $x$  und  $y$ ) das Taylorpolynom zweiten Grades ermitteln. Wir haben also

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \vec{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

und das Taylorpolynom ergibt sich dann zu

$$\begin{aligned}
 & f(\vec{x}_0) + \begin{bmatrix} \frac{\partial f}{\partial x}(\vec{x}_0) & \frac{\partial f}{\partial y}(\vec{x}_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) & \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\vec{x}_0) & \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\
 &= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(\vec{x}_0) \cdot (y - y_0) \\
 &\quad + \frac{1}{2} \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot (y - y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \cdot (y - y_0) \end{bmatrix} \\
 &= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(\vec{x}_0) \cdot (y - y_0) \\
 &\quad + \frac{1}{2} \left( \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \cdot (x - x_0)^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot (x - x_0) \cdot (y - y_0) + \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \cdot (y - y_0)^2 \right) \\
 &= \frac{f(\vec{x}_0)}{0!0!} \cdot (x - x_0)^0 \cdot (y - y_0)^0 + \frac{\frac{\partial f}{\partial x}(\vec{x}_0)}{1!0!} \cdot (x - x_0)^1 \cdot (y - y_0)^0 + \frac{\frac{\partial f}{\partial y}(\vec{x}_0)}{0!1!} \cdot (x - x_0)^0 \cdot (y - y_0)^1 \\
 &\quad + \frac{\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)}{2!0!} \cdot (x - x_0)^2 \cdot (y - y_0)^0 + \frac{\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0)}{1!1!} \cdot (x - x_0)^1 \cdot (y - y_0)^1 + \frac{\frac{\partial^2 f}{\partial y^2}(\vec{x}_0)}{0!2!} \cdot (x - x_0)^0 \cdot (y - y_0)^2.
 \end{aligned}$$

Hierbei haben wir verwendet:  $0! = 1$ ,  $1! = 1$ ,  $2! = 1 \cdot 2 = 2$  usw. sowie  $s^0 = 1$  für alle  $s \in \mathbb{R} \setminus \{0\}$ , und im Zusammenhang mit Taylorpolynomen vereinbaren wir  $0^0 := 1$  (normalerweise wird  $0^0$  als undefiniert angesehen).

Nach diesem Bildungsprinzip werden dann auch Taylorpolynome noch höheren Grades zusammengesetzt.

**Beispiel 3.4:**

Wir wollen für die Funktion  $f$  mit der Gleichung  $f(x, y) = (1 + 3y^2)e^x$  im Punkt  $[x_0, y_0]^T = [0, 0]^T$  das Taylorpolynom zweiten Grades aufstellen. Die Ableitungen und deren Werte sind:

$$\begin{array}{ll}
 f(0, 0) = 1, & \\
 \frac{\partial f}{\partial x}(x, y) = (1 + 3y^2)e^x, & \frac{\partial f}{\partial x}(0, 0) = 1, \\
 \frac{\partial f}{\partial y}(x, y) = 6ye^x, & \frac{\partial f}{\partial y}(0, 0) = 0, \\
 \frac{\partial^2 f}{\partial x^2}(x, y) = (1 + 3y^2)e^x, & \frac{\partial^2 f}{\partial x^2}(0, 0) = 1, \\
 \frac{\partial^2 f}{\partial x \partial y}(x, y) = 6ye^x, & \frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial^2 f}{\partial y \partial x}(0, 0) = 0, \\
 \frac{\partial^2 f}{\partial y^2}(x, y) = 6e^x, & \frac{\partial^2 f}{\partial y^2}(0, 0) = 6.
 \end{array}$$

Damit erhalten wir folgendes Taylorpolynom zweiten Grades:

$$\begin{aligned}
 & f(0, 0) + \frac{\partial f}{\partial x}(0, 0) \cdot (x - 0) + \frac{\partial f}{\partial y}(0, 0) \cdot (y - 0) \\
 &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0) \cdot (x - 0)^2 + \frac{\partial^2 f}{\partial x \partial y}(0, 0) \cdot (x - 0) \cdot (y - 0) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0) \cdot (y - 0)^2 \\
 &= 1 + 1 \cdot x + 0 \cdot y + \frac{1}{2} \cdot 1 \cdot x^2 + 0 \cdot x \cdot y + \frac{1}{2} \cdot 6 \cdot y^2 \\
 &= 1 + x + \frac{1}{2}x^2 + 3y^2.
 \end{aligned}$$

Wir hätten auch anders rechnen können, indem wir uns an die Reihe für  $e^x$  aus dem ersten Semester erinnern:

$$\begin{aligned}
 f(x, y) &= (1 + 3y^2)e^x = (1 + 3y^2) \cdot \left( 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots \right) \quad \left| \text{ausmultiplizieren} \right. \\
 &= \underbrace{1 + x + \frac{x^2}{2} + 3y^2}_{\text{alle Terme bis Grad 2}} + \underbrace{\dots}_{\text{alle Terme ab Grad 3}}.
 \end{aligned}$$

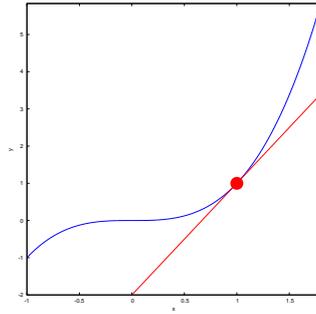


Abbildung 3.14: Die Tangente an den Graph der Funktion  $y = f(x)$ , genommen im Punkt  $[x_0, f(x_0)]^\top$ , hört auf die Gleichung  $y = f(x_0) + f'(x_0) \cdot (x - x_0)$ . Wir bekommen diese Tangentengleichung, indem wir die Taylorentwicklung hinschreiben und alles ab dem quadratischen Term wegwerfen. Die Tangente ist diejenige Gerade, die den Graphen der Funktion  $f$  in der Nähe von  $x_0$  bestmöglich annähert. Eine Funktion  $f$  ist genau dann im Punkt  $x_0$  differenzierbar, wenn es dort eine eindeutig bestimmte Tangente gibt.

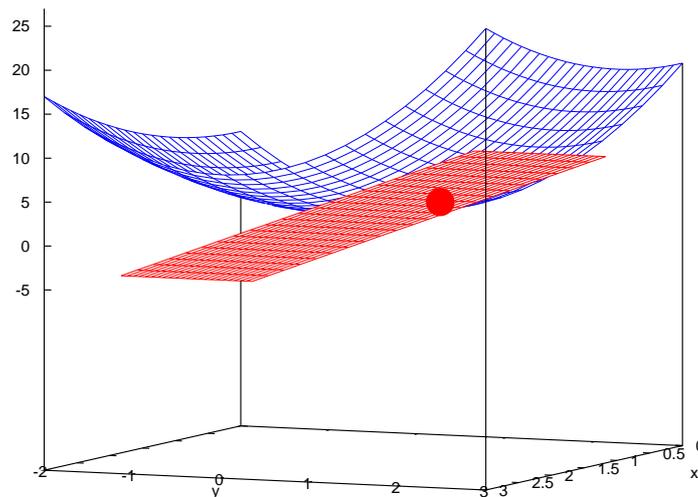


Abbildung 3.15: Die Tangentialebene an den Graph der Funktion  $y = f(\vec{x})$ , genommen im Punkt  $[\vec{x}_0^\top, f(\vec{x}_0)]^\top$ , hört auf die Gleichung  $y = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$ . Wir bekommen diese Ebenengleichung, indem wir die Taylorentwicklung hinschreiben und alles ab den quadratischen Termen wegwerfen. Die Tangentialebene ist diejenige Ebene, die den Graphen der Funktion  $f$  in der Nähe von  $\vec{x}_0$  bestmöglich annähert. Eine Funktion  $f$  ist genau dann im Punkt  $\vec{x}_0$  total differenzierbar, wenn es dort eine eindeutig bestimmte Tangentialebene gibt.

Die sprachlichen Parallelen in den Bildunterschriften sind didaktische Absicht !

**Definition 3.5 (Extremwerte):**

Sei  $M \subseteq \mathbb{R}^n$  eine (nicht unbedingt offene) Menge und  $f: M \rightarrow \mathbb{R}$ . Ein Punkt  $\vec{x}_0 \in M$  heißt

**globales Maximum von  $f$** , wenn  $f(\vec{x}) \leq f(\vec{x}_0)$  für jedes  $\vec{x} \in M$  gilt,

**globales Minimum von  $f$** , wenn  $f(\vec{x}) \geq f(\vec{x}_0)$  für jedes  $\vec{x} \in M$  gilt,

**lokales Maximum von  $f$** , wenn es einen Radius  $\delta > 0$  gibt mit  $f(\vec{x}) \leq f(\vec{x}_0)$  für jedes  $\vec{x} \in U_\delta(\vec{x}_0) \cap M$ ,

**lokales Minimum von  $f$** , wenn es einen Radius  $\delta > 0$  gibt mit  $f(\vec{x}) \geq f(\vec{x}_0)$  für jedes  $\vec{x} \in U_\delta(\vec{x}_0) \cap M$ ,

**globales Extremum**, wenn  $\vec{x}_0$  ein globales Maximum oder ein globales Minimum von  $f$  ist,

**lokales Extremum**, wenn  $\vec{x}_0$  ein lokales Maximum oder ein lokales Minimum von  $f$  ist.

Wir **erinnern** an den Schulstoff: für Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  gilt:

$$\begin{aligned} f'(x_0) = 0 \quad \text{und} \quad f''(x_0) > 0 &\implies x_0 \text{ ist lokales Minimum,} \\ f'(x_0) = 0 \quad \text{und} \quad f''(x_0) < 0 &\implies x_0 \text{ ist lokales Maximum.} \end{aligned}$$

**Satz 3.6 (Notwendige Bedingung für Extrema):**

Sei  $M \subseteq \mathbb{R}^n$ ,  $\vec{x}_0$  ein innerer Punkt von  $M$ , und sei  $f: M \rightarrow \mathbb{R}$  total differenzierbar in  $\vec{x}_0$  mit lokalem Extremum in  $\vec{x}_0$ . Dann gilt  $\nabla f(\vec{x}_0) = \vec{0}^\top$ .

Wir führen eine Sprechweise ein: wenn eine Funktion  $f$  in  $\vec{x}_0$  total differenzierbar ist und wenn dort  $\nabla f(\vec{x}_0) = \vec{0}^\top$  gilt, dann nennen wir  $\vec{x}_0$  einen *kritischen Punkt*.

Damit haben wir: jedes innere lokale Extremum einer total differenzierbaren Funktion ist ein kritischer Punkt.

**Warnung 3.7:**

Wie im Eindimensionalen gilt: nicht jeder kritische Punkt ist eine Extremstelle! Man denke z.B. an die Funktion  $f$  mit der Gleichung  $f(x, y) = x^2 - y^2$  (siehe Bild 1.1 (links)) oder an die „Affensattelfunktion“  $f(x, y) = x^3 - 3xy^2$  (siehe Bild 3.16).

Kritische Punkte, die keine Extrema sind, nennt man *Sattelpunkte*.

**Bemerkung 3.8 (Vorüberlegung für hinreichende Bedingungen):**

Sei  $M \subseteq \mathbb{R}^n$  eine offene Menge,  $f: M \rightarrow \mathbb{R}$  sei zweimal stetig partiell differenzierbar, und für  $\vec{x}_0 \in M$  gelte  $\nabla f(\vec{x}_0) = \vec{0}^\top$ . Der Satz von Taylor sagt uns dann: für jedes  $\vec{x}$  in der Nähe von  $\vec{x}_0$  gibt es dann einen Punkt  $\vec{\xi}$  auf der Verbindungsstrecke von  $\vec{x}$  und  $\vec{x}_0$ , sodass

$$f(\vec{x}) = f(\vec{x}_0) + \underbrace{\nabla f(\vec{x}_0)}_{=\vec{0}^\top} \cdot (\vec{x} - \vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^\top \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0).$$

Also haben wir dann

$$f(\vec{x}) - f(\vec{x}_0) = \frac{1}{2}(\vec{x} - \vec{x}_0)^\top \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0). \quad (\spadesuit)$$

Nach dem Satz von Schwarz ist die Hesse-Matrix  $H_f(\vec{\xi}) \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix. Nach Voraussetzung der zweimaligen stetigen Differenzierbarkeit ist  $H_f(\vec{\xi})$  aber auch eine stetige Matrix-Funktion in der Vektorvariablen  $\vec{\xi}$ . Das bedeutet also  $H_f(\vec{\xi}) \approx H_f(\vec{x}_0)$  für  $\vec{\xi}$  in der Nähe von  $\vec{x}_0$ .

Wenn die rechte Seite von  $(\spadesuit)$  immer  $\geq 0$  ist für alle  $\vec{\xi}$  in der Nähe von  $\vec{x}_0$ , dann ist  $\vec{x}_0$  ein lokales Minimum von  $f$ . Und wenn die rechte Seite immer  $\leq 0$  ist für alle  $\vec{\xi}$  in der Nähe von  $\vec{x}_0$ , dann ist  $\vec{x}_0$  ein lokales Maximum von  $f$ .

Damit ist die nächste Definition motiviert:

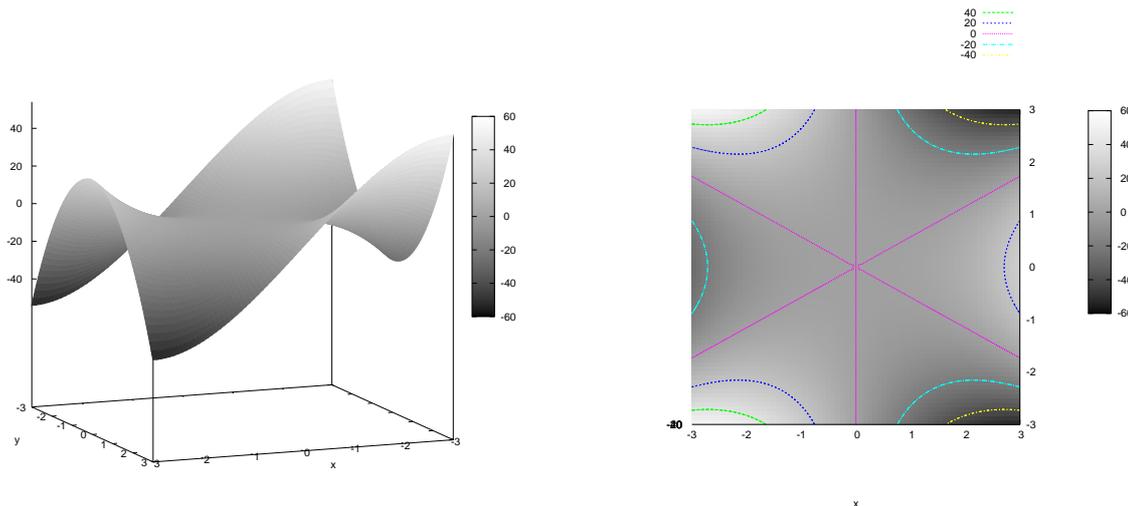


Abbildung 3.16: Links: die Affensattelfunktion  $f(x, y) = x^3 - 3xy^2$ . Wir sehen drei Richtungen, wo die Funktion „runtergeht“, und drei Richtungen, wo die Funktion „raufgeht“. Rechts der dazugehörige Höhenlinienplot.

**Definition 3.9:**

Sei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix. Dann heißt

$A$  **positiv definit** genau dann, wenn  $\vec{x}^\top \cdot A \cdot \vec{x} > 0$  für alle  $\vec{x} \neq \vec{0}$ ,

$A$  **negativ definit** genau dann, wenn  $\vec{x}^\top \cdot A \cdot \vec{x} < 0$  für alle  $\vec{x} \neq \vec{0}$ ,

$A$  **indefinit** genau dann, wenn es zwei Vektoren  $\vec{x}$  und  $\vec{y}$  aus dem  $\mathbb{R}^n$  gibt mit  $\vec{x}^\top \cdot A \cdot \vec{x} > 0$  und  $\vec{y}^\top \cdot A \cdot \vec{y} < 0$ .

**Satz 3.10 (Hinreichende Bedingung für Extrema):**

Sei  $M \subseteq \mathbb{R}^n$  offen,  $f: M \rightarrow \mathbb{R}$  sei zweimal stetig partiell differenzierbar, und  $\vec{x}_0 \in M$  sei ein kritischer Punkt von  $f$ . Dann gilt:

**Wenn  $H_f(\vec{x}_0)$  positiv definit ist,** dann ist  $\vec{x}_0$  ein lokales Minimum von  $f$ .

**Wenn  $H_f(\vec{x}_0)$  negativ definit ist,** dann ist  $\vec{x}_0$  ein lokales Maximum von  $f$ .

**Wenn  $H_f(\vec{x}_0)$  indefinit ist,** dann ist  $\vec{x}_0$  ein Sattelpunkt von  $f$ .

**Bemerkung 3.11:**

Wie schon im Falle von  $n = 1$  bekannt, gibt es auch kritische Punkte, die von diesem Kriterium nicht erfasst werden („unentscheidbare Fälle“). Das kann zum Beispiel aussehen wie folgt:

$$\begin{aligned} \text{Es ist } \vec{x}^\top \cdot H_f(\vec{x}_0) \cdot \vec{x} &\geq 0 \quad \text{für alle } \vec{x}, \\ \text{aber es gibt auch ein } \vec{y} \neq 0 \quad &\text{mit } \vec{y}^\top \cdot H_f(\vec{x}_0) \cdot \vec{y} = 0. \end{aligned}$$

In diesem Fall nennt man die Hesse-Matrix  $H_f(\vec{x}_0)$  *positiv semi-definit*.

Eine analoge Begriffsbildung gibt es mit  $\leq 0$ , *negativ semidefinit*.

Das Affensattelbeispiel wird von unserem hinreichenden Kriterium übrigens auch nicht erfasst, weil dann für  $\vec{x}_0 = [0, 0]^\top$  die Hesse-Matrix  $H_f(\vec{x}_0)$  gleich der  $2 \times 2$ -Nullmatrix ist.

**Satz 3.12 (Praktischer Nachweis der Definitheit):**

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann gilt:

(a)

$A$  ist positiv definit

$\iff$  alle Eigenwerte von  $A$  sind strikt positiv

$\iff$  die Unterdeterminanten  $|a_{11}|$ ,  $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ ,  $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \dots$  sind alle positiv.

Diese Unterdeterminanten nennt man *Unteterminoren*. (Bemerkung: Die Vorzeichenreihenfolge der Unteterminoren kann man sich merken, indem man an die Einheitsmatrix  $E_n$  für  $A$  denkt.)

(b)

$A$  ist negativ definit

$\iff$  alle Eigenwerte von  $A$  sind strikt negativ

$\iff$  die Unteterminoren sind abwechselnd negativ, positiv, negativ, positiv usw.

(Bemerkung: Die Vorzeichenreihenfolge der Unteterminoren kann man sich merken, indem man an die Matrix  $-E_n$  für  $A$  denkt, denn  $-E_n$  ist negativ definit und hat die Unteterminoren  $-1, +1, -1, \dots$ )

(c)

$A$  ist indefinit

$\iff$  es gibt einen strikt positiven und einen strikt negativen Eigenwert von  $A$

$\iff$  die Unteterminoren sind alle  $\neq 0$  und passen in keines der beiden obigen Schemata.

(Achtung: in der letzten Zeile gilt wirklich nur  $\iff$ , nicht  $\implies$ .)

**Beispiele 3.13:** (a) Die Matrix  $A = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}$  ist symmetrisch. Die Unteterminoren sind

$$\det [3] = 3 > 0, \quad \det \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix} = 12 - 4 = 8 > 0,$$

also ist die Matrix  $A$  positiv definit.

(b) Wir betrachten die Matrix

$$B = \begin{bmatrix} -1 & -1 & 0 \\ -1 & -3 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Diese ist symmetrisch, und die Unteterminoren sind

$$\det [-1] = -1 < 0, \quad \det \begin{bmatrix} -1 & -1 \\ -1 & -3 \end{bmatrix} = 3 - 1 = 2 > 0,$$

$$\det \begin{bmatrix} -1 & -1 & 0 \\ -1 & -3 & 1 \\ 0 & 1 & -1 \end{bmatrix} = \det \begin{bmatrix} -1 & -1 & 0 \\ -1 & -2 & 1 \\ 0 & 0 & -1 \end{bmatrix} = -\det \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix} = -(2 - 1) = -1 < 0.$$

Also ist  $B$  negativ definit.

(c) Wir betrachten die symmetrische Matrix  $C = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ . Die Unteterminoren sind

$$\det [1] = 1 > 0, \quad \det \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = 2 - 9 = -7 < 0.$$

Also ist  $C$  indefinit.

**Beispiel 3.14:**

Für die Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = (x^2 + 2y^2)e^{-x^2-y^2}$  wollen wir alle Extrema finden. Zuerst suchen wir die kritischen Punkte; also alle diejenigen  $[x, y]^T$  mit  $\nabla f(x, y) = [0, 0]$ . Dazu bestimmen wir  $\nabla f$ :

$$\begin{aligned}\nabla f(x, y) &= \left[ 2xe^{-x^2-y^2} + (x^2 + 2y^2)e^{-x^2-y^2} \cdot (-2x), \quad 4ye^{-x^2-y^2} + (x^2 + 2y^2)e^{-x^2-y^2} \cdot (-2y) \right] \\ &= \left[ 2xe^{-x^2-y^2}(1 - x^2 - 2y^2), 2ye^{-x^2-y^2}(2 - x^2 - 2y^2) \right] \stackrel{!}{=} [0, 0].\end{aligned}$$

Weil  $e^{-x^2-y^2}$  niemals Null werden kann, müssen kritische Punkte  $[x, y]^T$  folgende beiden Gleichungen lösen:

$$x(1 - x^2 - 2y^2) = 0, \tag{I}$$

$$y(2 - x^2 - 2y^2) = 0. \tag{II}$$

**Sei  $x = 0$ :** dann ist (I) erfüllt, und aus (II) folgt  $y(2 - 2y^2) = 0$ , also  $0 = y(1 - y^2) = y(1 - y)(1 + y)$ , also  $y = 0$  oder  $y = 1$  oder  $y = -1$ .

**Sei  $x \neq 0$ :** dann liefert (I) uns  $(1 - x^2 - 2y^2) = 0$ , und aus (II) bekommen wir

$$0 = y \cdot (2 - x^2 - 2y^2) = y \cdot (1 + (1 - x^2 - 2y^2)) = y \cdot (1 + 0) = y,$$

also  $y = 0$ . Damit gehen wir wieder zu (I) zurück und bekommen  $(1 - x^2) = 0$ , also  $x = 1$  oder  $x = -1$ .

Insgesamt finden wir die folgenden kritischen Stellen:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Jetzt ermitteln wir den ersten Eintrag der Hesse-Matrix:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2}(x, y) &= 2e^{-x^2-y^2}(1 - x^2 - 2y^2) + 2xe^{-x^2-y^2} \cdot (-2x) \cdot (1 - x^2 - 2y^2) + 2xe^{-x^2-y^2} \cdot (-2x) \\ &= 2e^{-x^2-y^2} \cdot (1 - x^2 - 2y^2 - 4x^2(1 - x^2 - 2y^2) - 4x^2) \\ &= 2e^{-x^2-y^2} (1 - x^2 - 2y^2 - 4x^2 + 4x^4 + 8x^2y^2 - 4x^2) \\ &= 2e^{-x^2-y^2} (1 - 9x^2 - 2y^2 + 4x^4 + 8x^2y^2).\end{aligned}$$

Auf ähnlichem Wege findet man

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = -4xye^{-x^2-y^2} (3 - x^2 - 2y^2), \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 2e^{-x^2-y^2} (2 - x^2 - 10y^2 + 2x^2y^2 + 4y^4),$$

und somit finden wir

$$H_f(x, y) = 2e^{-x^2-y^2} \begin{bmatrix} 1 - 9x^2 - 2y^2 + 4x^4 + 8x^2y^2 & -2xy(3 - x^2 - 2y^2) \\ -2xy(3 - x^2 - 2y^2) & 2 - x^2 - 10y^2 + 2x^2y^2 + 4y^4 \end{bmatrix}.$$

Wir setzen darin die kritischen Stellen ein:

$$H_f(0, 0) = 2 \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{mit Eigenwerten 2 und 4, also hat } f \text{ lokales Minimum in } \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$H_f(0, 1) = \frac{2}{e} \begin{bmatrix} 1 - 2 & 0 \\ 0 & 2 - 10 + 4 \end{bmatrix} = \frac{2}{e} \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}$$

$$\text{mit Eigenwerten } -\frac{2}{e} \text{ und } -\frac{8}{e}, \text{ also hat } f \text{ lokales Maximum in } \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$H_f(0, -1) = H_f(0, 1), \quad \text{also ebenso lokales Maximum in } \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

$$H_f(1, 0) = \frac{2}{e} \begin{bmatrix} 1 - 9 + 4 & 0 \\ 0 & 2 - 1 \end{bmatrix} = \frac{2}{e} \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{mit Eigenwerten } -\frac{8}{e} \text{ und } \frac{2}{e}, \text{ also einem Sattelpunkt in } \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$H_f(-1, 0) = H_f(1, 0), \quad \text{also ebenso ein Sattelpunkt in } \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

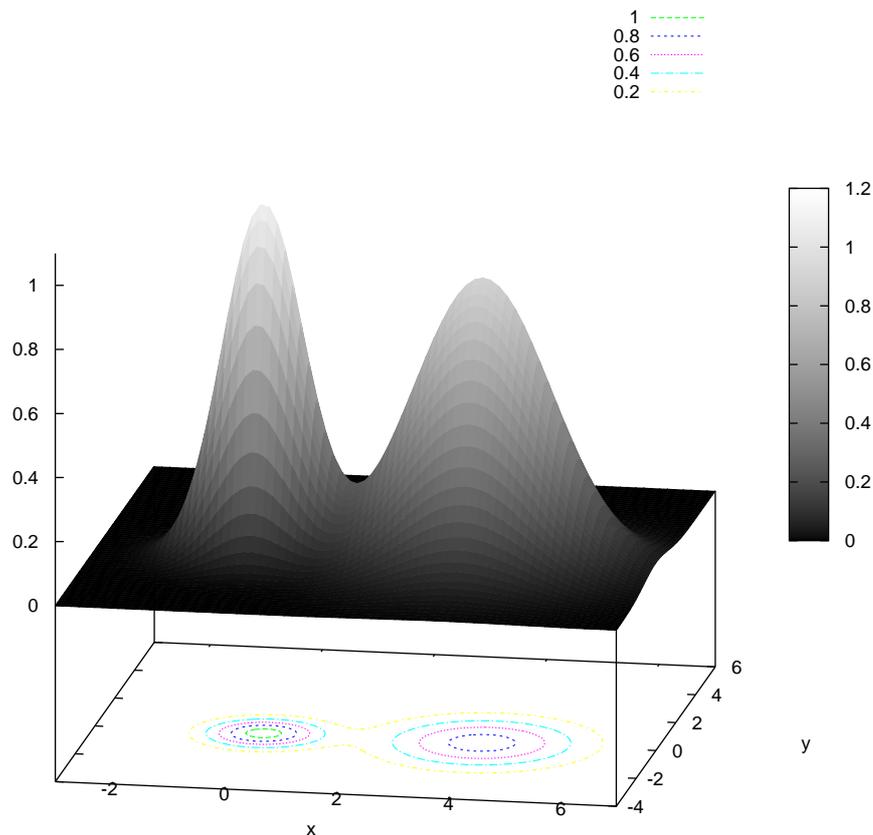


Abbildung 3.17: Achtung: diese Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , gegeben durch die Funktionsgleichung  $f(x, y) = 1.1e^{-x^2-y^2} + 0.9e^{-(x-3.9)^2/2-y^2/3}$  hat zwei Maxima, aber dazwischen kein Minimum (denn sie hat überhaupt kein Minimum auf dem  $\mathbb{R}^2$ ). Funktionen im  $\mathbb{R}^2$  können sich also anders verhalten als Funktionen im  $\mathbb{R}^1$ .

**Bemerkung 3.15 (Die Gradientenmethode):**

Es gibt ein numerisches Verfahren zur Suche nach lokalen Extrema, und wir zeigen jetzt für den Maximumfall, wie es funktioniert.

Gegeben sei eine stetig partiell differenzierbare Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , und wir suchen ein lokales Maximum von  $f$ .

Wir erinnern daran, dass  $\nabla f(\vec{x}_0)$  in Richtung des steilsten Anstiegs von  $f$  im Punkt  $\vec{x}_0$  zeigt.

**Schritt 1:** wähle einen Startwert  $\vec{x}_0$ . Es empfiehlt sich, diesen Startwert so zu wählen, dass er schon eine gute Näherung des Maximums ergibt (indem man anderweitig gewonnenes Wissen über die Funktion  $f$  heranzieht, um die Lage des Maximums einzugrenzen).

**Schritt 2:** wähle eine Schrittweite  $h$  und setze  $\vec{x}_1 := \vec{x}_0 + h \cdot \nabla f(\vec{x}_0)$ .

**Schritt 3:** Wenn  $f(\vec{x}_1) \geq f(\vec{x}_0)$ , dann iteriere das Verfahren (gehe also zurück zu Schritt 2 mit  $\vec{x}_1$  anstatt von  $\vec{x}_0$ ). Wenn  $f(\vec{x}_1) < f(\vec{x}_0)$  (also wenn man zu weit gelaufen ist), dann wiederhole Schritt 2 mit kleinerer Schrittweite  $h$  (z.B.  $h/2$  anstatt  $h$ ).

**Abbruchbedingung:**  $\nabla f(\vec{x}_k)$  ist nahe genug bei  $\vec{0}$ . Die Toleranz dafür ist vorher festzulegen.

Wir kommen jetzt zu einem neuen Thema.

**Beispiel 3.16 (Extremwerte unter Nebenbedingungen):** (a) Wir suchen Extrema der Funktion  $f(x, y) = x^3 - y^2$  auf der Menge

$$N = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\}.$$

Die Gleichung  $x^2 + y^2 = 1$  nennen wir *Nebenbedingung*. Siehe Bild 3.18.

Zunächst **beobachten** wir: die Funktion  $f$  ist stetig und die Menge  $N$  ist kompakt. Also existieren tatsächlich Minima und Maxima der Funktion  $f$  auf  $N$ ; wir müssen diese Extrema „nur noch“ finden. Bei dieser Aufgabe kommen wir mit **Einsetzen** zum Ziel: Für alle  $[x, y]^T \in N$  ist  $y^2 = 1 - x^2$ , also ist dort

$$f(x, y) = x^3 - y^2 = x^3 - 1 + x^2 =: F(x),$$

und jetzt suchen wir kritische Punkte von  $F$ :

$$\begin{aligned} 0 &\stackrel{!}{=} F'(x) = 3x^2 + 2x && \iff && x = 0 \text{ oder } 3x + 2 = 0 \\ &&& \iff && x = 0 \text{ oder } x = -\frac{2}{3}. \end{aligned}$$

Wir setzen diese Werte in  $F$  ein: es ist  $F(0) = -1$  und

$$F\left(-\frac{2}{3}\right) = \left(-\frac{2}{3}\right)^3 - 1 + \left(-\frac{2}{3}\right)^2 = -\frac{8}{27} - 1 + \frac{4}{9} = \frac{-8 - 27 + 12}{27} = -\frac{23}{27}.$$

Die Maxima und Minima müssen wegen unserer Beobachtung existieren, also liegt das Maximum von  $f$  vor für  $x = -\frac{2}{3}$ ; und das dazugehörige  $y$  ergibt sich dann aus der Gleichung  $x^2 + y^2 = 1$ , also  $\frac{4}{9} + y^2 = 1$ , also  $y = \pm\frac{\sqrt{5}}{3}$ . Es hat also  $f$  Maxima in den Punkten

$$\begin{bmatrix} -\frac{2}{3} \\ \frac{\sqrt{5}}{3} \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} -\frac{2}{3} \\ -\frac{\sqrt{5}}{3} \end{bmatrix}.$$

Die Minima gibt es dann für  $x = 0$ , und das dazugehörige  $y$  ist gegeben über die Gleichung  $x^2 + y^2 = 1$ , also  $0^2 + y^2 = 1$ , also  $y = \pm 1$ . Es hat also  $f$  Minima in den Punkten

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

- (b) Jetzt suchen wir Extrema der Funktion  $f(x, y) = x^3 - y^3$  auf derselben Menge  $N$  wie in (a). Wir können die Nebenbedingung  $x^2 + y^2 = 1$  in die Gestalt bringen  $g(x, y) = 0$ , mit  $g(x, y) := x^2 + y^2 - 1$ . Aber leider ist der Rechenweg aus (a) jetzt nicht mehr so schön durchführbar, wie man schnell bemerkt. Wir wollen uns eine andere Herangehensweise überlegen.

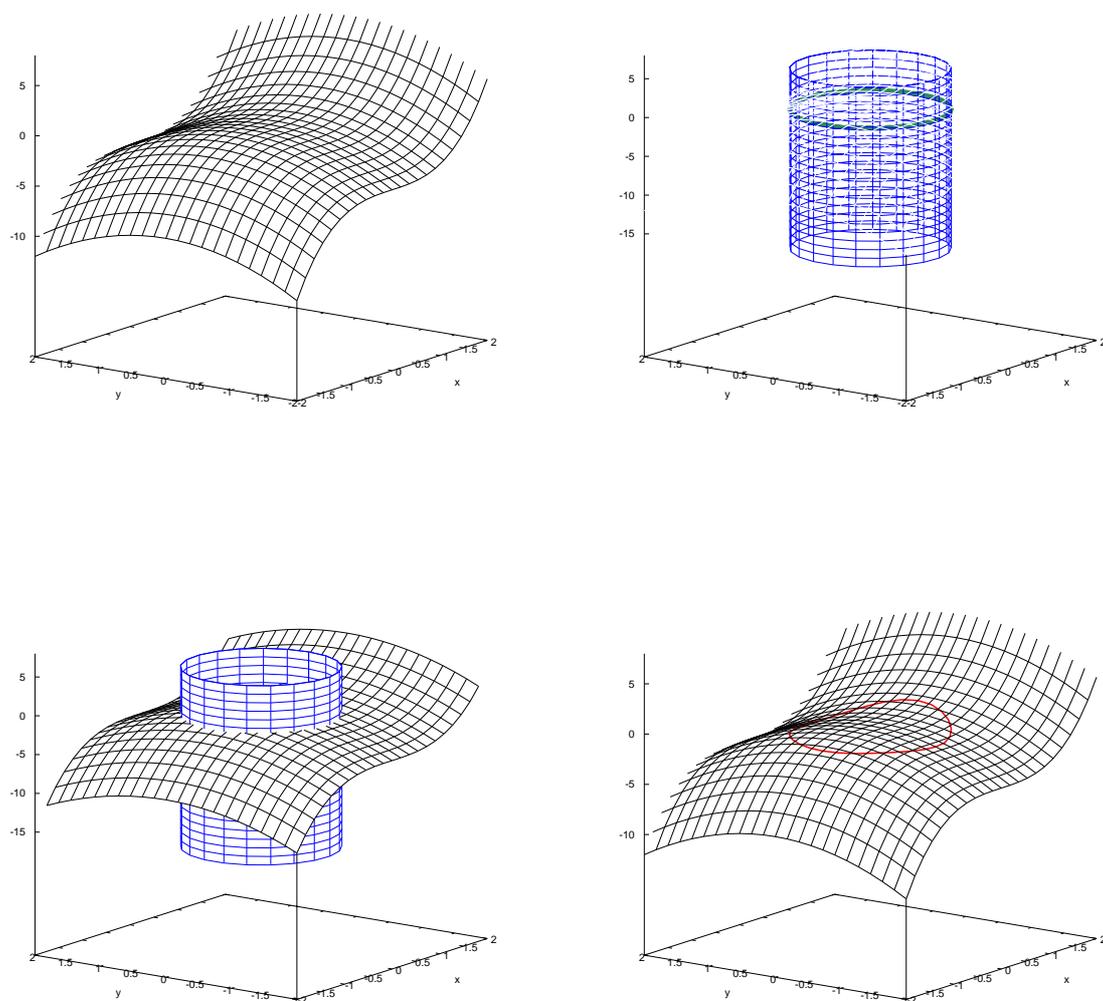


Abbildung 3.18: Links oben: die durch die Funktion  $z = f(x, y) = x^3 - y^2$  beschriebene Fläche. Rechts oben: die Nebenbedingung  $x^2 + y^2 = 1$  beschreibt in der  $xy$ -Ebene einen Kreis (grün). Wenn wir diesen entlang der  $z$ -Achse parallel verschieben, entsteht ein blauer Zylinder. Links unten: wir schneiden den blauen Zylinder mit der Fläche zur Funktion  $z = f(x, y) = x^3 - y^2$ . Rechts unten: als Schnittfigur entsteht eine rote Kurve. Von dieser Kurve suchen wir die lokal höchsten und tiefsten Punkte. Siehe Beispiel 3.16 (a).

**Bemerkung 3.17 (Optimierung unter Nebenbedingungen):**

Wir wollen folgendes allgemeine **Problem** lösen: Seien  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  und  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  gegebene Funktionen. Wir suchen Extrema von  $f$  auf der Menge  $N := \{\vec{x} \in \mathbb{R}^n : g(\vec{x}) = 0\}$ . Die Gleichung  $g(\vec{x}) \stackrel{!}{=} 0$  nennen wir *Nebenbedingung*.

Der Abbildung 3.19 und ihrem Begleittext entnehmen wir eine heuristische Überlegung, die uns hilft, die gesuchten Extrema zu finden. Wir machen uns plausibel, dass die gesuchten Extrema von  $f$  auf  $N$  nur an denjenigen Stellen überhaupt liegen können, an denen

- die Höhenlinien von  $f$  und  $g$  einander berühren, ohne einander zu schneiden,
- also die Tangenten an die Höhenlinien zueinander parallel sind.

Wir **erinnern** uns: es steht  $\nabla f(\vec{x})$  senkrecht auf der Höhenlinie zu  $f$  durch den Punkt  $\vec{x}$ , und entsprechendes gilt für die Funktion  $g$ .

Das bedeutet: Extrema von  $f$  auf  $N$  kann es nur an den Stellen geben, an denen  $\nabla f(\vec{x})$  und  $\nabla g(\vec{x})$  parallel zueinander sind. Es muss also ein  $\lambda \in \mathbb{R}$  geben mit  $\nabla f(\vec{x}) = \lambda \nabla g(\vec{x})$ .

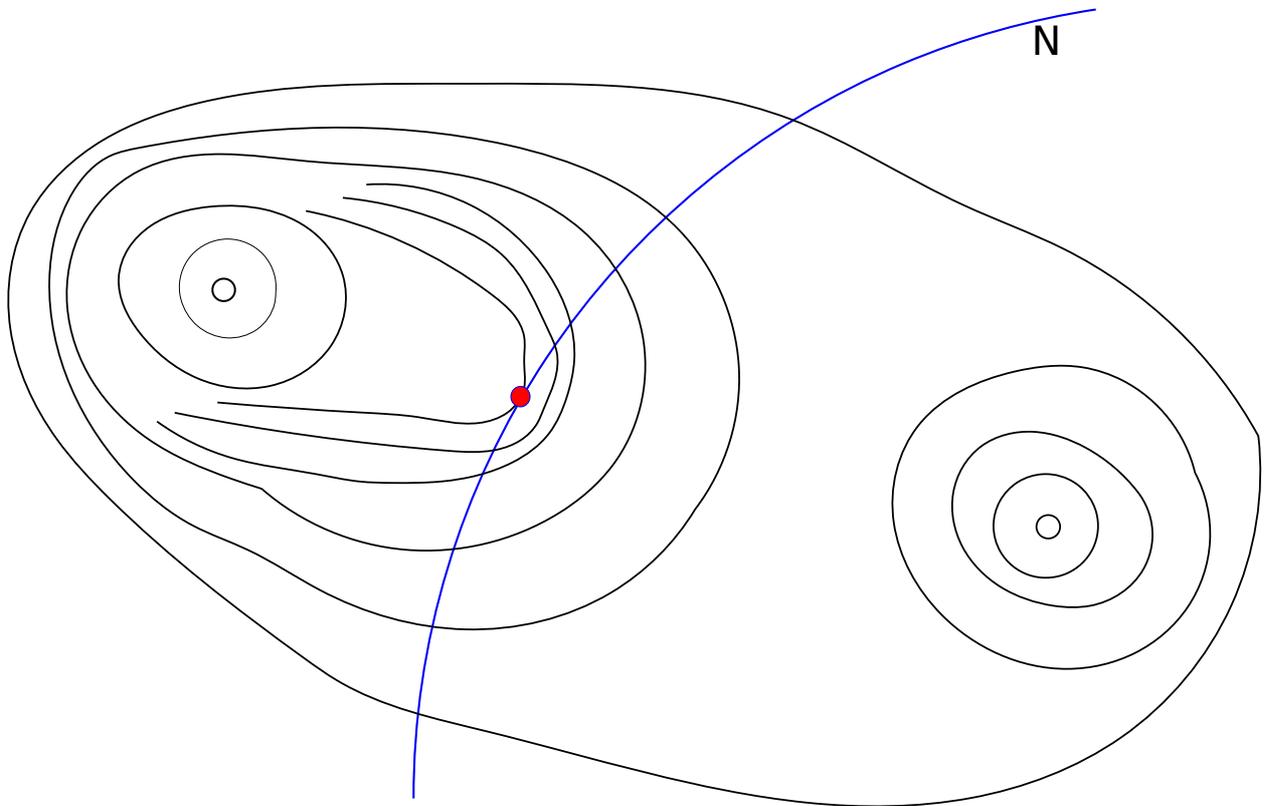


Abbildung 3.19: Eine handgezeichnete Skizze zur Suche nach Maxima der Funktion  $f(\vec{x})$  für  $\vec{x} \in \mathbb{R}^2$  unter der Nebenbedingung  $g(\vec{x}) = 0$ . Die Menge aller derjenigen  $\vec{x} \in \mathbb{R}^2$  mit  $g(\vec{x}) = 0$  ist gerade die Menge  $N$  (blaue Kurve), und dies ist die Höhenlinie der Funktion  $g$  zur Höhe Null. Die Höhenlinien der Funktion  $f$  sind schwarz skizziert, und in diesem Beispiel soll  $f$  genau zwei lokale Maxima besitzen. Der höchste Punkt von  $f$  auf der blauen Linie ist rot markiert. Diesen roten Punkt suchen wir. Am roten Punkt gilt: wenn wir die blaue Kurve nach „rechts oben“ entlanggehen, dann treffen wir dort nur niedrigere Niveaulinien von  $f$ . Und wenn wir vom roten Punkt aus die blaue Kurve nach „links unten“ entlanggehen, dann treffen wir ebenfalls nur niedrigere Niveaulinien von  $f$ . Also berührt die durch den roten Punkt laufende  $f$ -Niveaulinie die blaue Kurve  $N$  tangential.

**Satz 3.18 (Satz von Lagrange):**

Seien  $M \subseteq \mathbb{R}^n$  offen,  $f: M \rightarrow \mathbb{R}$  und  $g: M \rightarrow \mathbb{R}$  beide stetig partiell differenzierbar, und sei  $\nabla g(\vec{x}) \neq \vec{0}^\top$  für alle  $\vec{x} \in M$ .

Dann gilt: falls  $f$  in  $\vec{x}_0 \in M$  ein Extremum unter der Nebenbedingung  $g = 0$  hat, so existiert ein  $\lambda \in \mathbb{R}$  mit  $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$ .

Wir kehren zurück zu Beispiel 3.16 (b).

**Beispiel 3.19:**

Wir suchen Extrema von  $f(x, y) = x^3 - y^3$  unter der Nebenbedingung  $g(x, y) = x^2 + y^2 - 1 = 0$ .

Die Lagrange-Bedingung lautet: es gibt ein  $\lambda \in \mathbb{R}$  mit

$$\nabla f(x, y) = \lambda \nabla g(x, y) \iff \begin{bmatrix} 3x^2 & -3y^2 \end{bmatrix} = \lambda \begin{bmatrix} 2x & 2y \end{bmatrix},$$

und das führt uns zum Gleichungssystem

$$\begin{aligned} 3x^2 &= 2\lambda x, \\ -3y^2 &= 2\lambda y, \\ x^2 + y^2 - 1 &= 0. \end{aligned}$$

(Beachte: dies sind zwar drei Gleichungen für die drei Unbekannten  $x$ ,  $y$  und  $\lambda$ , aber damit ist über die Lösbarkeit dieses Gleichungssystems noch gar nichts gesagt.)

Dieses Gleichungssystem kann gelöst werden, und wir finden die Lösungen

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Dies sind alle Stellen, an denen es überhaupt Extrema geben könnte.

Allerdings ist  $f$  stetig auf der Menge

$$N = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : g(x, y) = 0 \right\} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\},$$

und  $N$  ist die Einheitskreislinie im  $\mathbb{R}^2$ . Offenkundig ist  $N$  beschränkt und abgeschlossen, also kompakt. Nach Satz 1.17 besitzt  $f$  auf  $N$  tatsächlich ein Maximum und ein Minimum. Wir brauchen also bloß noch die Werte von  $f$  an den extremwertverdächtigen Stellen ausrechnen und vergleichen:

$$\begin{aligned} f(1, 0) &= 1, & f(-1, 0) &= 1, & f(0, 1) &= -1, & f(0, -1) &= -1, \\ f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) &= \frac{1}{\sqrt{2}^3} + \frac{1}{\sqrt{2}^3} = \frac{2}{\sqrt{2}^3} = \frac{1}{\sqrt{2}}, \\ f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) &= -\frac{1}{\sqrt{2}^3} - \frac{1}{\sqrt{2}^3} = -\frac{2}{\sqrt{2}^3} = -\frac{1}{\sqrt{2}}, \end{aligned}$$

und deshalb besitzt  $f$  auf der Menge  $N$

$$\begin{aligned} \text{Maxima in } \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix} & \quad \text{mit Funktionswert } 1, \\ \text{Minima in } \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix} & \quad \text{mit Funktionswert } -1. \end{aligned}$$



# Kapitel 3

## Integration in mehreren Variablen

### 1 Parameterintegrale

**Definition 1.1 (Parameterintegral):**

Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  und  $a, b \in \mathbb{R}$ . Weiter sei für jedes  $x \in \mathbb{R}$  die Abbildung  $y \mapsto f(x, y)$  stückweise stetig. Dann definieren wir

$$g(x) := \int_{y=a}^b f(x, y) \, dy$$

und nennen dies *Parameterintegral*. Die Schreibweise  $\int_a^b f(x, y) \, dy$  ist auch gebräuchlich.

**Beispiele 1.2:** (a) Wir setzen

$$g(x) := \int_{y=1}^2 y e^{xy} \, dy, \quad x \in \mathbb{R},$$

und fragen uns nach  $\int_{x=c}^d g(x) \, dx$  sowie  $g'(x)$ .

(b) Sei die Funktion  $u_0: \mathbb{R} \rightarrow \mathbb{R}$  stetig. Wir definieren die *Gaußglockenfunktion*

$$G(t, x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x|^2}{4t}}$$

und betrachten das uneigentliche Parameterintegral

$$u(t, x) := \int_{y=-\infty}^{\infty} G(t, x-y) u_0(y) \, dy = \frac{1}{\sqrt{4\pi t}} \int_{y=-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}} u_0(y) \, dy. \quad (\spadesuit)$$

Hier haben wir sogar zwei Parameter, nämlich  $t$  und  $x$ .

Die **physikalische Bedeutung** ist: sei  $u_0(x)$  die Temperatur zur Zeit  $t = 0$  am Ort  $x$ , dann ist  $u(t, x)$  die Temperatur zur Zeit  $t$  am Ort  $x$  bei einem ortsunabhängigen Wärmeleitungskoeffizienten. Das bedeutet, dass die Funktion  $u(t, x)$  folgende Wärmeleitungsgleichung löst:

$$\frac{\partial u}{\partial t}(t, x) = \frac{\partial^2 u}{\partial x^2}(t, x), \quad u(0, x) = u_0(x). \quad (\heartsuit)$$

Man nennt die Funktion  $G$  auch *Fundamentallösung* der Wärmeleitungsgleichung. Um nachzuprüfen, dass unser  $u$  tatsächlich die Wärmeleitungsdifferentialgleichung  $(\heartsuit)$  löst, muss das Parameterintegral in  $(\spadesuit)$  einmal nach  $t$  bzw. zweimal nach  $x$  differenziert werden.

Wir haben also die Frage noch zu beantworten, wie man Parameterintegrale nach dem Parameter differenziert.

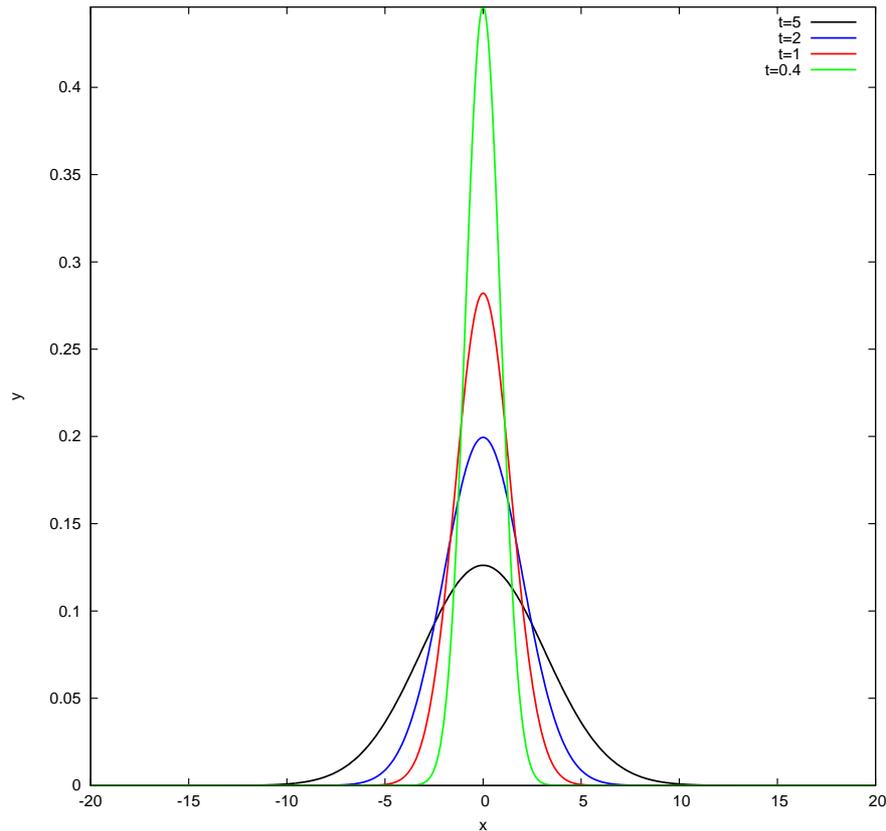


Abbildung 1.1: Die Gaußglockenfunktion  $G(t, x)$  als Funktion von  $x$ , für verschiedene Werte von  $t$ . Der Flächeninhalt unter dem Graphen ist immer gleich Eins.

**Satz 1.3 (Integration von Parameterintegralen):**

Seien  $a, b, c, d \in \mathbb{R}$ . Wenn sowohl für jedes  $x \in [c, d]$  die Funktionen  $y \mapsto f(x, y)$  als auch für jedes  $y \in [a, b]$  die Funktionen  $x \mapsto f(x, y)$  stückweise stetig sind, so gilt

$$\int_{x=c}^d g(x) dx = \int_{x=c}^d \int_{y=a}^b f(x, y) dy dx = \int_{y=a}^b \int_{x=c}^d f(x, y) dx dy,$$

was bedeutet, dass die beiden Integral getauscht werden dürfen.

**Beispiele 1.4:**

Wir greifen Beispiel 1.2 noch mal auf.

(a) Es ist (zur Verdeutlichung setzen wir Extra-Klammern)

$$\begin{aligned} \int_{x=0}^1 g(x) dx &= \int_{x=0}^1 \left( \int_{y=0}^2 ye^{xy} dy \right) dx && \text{nutze Satz 1.3} \\ &= \int_{y=0}^2 \left( \int_{x=0}^1 ye^{xy} dx \right) dy = \int_{y=0}^2 y \cdot \left( \int_{x=0}^1 e^{xy} dx \right) dy \\ &= \int_{y=0}^2 y \cdot \left( \frac{1}{y} e^{xy} \Big|_{x=0}^{x=1} \right) dy = \int_{y=0}^2 e^{xy} \Big|_{x=0}^{x=1} dy = \int_{y=0}^2 (e^y - 1) dy \\ &= (e^y - y) \Big|_{y=0}^{y=2} = e^2 - 2 - 1 = e^2 - 3. \end{aligned}$$

(b) Wir betrachten die Funktion  $u(t, x)$ , die gegeben ist durch

$$u(t, x) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}} u_0(y) dy.$$

Die Gesamt-Wärmemenge zur Zeit  $t$  berechnen wir dann wie folgt:

$$\begin{aligned} \int_{x=-\infty}^{\infty} u(t, x) dx &= \int_{x=-\infty}^{\infty} \left( \frac{1}{\sqrt{4\pi t}} \int_{y=-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}} u_0(y) dy \right) dx && \left| \text{nutze Satz 1.3} \right. \\ &= \int_{y=-\infty}^{\infty} \left( \int_{x=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}} u_0(y) dx \right) dy \\ &= \int_{y=-\infty}^{\infty} u_0(y) \underbrace{\left( \int_{x=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}} dx \right)}_{=1 \quad (\text{Eigenschaft der Gau\ss-Glocke})} dy \\ &= \int_{y=-\infty}^{\infty} u_0(y) dy, \end{aligned}$$

und das ist aber die Gesamtwärmemenge zur Zeit 0. Auf diesem Wege haben wir die Erhaltung der Energie durch Nachrechnen wiederentdeckt !

**Satz 1.5 (Differentiation von Parameterintegralen):**

Sei  $f \in: \mathbb{R}^2 \rightarrow \mathbb{R}$  und seien  $a, b \in \mathbb{R}$ . Außerdem sei wieder für jedes  $x \in \mathbb{R}$  die Funktion  $y \mapsto f(x, y)$  stückweise stetig, sodass das Parameterintegral

$$g(x) = \int_{y=a}^b f(x, y) dy$$

existiert. Wenn zusätzlich  $f$  nach  $x$  stetig partiell differenzierbar ist, dann ist auch die Funktion  $g$  differenzierbar, und es ist

$$g'(x) = \int_{y=a}^b \frac{\partial f}{\partial x}(x, y) dy,$$

also dürfen Differentiation nach dem Parameter  $x$  und Integration bezüglich der Variablen  $y$  vertauscht werden.

**Beispiel 1.6:**

Es sei

$$g(x) := \int_{y=1}^2 \frac{e^{xy} - e^y}{y} dy.$$

Da  $f(x, y) = \frac{e^{xy} - e^y}{y}$  stetig partiell nach  $x$  differenzierbar ist (beachte, dass wir wegen  $1 \leq y \leq 2$  keine Gefahr laufen, durch Null zu dividieren), gilt für  $x \neq 0$ , dass

$$\begin{aligned} g'(x) &= \int_{y=1}^2 \frac{\partial}{\partial x} \left( \frac{e^{xy} - e^y}{y} \right) dy = \int_{y=1}^2 \frac{1}{y} \cdot y \cdot e^{xy} dy = \int_{y=1}^2 e^{xy} dy && \left| \text{benutze jetzt } x \neq 0 \right. \\ &= \frac{1}{x} e^{xy} \Big|_{y=1}^{y=2} = \frac{1}{x} (e^{2x} - e^x). \end{aligned}$$

**Bemerkung 1.7 (Variable Integrationsgrenzen):**

Jetzt sei

$$g(x) = \int_{y=x+1}^{y=x^2} \frac{e^{xy} - e^y}{y} dy, \quad x > 0,$$

und wir suchen  $g'(x)$ . Wir beobachten, dass das Integral auf der rechten Seite an drei Stellen vom Parameter  $x$  abhängt:

- im Integranden,
- in der oberen Grenze,
- in der unteren Grenze.

Zur Auswertung der Funktion  $g$  benutzen wir eine Hilfsfunktion  $G$ , gegeben durch

$$G(x, u, v) := \int_{y=u}^{y=v} \frac{e^{xy} - e^y}{y} dy,$$

und dann ist  $g(x) = G(x, x+1, x^2)$ . Gemäß Kettenregel gilt dann

$$g'(x) = \nabla G(x, x+1, x^2) \cdot \begin{bmatrix} (x)' \\ (x+1)' \\ (x^2)' \end{bmatrix} = \nabla G(x, x+1, x^2) \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix}.$$

Hierbei bedeutet  $\nabla G(x, x+1, x^2)$ , dass wir zuerst  $\nabla$  auf  $G(x, u, v)$  anwenden, also

$$\nabla G(x, u, v) = \left[ \frac{\partial G}{\partial x} \quad \frac{\partial G}{\partial u} \quad \frac{\partial G}{\partial v} \right]$$

ermitteln, und danach setzen wir  $(x, x+1, x^2)$  ein für  $(x, u, v)$ . Das machen wir jetzt. Wegen Beispiel 1.6 haben wir

$$\begin{aligned} \frac{\partial G}{\partial x}(x, u, v) &= \int_{y=u}^v \frac{\partial}{\partial x} \left( \frac{e^{xy} - e^y}{y} \right) dy = \int_{y=u}^v \frac{1}{y} \cdot y \cdot e^{xy} dy = \int_{y=u}^v e^{xy} dy \\ &= \frac{1}{x} e^{xy} \Big|_{y=u}^{y=v} = \frac{e^{vx} - e^{ux}}{x}. \end{aligned}$$

Nach Hauptsatz der Differential- und Integralrechnung ist

$$\frac{\partial G}{\partial u}(x, u, v) = -\frac{e^{xu} - e^u}{u}, \quad \frac{\partial G}{\partial v}(x, u, v) = +\frac{e^{xv} - e^v}{v}.$$

Und jetzt setzen wir anstatt  $(x, u, v)$  wieder  $(x, x+1, x^2)$  ein:

$$\begin{aligned} g'(x) &= \nabla G(x, x+1, x^2) \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix} \\ &= \left[ \frac{e^{x^2 \cdot x} - e^{(x+1) \cdot x}}{x} \quad -\frac{e^{x \cdot (x+1)} - e^{x+1}}{x+1} \quad \frac{e^{x \cdot x^2} - e^{x^2}}{x^2} \right] \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix} \\ &= \frac{1}{x} (e^{x^3} - e^{x^2} e^x) - \frac{1}{x+1} (e^{x^2} e^x - e^x e) + \frac{2}{x} (e^{x^3} - e^{x^2}) \\ &= \frac{1}{x} (3e^{x^3} - e^{x^2} (e^x + 2)) - \frac{e^x}{x+1} (e^{x^2} - e). \end{aligned}$$

## 2 Kurvenintegrale

### Bemerkung 2.1 (Motivation):

Wenn ein Teilchen entlang einer geraden Strecke mit gleichförmiger Geschwindigkeit gezogen wird und wenn dabei gegen eine konstante Reibungskraft gearbeitet werden muss, dann ist bekanntlich die geleistete mechanische Arbeit gleich dem Produkt aus Kraft und Weglänge.

Jetzt wollen wir dies verallgemeinern, siehe Bild 2.2:

- der Weg braucht nicht mehr eine gerade Strecke sein — auch eine geschwungene Kurve ist jetzt erlaubt;
- die zu überwindende Kraft braucht nicht mehr konstant sein — jetzt wollen wir auch eine variable Kraft zulassen.

Anschaulich bedeutet dies: wir werden die Kraftfunktion  $\vec{f}$  entlang einer Kurve integrieren.

Anschaulich ist auch klar: eine Kurve entspricht einem verbogenen Stück des  $\mathbb{R}$ , das man in den  $\mathbb{R}^2$  oder den  $\mathbb{R}^n$  hineingelegt hat.

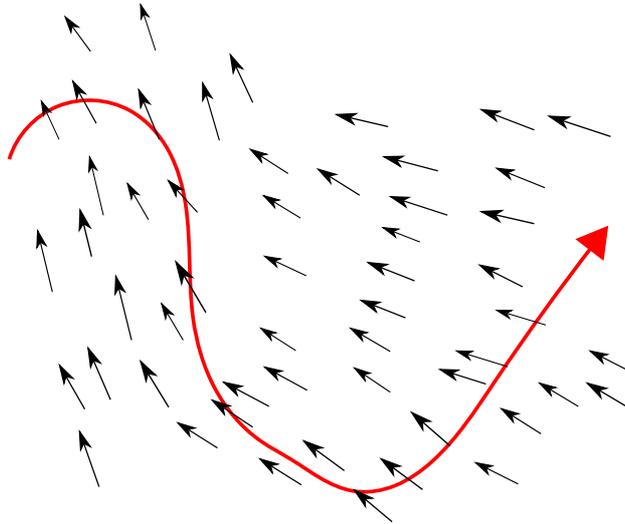


Abbildung 2.2: Ein Teilchen soll entlang des roten Wegs gezogen werden, und dabei ist die angegebene Kraft zu überwinden.

**Definition 2.2 (Kurven):**

Eine Funktion  $\gamma: [a, b] \rightarrow \mathbb{R}^n$ , die stetig und stückweise differenzierbar ist, heißt *Kurve im  $\mathbb{R}^n$* .

Die Punktmenge

$$\text{spur}(\gamma) := \{\gamma(t) : a \leq t \leq b\}$$

heißt *Spur der Kurve*, und  $\gamma(a) \in \mathbb{R}^n$  heißt *Anfangspunkt der Kurve*. Entsprechend heißt  $\gamma(b) \in \mathbb{R}^n$  *Endpunkt der Kurve*.

Man erhält eine gute Vorstellung, wenn man  $t$  als die Zeitvariable liest und  $\gamma(t)$  als den Ort innerhalb des  $\mathbb{R}^n$ , an dem ein Teilchen zur Zeit  $t$  ist.

Was ist der Unterschied zwischen einer Kurve  $\gamma$  und  $\text{spur}(\gamma)$ ? Wir überlegen uns, dass jedes mathematische Objekt einen Typ hat (wir kennen schon die Typen „reelle Zahl“, „komplexe Zahl“, „Vektor“, „Matrix“ usw.). Die Kurve  $\gamma$  hat den Typ „Funktion von  $[a, b]$  in den  $\mathbb{R}^n$ “, und  $\text{spur}(\gamma)$  hat den Typ „Punktmenge des  $\mathbb{R}^n$ “. Wir stellen uns ein Auto vor, dass mit nassen Reifen über eine Straße fährt. Dann ist  $[a, b]$  das Zeitintervall, und  $\gamma$  gibt an, wo das Auto wann ist. Im Gegensatz dazu ist  $\text{spur}(\gamma)$  die Spur der nassen Reifen auf dem Asphalt. Die Information, wann das Auto wo war, ist in  $\text{spur}(\gamma)$  nicht mehr enthalten.

**Beispiele 2.3:** (a) Seien  $\vec{x}, \vec{y} \in \mathbb{R}^n$ . Dann beschreibt die Funktion  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$ , gegeben durch die Gleichung

$$\gamma(t) := (1 - t)\vec{x} + t\vec{y},$$

die *Verbindungsstrecke* von  $\vec{x}$  nach  $\vec{y}$ . Das Teilchen beginnt seine Fahrt zur Zeit 0 im Punkt  $\vec{x}$  und erreicht den Punkt  $\vec{y}$  zur Zeit 1.

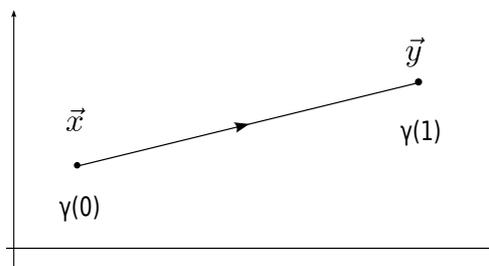


Abbildung 2.3: Eine Verbindungsstrecke.

(b) Die Funktion

$$\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) := \begin{bmatrix} r \cos(t) \\ r \sin(t) \end{bmatrix}$$

beschreibt einen im Gegenuhrzeigersinn durchlaufenen Kreis vom Radius  $r$ . Wir haben  $\gamma(0) = [r, 0]^T = \gamma(2\pi)$ .

Eine Kurve  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  mit  $\gamma(a) = \gamma(b)$  heißt *geschlossene Kurve*.

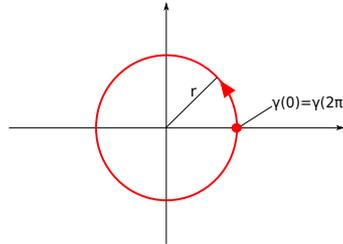


Abbildung 2.4: Ein Kreis vom Radius  $r$ .

(c) Die Funktion

$$\gamma: [0, \pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

beschreibt den Einheitshalbkreis in der oberen Halbebene, denn es ist  $(\cos(t))^2 + (\sin(t))^2 = 1$  und  $\sin(t) \geq 0$  für  $t \in [0, \pi]$ .

Die Funktion

$$\hat{\gamma}: [-1, 1] \rightarrow \mathbb{R}^2, \quad \hat{\gamma}(t) := \begin{bmatrix} t \\ \sqrt{1-t^2} \end{bmatrix}$$

beschreibt ebenfalls den Einheitshalbkreis in der oberen Halbebene, denn es ist  $t^2 + (\sqrt{1-t^2})^2 = 1$  und  $\sqrt{1-t^2} \geq 0$ . Wir erkennen also

$$\text{spur}(\gamma) = \text{spur}(\hat{\gamma}),$$

aber die Kurven  $\gamma$  und  $\hat{\gamma}$  sind sehr unterschiedlich.

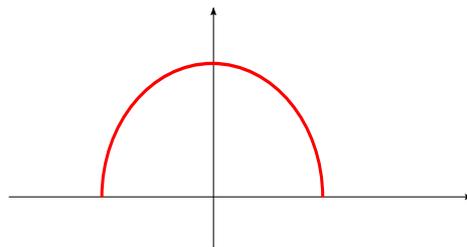
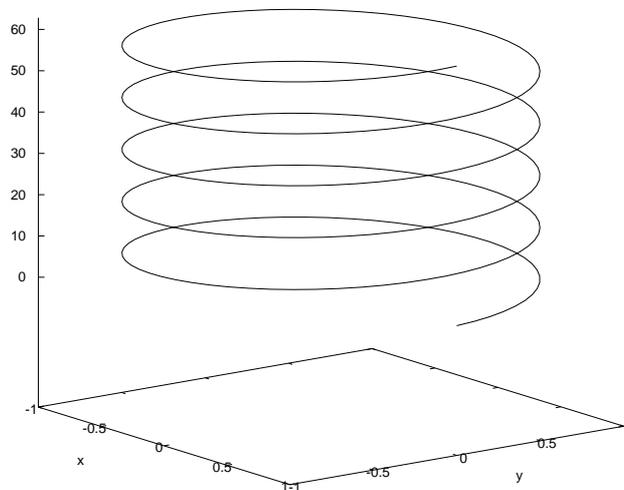


Abbildung 2.5: Der Einheitshalbkreis in der oberen Halbebene.

(d) Die Funktion

$$\gamma: [0, 10\pi] \rightarrow \mathbb{R}^3, \quad \gamma(t) := \begin{bmatrix} r \cos(t) \\ r \sin(t) \\ ct \end{bmatrix}$$

beschreibt eine Schraubenlinie im  $\mathbb{R}^3$ , siehe Bild 2.6.

Abbildung 2.6: Eine Schraubenlinie mit  $r = 1$  und  $c = 2$ .**Bemerkung 2.4 (Tangente an eine Kurve):**

Sei  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  eine Kurve und  $t_0 \in (a, b)$  ein Zeitpunkt im Inneren des Intervalls  $[a, b]$ .

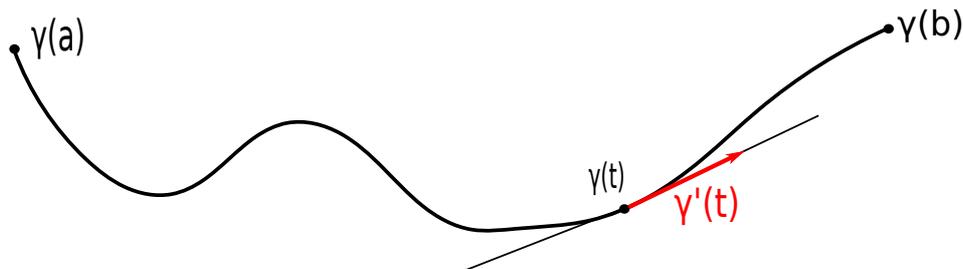
Die lineare Approximation im Punkt  $\gamma(t_0)$  durch ein Taylorpolynom ist

$$\gamma(t_0) + D\gamma(t_0) \cdot (t - t_0) = \underbrace{\gamma(t_0)}_{\text{Aufpunkt}} + (t - t_0) \cdot \underbrace{D\gamma(t_0)}_{\text{Richtungsvektor}}.$$

Hierbei ist  $D\gamma$  die totale Ableitung der Funktion  $\gamma$ . Anstatt  $D\gamma(t_0)$  schreiben wir üblicherweise  $\gamma'(t_0)$ , und es ist

$$\gamma'(t_0) = \begin{bmatrix} \gamma'_1(t_0) \\ \gamma'_2(t_0) \\ \vdots \\ \gamma'_n(t_0) \end{bmatrix}.$$

Die **Anschauung** dazu ist: für das Teilchen, das auf  $\text{spur}(\gamma)$  entlang fliegt, ist zur Zeit  $t_0$  der Geschwindigkeitsvektor gleich  $\gamma'(t_0)$ , und der Betrag der Geschwindigkeit ist  $\|\gamma'(t_0)\|$ . Der Geschwindigkeitsvektor verläuft tangential an  $\text{spur}(\gamma)$ .

Abbildung 2.7: Die Spur einer Kurve mit einem Tangentenvektor  $\gamma'(t)$  am Punkt  $\gamma(t)$ .

**Beispiele 2.5:** (a) Für die Verbindungsstrecke von  $\vec{x}$  und  $\vec{y} \in \mathbb{R}^n$  ist  $\gamma(t) = (1-t)\vec{x} + t\vec{y}$ . Dann folgt

$$\gamma'(t) = -\vec{x} + \vec{y} = \vec{y} - \vec{x}.$$

Dieser Ausdruck hängt von der Zeit  $t$  gar nicht ab, es handelt sich also um eine gleichförmige lineare Bewegung. Innerhalb von einer Zeiteinheit wandert das Teilchen von  $\vec{x}$  nach  $\vec{y}$ , also ist der Betrag der Geschwindigkeit tatsächlich  $\frac{\|\vec{y}-\vec{x}\|}{1} = \|\gamma'(t)\|$ , wie vorhergesagt (bei gleichförmigen linearen Bewegungen ergibt sich der Geschwindigkeitsbetrag tatsächlich durch Division von Gesamtweglänge durch Gesamtzeit).

(b) Wir betrachten ein Teilchen auf dem Einheitskreis. Dann ist  $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$  mit

$$\gamma(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \quad \text{also} \quad \gamma'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}.$$

Die Länge des Tangentenvektors  $\gamma'(t)$  ist dann

$$\|\gamma'(t)\| = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

Es handelt sich um eine gleichmäßige Kreisbewegung. Die Richtung, in die  $\gamma'(t)$  zeigt, ist für jedes  $t$  eine andere.

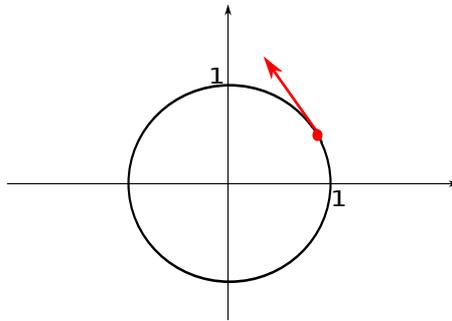


Abbildung 2.8: Der Einheitskreis mit einem Tangentenvektor.

**Bemerkung 2.6 (Länge einer Kurve):**

Sei  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  eine Kurve. Die zurückgelegte Strecke ist das Integral des Geschwindigkeitsbetrags über der Zeit, also ist die Länge der Spur von  $\gamma$  gleich

$$L(\gamma) := \int_a^b \|\gamma'(t)\| dt.$$

Achtung: wenn man von der Geschwindigkeit redet, dann meint man im Allgemeinen den Geschwindigkeitsvektor  $\gamma'(t) \in \mathbb{R}^n$ . Hier jedoch integrieren wir über den Geschwindigkeitsbetrag, welcher eine Zahl ( $\geq 0$ ) ist.

**Beispiele 2.7:** (a) Für die Kreisbahn mit Radius  $r$  haben wir  $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$  mit

$$\gamma(t) = \begin{bmatrix} r \cos(t) \\ r \sin(t) \end{bmatrix}, \quad r > 0.$$

Dann folgt

$$\begin{aligned} L(\gamma) &= \int_0^{2\pi} \|\gamma'(t)\| dt = \int_0^{2\pi} \left\| \begin{bmatrix} -r \sin t \\ r \cos t \end{bmatrix} \right\| dt = \int_0^{2\pi} \sqrt{r^2 \sin^2(t) + r^2 \cos^2(t)} dt \\ &= \int_0^{2\pi} r \sqrt{\sin^2(t) + \cos^2(t)} dt = \int_0^{2\pi} r dt = 2\pi r. \end{aligned}$$

Das entspricht unseren Erwartungen.

(b) Wir betrachten die Schraubenlinie auf dem Zeitintervall  $[0, \ell]$ . Dann ist  $\gamma: [0, \ell] \rightarrow \mathbb{R}^3$  mit

$$\gamma(t) = \begin{bmatrix} r \cos(t) \\ r \sin(t) \\ ct \end{bmatrix}, \quad \gamma'(t) = \begin{bmatrix} -r \sin(t) \\ r \cos(t) \\ c \end{bmatrix},$$

wobei wir natürlich  $\ell, r, c > 0$  haben. Es folgt dann

$$\begin{aligned} L(\gamma) &= \int_0^\ell \|\gamma'(t)\| dt = \int_0^\ell \sqrt{r^2 \sin^2(t) + r^2 \cos^2(t) + c^2} dt \\ &= \int_0^\ell \sqrt{r^2 + c^2} dt = \ell \cdot \sqrt{r^2 + c^2}. \end{aligned}$$

Dies hätten wir auch herausfinden können, indem wir den Zylindermantel, auf dem die Schraubenlinie aufgewickelt ist, verzerrungsfrei abwickeln.

**Definition 2.8 (Arbeitsintegral):**

Sei  $M \subseteq \mathbb{R}^n$ ,  $f: M \rightarrow \mathbb{R}^n$  ein Vektorfeld, und sei  $\gamma: [a, b] \rightarrow M$  eine Kurve. Dann heißt

$$\int_\gamma f(\vec{x}) d\vec{x} := \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle dt$$

das *Arbeitsintegral von  $f$  längs  $\text{spur}(\gamma)$*  oder auch *Kurvenintegral von  $f$  längs  $\text{spur}(\gamma)$* .

Wir stellen uns  $f$  als Kraftfeld vor und  $\text{spur}(\gamma)$  als den Weg des Teilchens. Siehe auch Bild 2.2.

**Bemerkung 2.9:** (a) Es ist auch die Notation

$$\int_a^b f(\gamma(t)) \cdot \gamma'(t) dt$$

gebräuchlich. Hierbei bezeichnet der Multiplikationspunkt  $\cdot$  das Skalarprodukt zweier Vektoren mit  $n$  Einträgen.

(b) Der Wert  $\int_\gamma f(\vec{x}) d\vec{x}$  hängt nur von  $\text{spur}(\gamma)$  ab und der Durchlaufrichtung, nicht aber vom konkreten  $\gamma$ . Das bedeutet:

- Wenn  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  und  $\hat{\gamma}: [c, d] \rightarrow \mathbb{R}^n$  Kurven sind mit  $\text{spur}(\gamma) = \text{spur}(\hat{\gamma})$  sowie  $\gamma(a) = \hat{\gamma}(c)$  und  $\gamma(b) = \hat{\gamma}(d)$ , dann ist  $\int_\gamma f(\vec{x}) d\vec{x} = \int_{\hat{\gamma}} f(\vec{x}) d\vec{x}$ .
- Wenn allerdings  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  und  $\hat{\gamma}: [c, d] \rightarrow \mathbb{R}^n$  Kurven sind mit  $\text{spur}(\gamma) = \text{spur}(\hat{\gamma})$  sowie  $\gamma(a) = \hat{\gamma}(d)$  und  $\gamma(b) = \hat{\gamma}(c)$  (wenn also die Durchlaufrichtung umgekehrt wurde), dann gilt  $\int_\gamma f(\vec{x}) d\vec{x} = - \int_{\hat{\gamma}} f(\vec{x}) d\vec{x}$ .

Wir kommen zu einem neuen Gedanken.

**Beispiele 2.10:**

Wir integrieren ein Vektorfeld  $f$  entlang zweier unterschiedlicher Kurven  $\gamma$  und  $\hat{\gamma}$ , für die auch deren Spuren unterschiedlich sind. Siehe Bild 2.9.

(a) Sei

$$f(x, y) := \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix}, \quad \gamma(t) := \begin{bmatrix} t \cos(t) \\ t \sin(t) \end{bmatrix}, \quad t \in \left[0, \frac{\pi}{2}\right].$$

Dann haben wir

$$\begin{aligned} \int_{\gamma} f(\vec{x}) \, dx &= \int_0^{\pi/2} \langle f(\gamma(t)), \gamma'(t) \rangle \, dt \\ &= \int_0^{\pi/2} \left\langle \begin{bmatrix} 2t^2 \cos(t) \sin(t) \\ t^2 \cos^2(t) + t^2 \sin^2(t) \end{bmatrix}, \begin{bmatrix} \cos(t) - t \sin(t) \\ \sin(t) + t \cos(t) \end{bmatrix} \right\rangle \, dt \quad \left| \text{nutze } \cos^2(t) + \sin^2(t) = 1 \right. \\ &= \int_0^{\pi/2} (2t^2 \cos^2(t) \sin(t) - 2t^3 \cos(t) \sin^2(t) + t^2 \sin(t) + t^3 \cos(t)) \, dt \\ &= (\text{man rechnet einige Zeit}) = \frac{\pi^3}{24}. \end{aligned}$$

(b) Wir nehmen dieselbe Funktion  $f$ , jetzt allerdings die Kurve  $\hat{\gamma}$  mit  $\hat{\gamma}(t) := [0, t]^\top$  für  $t \in [0, \frac{\pi}{2}]$ . Wir beobachten  $\hat{\gamma}(0) = [0, 0]^\top = \gamma(0)$  und  $\hat{\gamma}(\frac{\pi}{2}) = [0, \frac{\pi}{2}]^\top = \gamma(\frac{\pi}{2})$ . Es haben also  $\text{spur}(\hat{\gamma})$  und  $\text{spur}(\gamma)$  dieselben Anfangs- und Endpunkte. Dann rechnen wir wie folgt:

$$\int_{\hat{\gamma}} f(\vec{x}) \, d\vec{x} = \int_0^{\pi/2} \langle f(\hat{\gamma}(t)), \hat{\gamma}'(t) \rangle \, dt = \int_0^{\pi/2} \left\langle \begin{bmatrix} 0 \\ t^2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle \, dt = \int_0^{\pi/2} t^2 \, dt = \frac{1}{3} t^3 \Big|_{t=0}^{t=\pi/2} = \frac{\pi^3}{24}.$$

Die Gleichheit der Integrale ist kein Zufall.

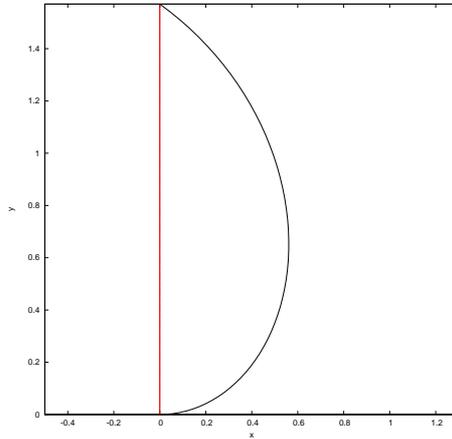


Abbildung 2.9:  $\text{spur}(\gamma)$  ist der schwarze Bogen,  $\text{spur}(\hat{\gamma})$  die rote Strecke.

### Definition 2.11 (Potential):

Sei  $M \subseteq \mathbb{R}^n$  und  $f: M \rightarrow \mathbb{R}^n$  ein Vektorfeld. Dann heißt ein Skalarfeld  $\varphi: M \rightarrow \mathbb{R}$  *Potential von  $f$* , falls gilt:

- $\varphi$  ist stetig partiell differenzierbar,
- $\nabla \varphi(\vec{x}) = f(\vec{x})^\top$  für jedes  $\vec{x} \in M$ .

In diesem Falle nennen wir  $f$  ein *Potentialfeld* oder ein *konservatives Feld* oder ein *Gradientenfeld*.

### Bemerkung 2.12 (Kurvenintegrale von Potentialfeldern):

Sei  $M \subseteq \mathbb{R}^n$  und  $f: M \rightarrow \mathbb{R}^n$  ein Potentialfeld mit Potential  $\varphi: M \rightarrow \mathbb{R}$ , das heißt  $f^\top = \nabla \varphi$ . Dann gilt für jede Kurve  $\gamma: [a, b] \rightarrow M$ , dass

$$\left( \varphi(\gamma(t)) \right)' \stackrel{\text{Kettenregel}}{=} \nabla \varphi(\gamma(t)) \cdot \gamma'(t) = f(\gamma(t))^\top \cdot \gamma'(t) = \langle f(\gamma(t)), \gamma'(t) \rangle,$$

also ergibt sich dann

$$\int_{\gamma} f(\vec{x}) \, d\vec{x} = \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt = \int_a^b \left( \varphi(\gamma(t)) \right)' \, dt \stackrel{\text{Hauptsatz}}{=} \varphi(\gamma(b)) - \varphi(\gamma(a)).$$

Das bedeutet: Wenn  $f$  ein Potentialfeld ist, dann hängt der Wert des Kurvenintegrals nur von der Lage der Anfangs- und Endpunkte der Kurve ab, aber nicht vom Verlauf der Spur dazwischen!

Man sagt auch: das Potential  $\varphi$  ist die *Stammfunktion* des Potentialfelds  $f$ .

Ein **wichtiger Spezialfall**: wenn  $f$  ein Potentialfeld ist und  $\gamma$  geschlossen, so ist

$$\int_{\gamma} f(\vec{x}) \, d\vec{x} = 0.$$

**Beispiele 2.13:** (a) Wir kehren zurück zu Beispiel 2.10. Dort war

$$f(x, y) = \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix},$$

und dies ist ein Potentialfeld, denn für die (im Moment noch erratene) Funktion  $\varphi(x, y) = x^2y + \frac{1}{3}y^3$  gilt tatsächlich

$$\nabla\varphi(x, y) = [2xy, x^2 + y^2] = f(x, y)^{\top}.$$

Dies erklärt, warum die beiden Integrale gleich waren.

(b) Auf der Menge  $M := \mathbb{R}^2 \setminus \{\vec{0}\}$  betrachten wir die Funktion

$$f(x, y) = \begin{bmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{bmatrix}.$$

Wir haben bei der Wahl von  $M$  den Nullpunkt ausgeschlossen, weil wir ansonsten durch Null dividiert hätten. Die Spur der Kurve sei der im Gegenuhrzeigersinn durchlaufene Einheitskreis, also

$$\gamma(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \quad t \in [0, 2\pi].$$

Es ist also  $\gamma$  eine geschlossene Kurve. Trotzdem ist unser Kurvenintegral nicht gleich Null, denn

$$\begin{aligned} \int_{\gamma} f(\vec{x}) \, d\vec{x} &= \int_0^{2\pi} \langle f(\gamma(t)), \gamma'(t) \rangle \, dt \\ &= \int_0^{2\pi} \begin{bmatrix} \frac{-\sin(t)}{\cos^2(t)+\sin^2(t)} & \frac{\cos(t)}{\cos^2(t)+\sin^2(t)} \end{bmatrix} \cdot \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \, dt \\ &= \int_0^{2\pi} (\sin^2(t) + \cos^2(t)) \, dt = 2\pi \neq 0 ! \end{aligned}$$

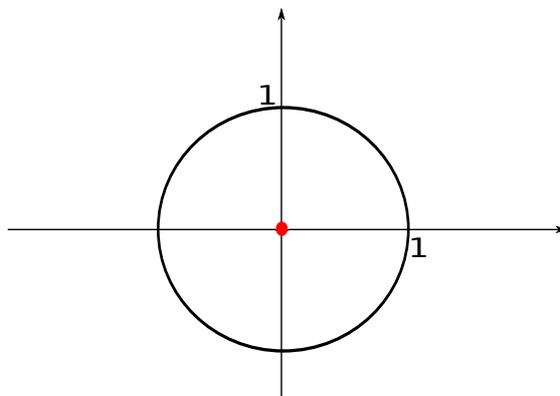


Abbildung 2.10: Der Integrationsweg für Beispiel 2.13 (b).

**Bemerkung 2.14 (Notwendige Bedingung für die Existenz eines Potentials):**

Sei  $M \subseteq \mathbb{R}^n$  und  $f = [f_1, \dots, f_n]^\top : M \rightarrow \mathbb{R}^n$  ein stetig differenzierbares Vektorfeld. Wenn  $f$  ein Potential  $\varphi$  besitzt, dann ist für jegliche  $i, j = 1, \dots, n$

$$f_i = \frac{\partial \varphi}{\partial x_i}, \quad f_j = \frac{\partial \varphi}{\partial x_j}.$$

Der Satz von Schwarz liefert dann (falls  $i \neq j$ )

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \frac{\partial^2 \varphi}{\partial x_j \partial x_i} = \frac{\partial f_j}{\partial x_i}.$$

Eine **Mindestvoraussetzung** für die Existenz eines Potentials ist also die **Integrabilitätsbedingung**

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}, \quad \text{für alle } i \neq j, \quad \text{bzw. die Jacobi-Matrix } J_f \text{ ist symmetrisch.}$$

Umgangssprachlich: „die Ableitungen über Kreuz müssen gleich sein“ !

**Beispiele 2.15:** (a) Wir kehren zu Beispiel 2.10 zurück. Da hatten wir

$$f(x, y) = \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix} = \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix}.$$

Wir prüfen die Integrabilitätsbedingung nach:

$$\frac{\partial f_1}{\partial y} \stackrel{?}{=} \frac{\partial f_2}{\partial x},$$

und dies stimmt tatsächlich, denn es ist  $2x = 2x$ . Also ergibt es Sinn, ein Potential  $\varphi$  zu suchen (von dem wir aber noch nicht wissen, ob es überhaupt existiert). Es soll gelten  $\frac{\partial \varphi}{\partial x} = 2xy$ , also folgt

$$\varphi(x, y) = \int 2xy \, dx + g(y) = x^2 y + c + g(y).$$

Hierbei sind  $c$  eine wählbare Konstante und  $g$  eine (noch) wählbare Funktion. Dann haben wir tatsächlich  $\frac{\partial \varphi}{\partial x} = 2xy$ , und wir wollen haben

$$x^2 + y^2 \stackrel{!}{=} \frac{\partial \varphi}{\partial y} = x^2 + 0 + g'(y).$$

Also muss  $g'(y) = y^2$  sein, d.h.  $g(y) = \frac{1}{3}y^3 + C$ . Zusammengefasst bekommen wir z.B.

$$\varphi(x, y) = x^2 y + \frac{1}{3}y^3,$$

wenn wir  $c = C = 0$  setzen.

(b) Potentiale sind etwas Seltenes und Kostbares ! Sei z.B.

$$f(x, y) = \begin{bmatrix} x \\ x + y \end{bmatrix}.$$

Dann ist  $\frac{\partial f_1}{\partial y} = 0 \neq 1 = \frac{\partial f_2}{\partial x}$ , also kann diese harmlos ausschauende Funktion gar kein Potential besitzen.

(c) Wir setzen Beispiel 2.13 (b) fort. Es war

$$f(x, y) = \begin{bmatrix} \frac{-y}{x^2 + y^2} \\ \frac{x}{x^2 + y^2} \end{bmatrix} = \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix}.$$

Wir testen, ob die Integrabilitätsbedingung erfüllt ist:

$$\begin{aligned} \frac{\partial f_1}{\partial y} &= \frac{-(x^2 + y^2) + 2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}, \\ \frac{\partial f_2}{\partial x} &= \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}, \end{aligned}$$

und dieser Test ist erfolgreich. Also suchen wir ein Potential  $\varphi$  auf ähnlichem Rechenweg wie eben, und es entsteht

$$\varphi(x, y) = \arctan\left(\frac{y}{x}\right) \quad \text{für } x \neq 0.$$

Wir haben jetzt ein Potential  $\varphi$ , aber trotzdem haben wir  $\int_{\gamma} f(\vec{x}) d\vec{x} \neq 0$  ausgerechnet. Der Grund ist, dass der Einheitskreis, den wir beim Integrieren entlang  $\gamma$  ablaufen, zwei Punkte besitzt mit  $x = 0$ , und dort haben wir leider kein Potential.

Der Integrand  $f$  besitzt im Ursprung  $[0, 0]^T$  eine Polstelle (auch Singularität genannt), und diese Singularität „strahlt aus“.

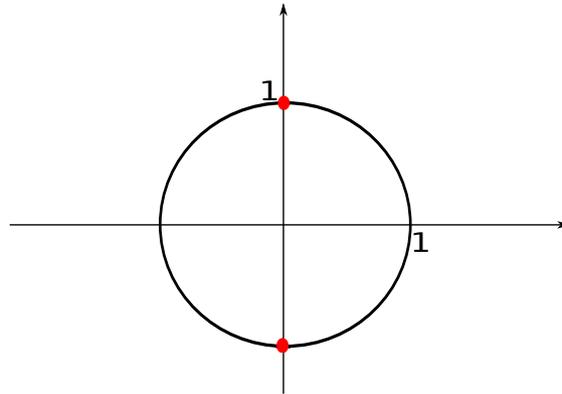


Abbildung 2.11: Der Integrationsweg mit zwei Stellen ohne Potential.

Wir erkennen: die Integrabilitätsbedingung ist zwar notwendig dafür, dass das Vektorfeld  $f$  ein Potential besitzt überall auf  $\text{spur}(\gamma)$ , aber nicht hinreichend. Eine Abhilfe geben wir im nächsten Satz.

**Satz 2.16:**

Sei  $M \subseteq \mathbb{R}^n$  und das Vektorfeld  $f: M \rightarrow \mathbb{R}^n$  erfülle die Integrabilitätsbedingung. Wenn die Menge  $M$  offen ist und *einfach zusammenhängend*, dann besitzt  $f$  auf ganz  $M$  ein Potential.

Hierbei bedeutet „einfach zusammenhängend“, dass  $M$  keine Löcher hat.

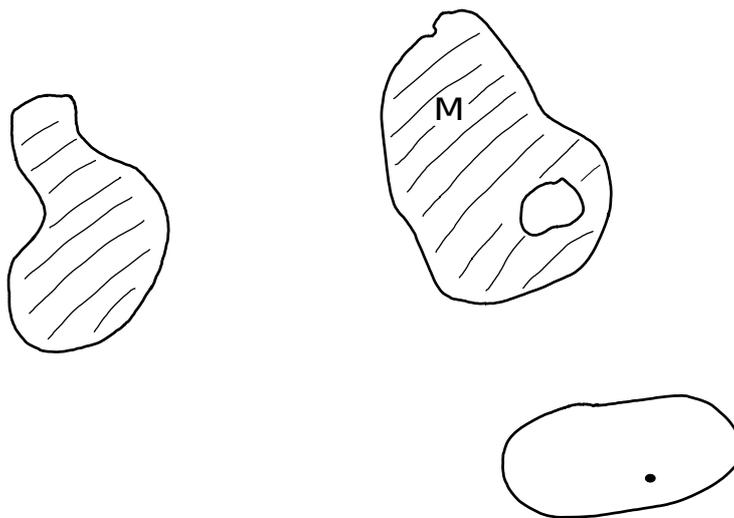


Abbildung 2.12: Dies sind drei Mengen im  $\mathbb{R}^2$ . Die linke Menge ist einfach zusammenhängend. Die Menge rechts oben ist es nicht, denn sie hat ein Loch. Die Menge rechts unten ist auch nicht einfach zusammenhängend, denn bei ihr ist ein Punkt herausgestanzt worden.

### 3 Integrale im $\mathbb{R}^n$

Unser **Ziel** ist die Integration von Funktionen in mehreren Variablen. Ein **Problem** kann dabei darin liegen, dass die Integrationsgebiete (also diejenigen Mengen, wo die Integrationsvariable  $\vec{x}$  drin lebt) relativ wild sein können.

**Definition 3.1 (Intervalle im  $\mathbb{R}^n$ ):**

Seien  $\vec{a}, \vec{b} \in \mathbb{R}^n$  mit  $a_i \leq b_i$  für alle  $i = 1, \dots, n$ . Dann heißt

$$I := [\vec{a}, \vec{b}] := \{\vec{x} \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ für alle } i = 1, \dots, n\}$$

abgeschlossenes Intervall im  $\mathbb{R}^n$ .

Dann ist  $I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$ .

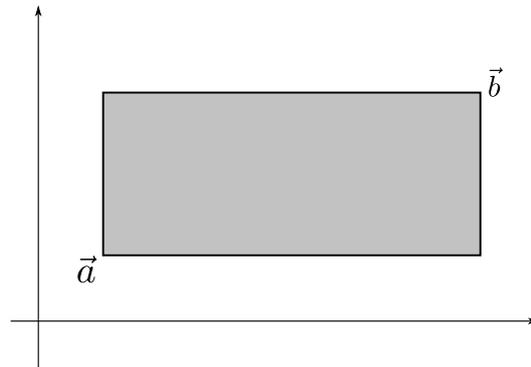


Abbildung 3.13: Ein Intervall im  $\mathbb{R}^2$ .

**Definition 3.2 (Integrale über Intervalle):**

Sei  $I = [\vec{a}, \vec{b}]$  ein Intervall im  $\mathbb{R}^n$ , und sei  $f: I \rightarrow \mathbb{R}$  stückweise stetig. Dann legen wir fest

$$\int_I f(\vec{x}) \, d\vec{x} := \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_n}^{b_n} f(x_1, x_2, \dots, x_n) \, dx_n \dots dx_2 \, dx_1.$$

Hierbei ist die Integrationsreihenfolge beliebig, also z.B. für  $n = 2$ :

$$\int_I f(x, y) \, d(x, y) = \int_a^b \int_c^d f(x, y) \, dy \, dx = \int_c^d \int_a^b f(x, y) \, dx \, dy,$$

wenn  $I = [a, b] \times [c, d]$ .

**Beispiele 3.3:** (a) Wenn wir die Funktion  $f(\vec{x}) = 1$  wählen, dann liefert das Integral uns das Volumen des Intervalls  $I$ :

$$\begin{aligned} \int_I 1 \, d\vec{x} &= \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} 1 \, dx_n \dots dx_1 = \int_{a_1}^{b_1} \dots \int_{a_{n-1}}^{b_{n-1}} (b_n - a_n) \, dx_{n-1} \dots dx_1 \\ &= (b_1 - a_1) \cdot (b_2 - a_2) \cdot \dots \cdot (b_n - a_n). \end{aligned}$$

(b) Ein schöner Spezialfall ergibt sich, wenn die Funktion  $f$  die Gestalt  $f(x, y) = g(x) \cdot h(y)$  hat. Dann entsteht

$$\begin{aligned} \int_a^b \int_c^d f(x, y) \, dy \, dx &= \int_a^b \int_c^d g(x)h(y) \, dy \, dx = \int_a^b g(x) \left( \int_c^d h(y) \, dy \right) \, dx \\ &= \left( \int_a^b g(x) \, dx \right) \cdot \left( \int_c^d h(y) \, dy \right). \end{aligned}$$

- (c) Sei  $f(x, y) = 2x \cos(x^2 - y)$ , und das Intervall sei  $I = [0, 1] \times [-1, 1]$ . Dann können wir einerseits substituieren wie folgt (wobei wir darauf achten, die Integrationsgrenzen mitzubesubstituieren):

$$\begin{aligned} & \int_{y=-1}^1 \int_{x=0}^1 2x \cos(x^2 - y) \, dx \, dy \quad \Bigg| \quad \text{substituiere } u = x^2, \quad \frac{du}{dx} = 2x \\ &= \int_{y=-1}^1 \int_{u=0}^1 \cos(u - y) \, du \, dy = \int_{y=-1}^1 \sin(u - y) \Big|_{u=0}^{u=1} \, dy \\ &= \int_{y=-1}^1 \left( \sin(1 - y) - \sin(-y) \right) \, dy = \left( \cos(1 - y) - \cos(-y) \right) \Big|_{y=-1}^{y=1} \\ &= \cos(0) - \cos(2) - \cos(-1) + \cos(1) = 1 - \cos(2) - \cos(1) + \cos(1) = 1 - \cos(2). \end{aligned}$$

Andererseits können wir aber auch die Integrale tauschen:

$$\begin{aligned} & \int_{x=0}^1 \int_{y=-1}^1 2x \cos(x^2 - y) \, dy \, dx = \int_{x=0}^1 2x \left( \int_{y=-1}^1 \cos(x^2 - y) \, dy \right) \, dx \\ &= (\dots \text{ähnliche Rechnung wie eben.} \dots) = 1 - \cos(2). \end{aligned}$$

Wir wollen jetzt auch über unregelmäßig berandeten Gebieten  $G$  integrieren, nicht bloß über rechteckigen Intervallen  $I$ .

**Definition 3.4 (Gebietsintegral):**

Sei  $G \subset \mathbb{R}^n$  kompakt mit stückweise glattem Rand, und die Funktion  $f: G \rightarrow \mathbb{R}$  sei stückweise stetig. Wir nennen  $G$  das *Integrationsgebiet*.

Wir wählen ein Intervall  $I \subset \mathbb{R}^n$  mit  $G \subseteq I$  und setzen

$$\tilde{f}(\vec{x}) := \begin{cases} f(\vec{x}) & : \text{wenn } \vec{x} \in G, \\ 0 & : \text{wenn } \vec{x} \notin G, \end{cases}$$

und diese Funktion  $\tilde{f}$  nennen wir *Fortsetzung von  $f$  durch Null*. Dann definieren wir

$$\int_G f(\vec{x}) \, d\vec{x} := \int_I \tilde{f}(\vec{x}) \, d\vec{x}.$$

Das Integral auf der linken Seite nennen wir *Gebietsintegral*.

Dies ist erstmal nur eine Definition zum Zwecke der Begriffsfestlegung. Die tatsächliche Berechnung eines solchen Integrals verlangt uns noch einige Arbeit ab, mit der wir jetzt loslegen.

Im Folgenden betrachten wir nur den Fall  $n = 2$ ; höhere Dimensionen können analog betrachtet werden.

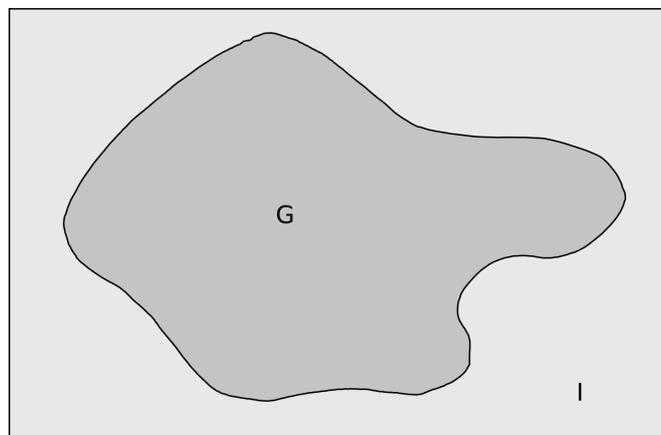


Abbildung 3.14: Ein beliebiges Integrationsgebiet  $G$  im  $\mathbb{R}^2$ .

**Definition 3.5 (Projizierbare Mengen):**

Sei  $G \subset \mathbb{R}^2$  eine gegebene Menge.

- (a) Wenn es stetige Funktionen  $c, d: [a, b] \rightarrow \mathbb{R}$  gibt mit  $c(x) \leq d(x)$  für jedes  $x \in [a, b]$  und der Eigenschaft

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : a \leq x \leq b \quad \text{und} \quad c(x) \leq y \leq d(x) \right\},$$

dann heißt  $G$   $y$ -projizierbar.

In diesem Fall haben wir

$$\int_G f(x, y) \, d(x, y) = \int_{x=a}^b \int_{y=c(x)}^{d(x)} f(x, y) \, dy \, dx.$$

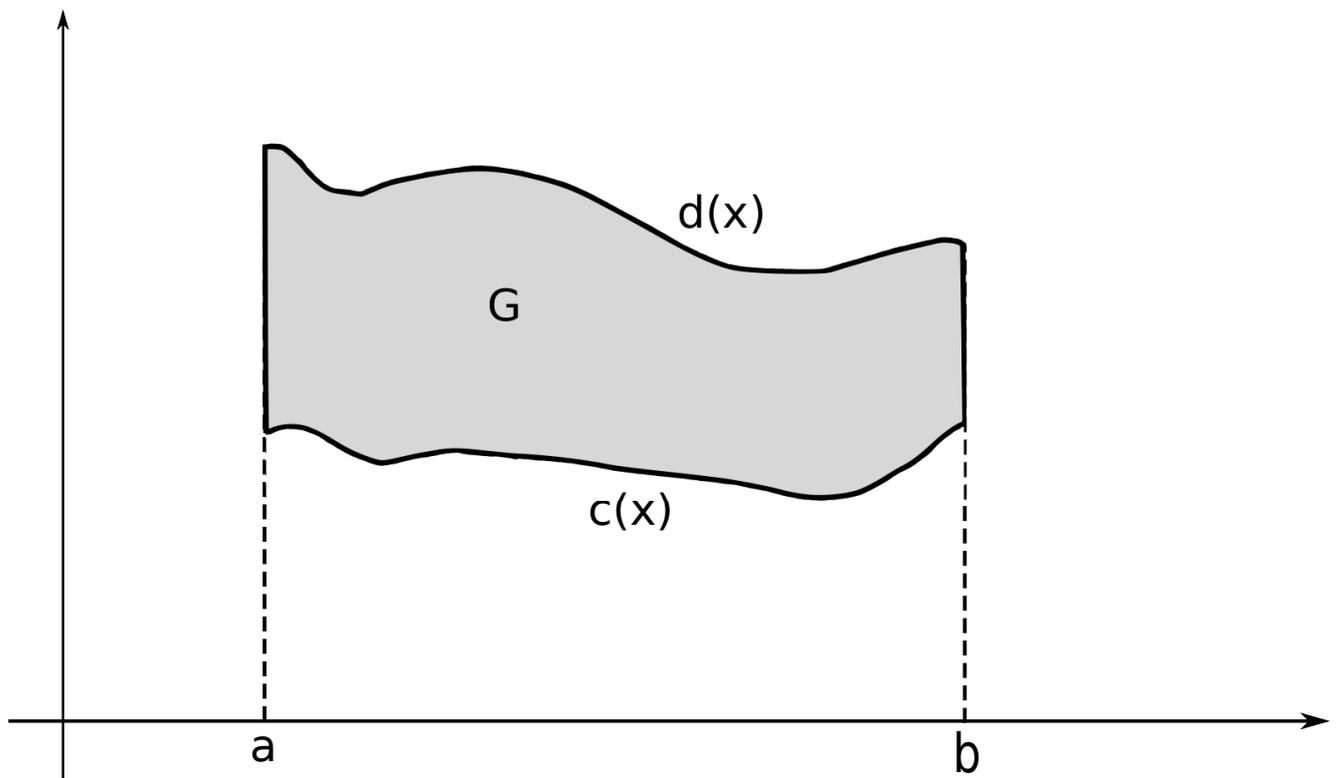


Abbildung 3.15: Eine  $y$ -projizierbare Menge  $G$ .

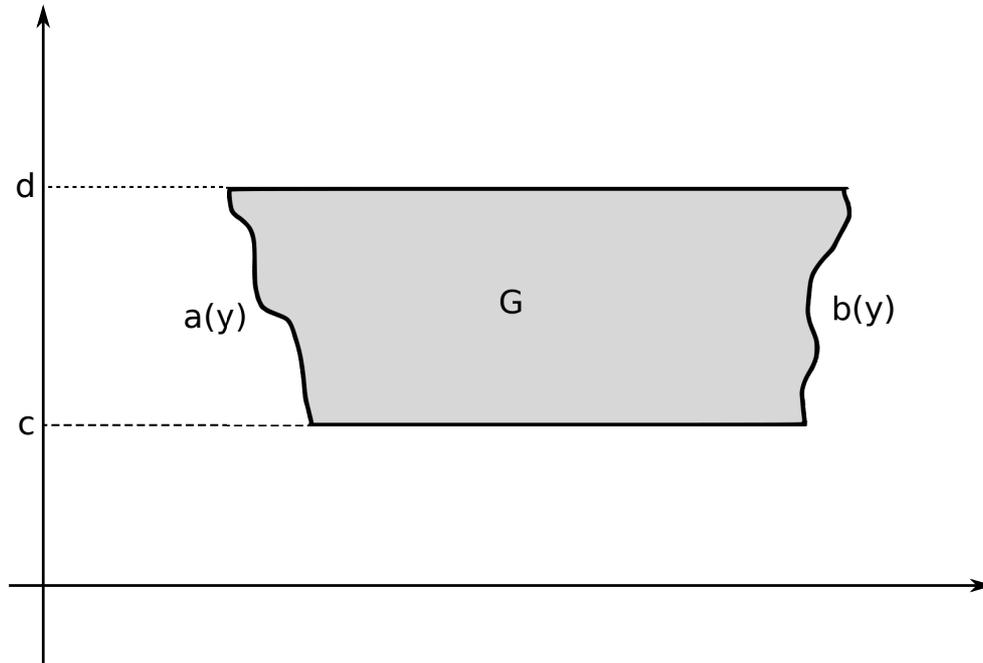
- (b) Wenn es stetige Funktionen  $a, b: [c, d] \rightarrow \mathbb{R}$  gibt mit  $a(y) \leq b(y)$  für jedes  $y \in [c, d]$  und der Eigenschaft

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : c \leq y \leq d \quad \text{und} \quad a(y) \leq x \leq b(y) \right\},$$

dann heißt  $G$   $x$ -projizierbar.

In diesem Fall haben wir

$$\int_G f(x, y) \, d(x, y) = \int_{y=c}^d \int_{x=a(y)}^{b(y)} f(x, y) \, dx \, dy.$$

Abbildung 3.16: Eine  $x$ -projizierbare Menge  $G$ .

**Beispiele 3.6:** (a) Wir betrachten die obere Halbkreisscheibe mit Radius  $r > 0$ :

$$G := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 \leq r^2 \quad \text{und} \quad y \geq 0 \right\} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : -r \leq x \leq r \quad \text{und} \quad 0 \leq y \leq \sqrt{r^2 - x^2} \right\}.$$

Also ist diese Menge  $G$   $y$ -projizierbar und wir haben

$$\int_G f(x, y) \, d(x, y) = \int_{x=-r}^r \int_{y=0}^{\sqrt{r^2-x^2}} f(x, y) \, dy \, dx.$$

Zum Beispiel für  $f(x, y) = 2x^2y$  ergibt sich dann:

$$\begin{aligned} \int_G 2x^2y \, d(x, y) &= \int_{x=-r}^r \int_{y=0}^{\sqrt{r^2-x^2}} 2x^2y \, dy \, dx = \int_{x=-r}^r x^2y^2 \Big|_{y=0}^{y=\sqrt{r^2-x^2}} \, dx \\ &= \int_{x=-r}^r x^2(r^2 - x^2) \, dx = \int_{x=-r}^r (r^2x^2 - x^4) \, dx \\ &= \left( \frac{1}{3}r^2x^3 - \frac{1}{5}x^5 \right) \Big|_{x=-r}^{x=r} \\ &= \frac{1}{3}r^5 + \frac{1}{3}r^5 - \frac{1}{5}r^5 - \frac{1}{5}r^5 = \left( \frac{2}{3} - \frac{2}{5} \right) r^5 = \frac{4}{15}r^5. \end{aligned}$$

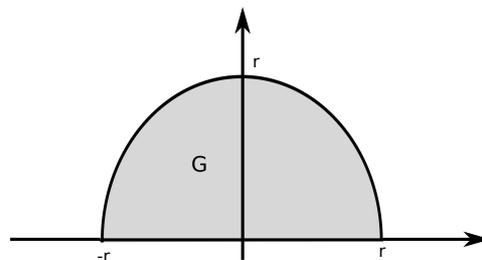


Abbildung 3.17: Eine Halbkreisscheibe.

- (b) Wir nehmen  $G$  als dasjenige gleichschenklige Dreieck in der Ebene, das begrenzt wird von den Geraden  $y = x$ ,  $y = 2 - x$  und  $y = 0$ . Sei  $f(x, y) = \frac{1}{1+x}$ . Diese Menge  $G$  ist  $x$ -projizierbar und  $y$ -projizierbar. Wir können uns also für die leichtere Rechnung entscheiden:

$$\begin{aligned} \int_G f(x, y) \, d(x, y) &= \int_{y=0}^1 \int_{x=y}^{x=2-y} \frac{1}{1+x} \, dx \, dy = \int_{y=0}^1 \ln(1+x) \Big|_{x=y}^{x=2-y} \, dy \\ &= \int_{y=0}^1 (\ln(3-y) - \ln(1+y)) \, dy = \dots = 3 \ln(3) - 4 \ln(2). \end{aligned}$$

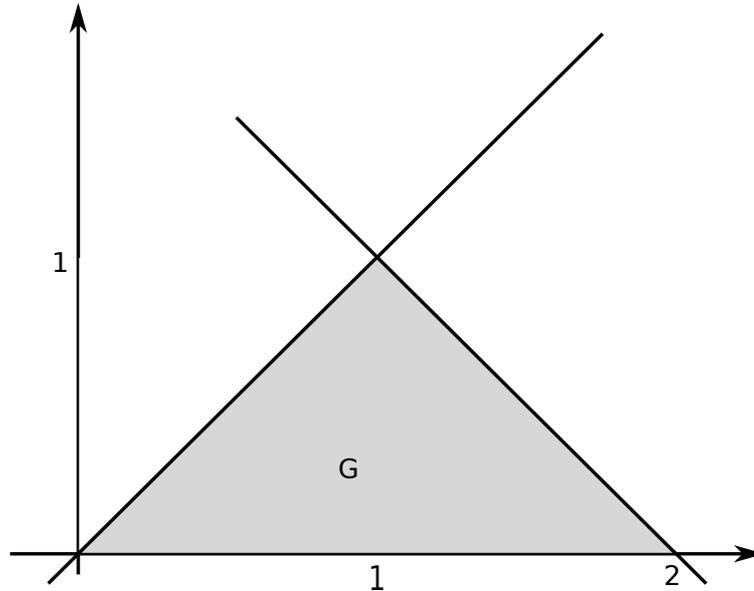


Abbildung 3.18: Ein Dreieck.

- (c) Wir betrachten einen Kreisring mit innerem Radius 1 und äußerem Radius 4:

$$G := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 1 \leq x^2 + y^2 \leq 4 \right\}.$$

Leider ist diese Menge weder  $x$ -projizierbar noch  $y$ -projizierbar. Es gibt aber mindestens zwei Auswege:

- Wir zerlegen  $G$  in einen „oberen“ und einen „unteren“ Halbring:

$$G = \left( G \cap \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : y \geq 0 \right\} \right) \cup \left( G \cap \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : y \leq 0 \right\} \right) = G_1 \cup G_2,$$

und dann berechnen wir

$$\int_G f(x, y) \, d(x, y) = \int_{G_1} f(x, y) \, d(x, y) + \int_{G_2} f(x, y) \, d(x, y).$$

Beachte, dass  $G_1$  und  $G_2$  beide  $y$ -projizierbar sind.

- Wir schreiben  $G$  als Differenzmenge zweier Kreisscheiben, also  $G = K_4(\vec{0}) \setminus K_1(\vec{0})$ , und dann berechnen wir

$$\int_G f(x, y) \, d(x, y) = \int_{K_4(\vec{0})} f(x, y) \, d(x, y) - \int_{K_1(\vec{0})} f(x, y) \, d(x, y).$$

Ein analoges Vorgehen ermöglicht für viele Mengen  $G$  die Rückführung auf projizierbare Mengen.

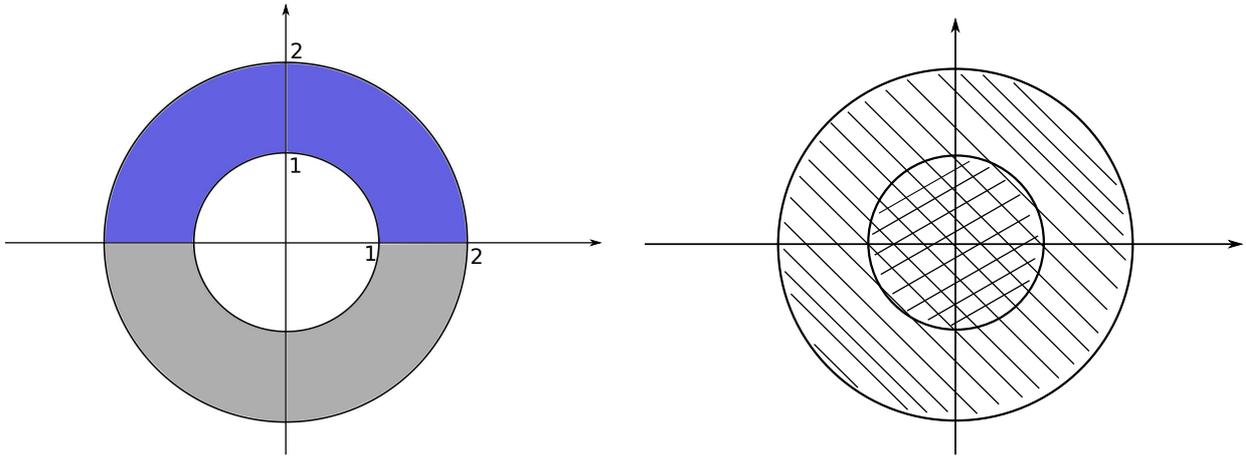


Abbildung 3.19: Wir zerlegen einen Kreisring in zwei Halbringe, oder wir interpretieren ihn als Differenzmenge zweier Kreisscheiben.

Ab jetzt gehen wir vom  $\mathbb{R}^2$  wieder zurück zum  $\mathbb{R}^n$  mit beliebigem  $n$ .

**Satz 3.7 (Substitutionsregel):**

Es seien  $G, \hat{G} \subset \mathbb{R}^n$  Integrationsgebiete,  $f: G \rightarrow \mathbb{R}$  sei stückweise stetig, und die Funktion  $g: \hat{G} \rightarrow G$  sei bijektiv und stetig differenzierbar. Dann gilt

$$\int_G f(\vec{x}) \, d\vec{x} = \int_{\hat{G}} f(g(\vec{y})) |\det(J_g(\vec{y}))| \, d\vec{y}.$$

Hierbei ist  $J_g$  die Jacobi-Matrix der Abbildung  $g$ .

Der Nutzen dieses Satzes wird für uns im wesentlichen darin liegen, ein schwierig zu beschreibendes Integrationsgebiet  $G$  in ein leichter zu beschreibendes Gebiet  $\hat{G}$  umzurechnen.

**Verfahren 3.8 (Umrechnung in Polarkoordinaten):**

Wir erinnern daran, dass wir im  $\mathbb{R}^2$  Polarkoordinaten eingeführt hatten gemäß

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{bmatrix}.$$

Dann ist  $g: [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2$  mit

$$g(r, \varphi) := \begin{bmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{bmatrix}$$

die Substitutionsfunktion bei der Umrechnung in Polarkoordinaten, und es ergibt sich

$$\int_G f(x, y) \, d(x, y) = \int_{\hat{G}} f(g(r, \varphi)) |\det J_g(r, \varphi)| \, d(r, \varphi),$$

wobei

$$\det J_g(r, \varphi) = \det \begin{bmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{bmatrix} = r \cos^2(\varphi) + r \sin^2(\varphi) = r,$$

was wir in folgende Merkregel gießen:

$$dx \, dy = r \, dr \, d\varphi.$$

**Beispiele 3.9:** (a) Sei  $G$  ein Kreis um den Ursprung mit Radius  $R$ , also

$$G = K_R(\vec{0}) = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 \leq R^2 \right\}.$$

Dieser Menge  $G$  entspricht in Polarkoordinaten die Menge

$$\hat{G} = \{(r, \varphi) \in [0, \infty) \times [0, 2\pi) : 0 \leq r \leq R\} = [0, R] \times [0, 2\pi),$$

und dieses  $\hat{G}$  ist erfreulicherweise ein Intervall, was das Verwalten der Integrationsgrenzen deutlich vereinfacht. Also ist

$$\begin{aligned} \int_{K_R(\vec{0})} f(x, y) \, d(x, y) &= \int_{\hat{G}} f(r \cos(\varphi), r \sin(\varphi)) r \, d(r, \varphi) \\ &= \int_{\varphi=0}^{2\pi} \int_{r=0}^R f(r \cos(\varphi), r \sin(\varphi)) r \, dr \, d\varphi. \end{aligned}$$

Zum Beispiel für  $f(x, y) = 1$  sollte sich der Flächeninhalt des Kreises ergeben. Das prüfen wir nach:

$$\begin{aligned} \int_{K_R(\vec{0})} 1 \, d(x, y) &= \int_{\varphi=0}^{2\pi} \int_{r=0}^R r \, dr \, d\varphi = \left( \int_{\varphi=0}^{2\pi} d\varphi \right) \cdot \left( \int_{r=0}^R r \, dr \right) \\ &= 2\pi \cdot \frac{1}{2} r^2 \Big|_{r=0}^{r=R} = 2\pi \cdot \frac{1}{2} R^2 = \pi R^2. \end{aligned}$$

(b) Wir greifen Beispiel 3.6 (c) auf:

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4 \right\}, \quad f(x, y) = 1 + y.$$

In Polarkoordinaten wird dies zu  $\hat{G} = \{(r, \varphi) : 1 \leq r \leq 2, \quad 0 \leq \varphi \leq 2\pi\}$  und  $f(r \cos(\varphi), r \sin(\varphi)) = 1 + r \sin(\varphi)$ , also

$$\begin{aligned} \int_G (1 + y) \, d(x, y) &= \int_{\varphi=0}^{2\pi} \int_{r=1}^2 (1 + r \sin(\varphi)) r \, dr \, d\varphi \\ &= \int_{\varphi=0}^{2\pi} \int_{r=1}^2 r \, dr \, d\varphi + \int_{\varphi=0}^{2\pi} \int_{r=1}^2 r^2 \sin(\varphi) \, dr \, d\varphi \\ &= \left( \int_{\varphi=0}^{2\pi} d\varphi \right) \cdot \left( \int_{r=1}^2 r \, dr \right) + \underbrace{\left( \int_{\varphi=0}^{2\pi} \sin(\varphi) \, d\varphi \right)}_{=0} \cdot \underbrace{\left( \int_{r=1}^2 r^2 \, dr \right)}_{\text{braucht man nicht berechnen}} \\ &= 2\pi \cdot \frac{1}{2} r^2 \Big|_{r=1}^{r=2} \\ &= 2\pi \cdot \left( 2 - \frac{1}{2} \right) = 3\pi. \end{aligned}$$

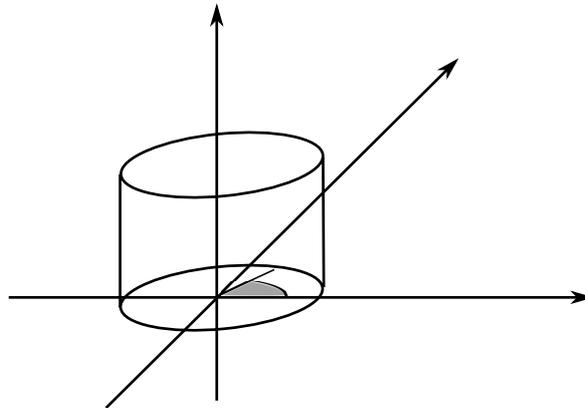


Abbildung 3.20: Der Winkel  $\varphi$  bei Zylinderkoordinaten.

**Verfahren 3.10 (Umrechnung in Zylinderkoordinaten):**

Dies sind Koordinaten, die zu einer Zylindersymmetrie im  $\mathbb{R}^3$  gut passen. Wir rechnen um

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} r \\ \varphi \\ z \end{bmatrix}$$

gemäß der Vorschrift

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \cos(\varphi) \\ r \sin(\varphi) \\ z \end{bmatrix} = g(r, \varphi, z).$$

Für die Umrechnung der Differentiale benötigen wir den Betrag der Determinante der Jacobi-Matrix:

$$|\det J_g(r, \varphi, z)| = \left| \det \begin{bmatrix} \cos(\varphi) & -r \sin(\varphi) & 0 \\ \sin(\varphi) & r \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = 1 \cdot \left| \det \begin{bmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{bmatrix} \right| = r,$$

analog zur Umrechnung in Polarkoordinaten. Wir erhalten also die Merkregel

$$dx dy dz = r dr d\varphi dz.$$

**Beispiel 3.11:**

Wir bestimmen das Volumen eines Kegels. Dieser steht auf der Spitze, hat die Höhe 1 und den Radius 1. Es ist also

$$G = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x^2 + y^2 \leq z^2, \quad 0 \leq z \leq 1 \right\}.$$

In Zylinderkoordinaten wird daraus dann  $0 \leq r \leq z$ ,  $0 \leq \varphi \leq 2\pi$  und  $0 \leq z \leq 1$ . Das Volumen des Kegels bestimmt sich dann zu (hierbei müssen wir aufpassen, weil die Grenzen für  $r$  von  $z$  abhängen, weshalb das  $r$ -Integral nach Innen gehört):

$$\begin{aligned} \int_G 1 d(x, y, z) &= \int_{z=0}^1 \int_{\varphi=0}^{2\pi} \int_{r=0}^z 1 \cdot r dr d\varphi dz = \int_{z=0}^1 \int_{\varphi=0}^{2\pi} \frac{1}{2} r^2 \Big|_{r=0}^{r=z} d\varphi dz \\ &= \int_{z=0}^1 \int_{\varphi=0}^{2\pi} \frac{1}{2} z^2 d\varphi dz = \frac{1}{2} \cdot 2\pi \cdot \int_{z=0}^1 z^2 dz \\ &= \pi \cdot \frac{1}{3} z^3 \Big|_{z=0}^{z=1} = \frac{1}{3} \pi. \end{aligned}$$

Das entspricht unseren Erwartungen („Grundflächeninhalt mal Höhe durch drei“).

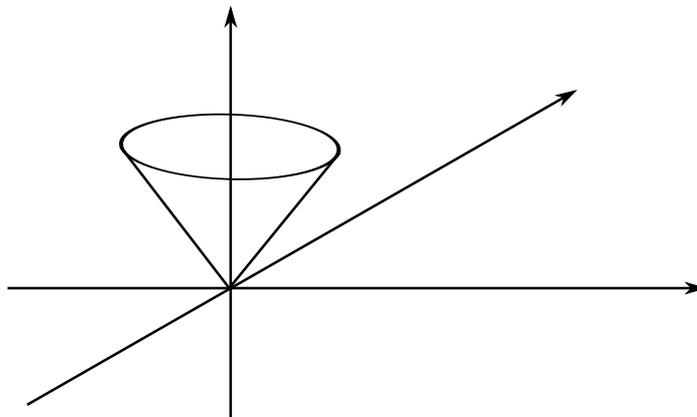


Abbildung 3.21: Wir suchen das Volumen dieses Kegels.

**Verfahren 3.12 (Umrechnung in Kugelkoordinaten):**

Dies sind Koordinaten, die zu einer Kugelsymmetrie im  $\mathbb{R}^3$  gut passen. Wir nehmen  $r \geq 0$  als den Abstand von  $[x, y, z]^\top$  zum Ursprung,  $\varphi \in [0, 2\pi)$  als den Winkel von  $[x, y]^\top$  zur positiven  $x$ -Achse (auch als geographische Länge bezeichnet), und  $\theta \in [0, \pi]$  ist der Winkel von  $[x, y, z]^\top$  zur positiven  $z$ -Achse. Für  $\theta = 0$  ergibt sich der Nordpol, für  $\theta = \pi$  der Südpol, und für  $\theta = \pi/2$  der Äquator. (Achtung: in manchen Büchern ist  $\theta \in [-\pi/2, \pi/2]$ , wobei  $\theta = 0$  den Äquator beschreibt. Dann ändern sich einige Formeln.)

Die Wortbeschreibung ergibt dann die Formeln

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = g(r, \varphi, \theta) = \begin{bmatrix} r \cos(\varphi) \sin(\theta) \\ r \sin(\varphi) \sin(\theta) \\ r \cos(\theta) \end{bmatrix}.$$

Für die Umrechnung der Differentiale benötigen wir die Determinante der Jacobi-Matrix:

$$\begin{aligned} \det J_g(r, \varphi, \theta) &= \begin{vmatrix} \cos(\varphi) \sin(\theta) & -r \sin(\varphi) \sin(\theta) & r \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \sin(\theta) & r \cos(\varphi) \sin(\theta) & r \sin(\varphi) \cos(\theta) \\ \cos(\theta) & 0 & -r \sin(\theta) \end{vmatrix} \\ &= (\dots \text{einige Rechnungen} \dots) = -r^2 \sin(\theta). \end{aligned}$$

Die Substitutionsformel verlangt jetzt allerdings den *Betrag* der Determinante der Jacobi-Matrix, also bekommen wir die Merkgel

$$dx \, dy \, dz = r^2 \sin(\theta) \, dr \, d\varphi \, d\theta.$$

**Beispiel 3.13:**

Wir bestimmen das Volumen der Einheitskugel im  $\mathbb{R}^3$ , also betrachten wir

$$G = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1 \right\}.$$

In Kugelkoordinaten wird dies zu  $0 \leq r \leq 1$ ,  $0 \leq \varphi \leq 2\pi$ ,  $0 \leq \theta \leq \pi$ . Das Volumen der Kugel berechnet sich dann so:

$$\begin{aligned} \int_G 1 \, d(x, y, z) &= \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \int_{r=0}^1 1 \cdot r^2 \sin(\theta) \, dr \, d\varphi \, d\theta \\ &= \left( \int_{\theta=0}^{\pi} \sin(\theta) \, d\theta \right) \cdot \left( \int_{\varphi=0}^{2\pi} d\varphi \right) \cdot \left( \int_{r=0}^1 r^2 \, dr \right) \\ &= -\cos(\theta) \Big|_{\theta=0}^{\theta=\pi} \cdot 2\pi \cdot \frac{1}{3} r^3 \Big|_{r=0}^{r=1} \\ &= 2 \cdot 2\pi \cdot \frac{1}{3} = \frac{4}{3}\pi. \end{aligned}$$

Das ist das erwartete Ergebnis.

## 4 Vektoranalysis

Unser **Ziel** ist es, Vektorfelder besser zu beschreiben. Wir denken dabei an elektrische oder magnetische Felder, an das Gravitationsfeld, an Strömungen oder an elastische Verformungen: sei z.B. ein Stahlträger gegeben, der sich unter Last ein wenig verbiegt. Das können wir uns so vorstellen, dass wir den Träger im Ruhezustand hernehmen und an jeden seiner Punkte einen Vektor anheften, der diesen Punkt im Ruhezustand mit dem neuen Punkt unter Last verbindet. Auf diesem Wege entsteht ein Vektorfeld, das *Verschiebungsfeld* genannt wird. Wie können wir dies mathematisch beschreiben?

**Definition 4.1 (Divergenz):**

Sei  $M \subseteq \mathbb{R}^n$  offen, und  $f: M \rightarrow \mathbb{R}^n$  sei ein stetig differenzierbares Vektorfeld. Dann nennen wir

$$\operatorname{div} f(\vec{x}) := \sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(\vec{x}) = \frac{\partial f_1}{\partial x_1}(\vec{x}) + \frac{\partial f_2}{\partial x_2}(\vec{x}) + \dots + \frac{\partial f_n}{\partial x_n}(\vec{x}), \quad \vec{x} \in M,$$

die *Divergenz von  $f$* . Beachte dabei: es ist  $f: M \rightarrow \mathbb{R}^n$  ein Vektorfeld, aber  $\operatorname{div} f: M \rightarrow \mathbb{R}$  ein Skalarfeld.

**Bemerkung 4.2 (Anschauliche Bedeutung):**

Die Funktion  $\operatorname{div} f$  misst den aus einer Volumeneinheit austretenden Fluss:

- Punkte  $\vec{x} \in M$  mit  $\operatorname{div} f(\vec{x}) > 0$  heißen *Quellen*,
- Punkte  $\vec{x} \in M$  mit  $\operatorname{div} f(\vec{x}) < 0$  heißen *Senken*.
- Wenn  $\operatorname{div} f(\vec{x}) = 0$  für alle  $\vec{x} \in M$  gilt, so nennt man das Vektorfeld  $f$  *quellenfrei*.



Abbildung 4.22: Links ist eine Quelle mit  $\operatorname{div} f > 0$ , und rechts ist eine Senke mit  $\operatorname{div} f < 0$ .

**Definition 4.3 (Rotation):**

Sei  $M \subset \mathbb{R}^3$  offen, und sei  $f: M \rightarrow \mathbb{R}^3$  ein stetig differenzierbares Vektorfeld. Dann heißt

$$\operatorname{rot} f(\vec{x}) := \begin{bmatrix} \frac{\partial f_3}{\partial x_2}(\vec{x}) - \frac{\partial f_2}{\partial x_3}(\vec{x}) \\ \frac{\partial f_1}{\partial x_3}(\vec{x}) - \frac{\partial f_3}{\partial x_1}(\vec{x}) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}) - \frac{\partial f_1}{\partial x_2}(\vec{x}) \end{bmatrix}$$

*Rotation von  $f$ .* Beachte dabei:  $f: M \rightarrow \mathbb{R}^3$  und  $\operatorname{rot} f: M \rightarrow \mathbb{R}^3$  sind beides Vektorfelder. Die Rotation gibt es nur im  $\mathbb{R}^3$ , aber nicht im  $\mathbb{R}^2$ .

**Bemerkung 4.4:** (a) Mit dem  $n$ -komponentigen *Nabla-Operator*

$$\vec{\nabla}_n := \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix}$$

haben wir formal

$$\operatorname{div} f(\vec{x}) = \underbrace{\vec{\nabla}_n^\top \cdot f(\vec{x})}_{\text{Skalarprodukt}}, \quad \operatorname{rot} f(\vec{x}) = \underbrace{(\vec{\nabla}_3 \times f)(\vec{x})}_{\text{Kreuzprodukt}}.$$

(b) Wir erinnern an die *Integrabilitätsbedingung* aus dem Abschnitt über Potenziale:

$$\frac{\partial f_i}{\partial x_j}(\vec{x}) = \frac{\partial f_j}{\partial x_i}(\vec{x}) \quad \text{für alle } i \neq j.$$

Diese Bedingung ist also erfüllt für  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  genau dann, wenn  $\operatorname{rot} f(\vec{x}) = \vec{0}$  für alle  $\vec{x} \in \mathbb{R}^3$  gilt.

(c) Es beschreibt  $\operatorname{rot} f$  die Wirbeldichte des Feldes.

- Gradientenfelder  $f$  haben immer  $\operatorname{rot} f = \vec{0}$ , denn: sei  $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}$  ein zweimal stetig differenzierbares Skalarfeld. Dann gilt

$$\operatorname{rot}(\nabla\varphi) = \operatorname{rot} \begin{bmatrix} \varphi_x \\ \varphi_y \\ \varphi_z \end{bmatrix} = \begin{bmatrix} \varphi_{zy} - \varphi_{yz} \\ \varphi_{xz} - \varphi_{zx} \\ \varphi_{yx} - \varphi_{xy} \end{bmatrix} = \vec{0}.$$

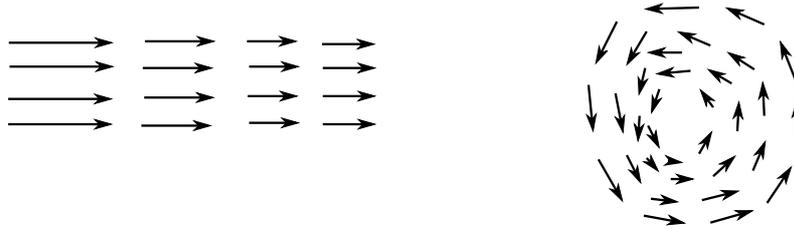


Abbildung 4.23: Links ist ein Vektorfeld  $f$  mit  $\operatorname{rot} f = \vec{0}$ , rechts ist ein Vektorfeld  $f$  mit  $\operatorname{rot} f \neq \vec{0}$ .

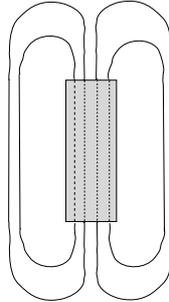


Abbildung 4.24: Die Feldlinien zum Magnetfeld eines Stabmagneten (sehr schematisch).

- Wenn  $\operatorname{rot} f \neq \vec{0}$  ist, so kann  $f$  kein Gradientenfeld sein. Typische Beispiele sind Magnetfelder.

Wir merken uns die anschauliche Interpretation: „ $\operatorname{rot} f$  misst die Abweichung des Vektorfeldes  $f$  von einem Gradientenfeld“.

**Bemerkung 4.5:**

Wir kombinieren die Operatoren  $\nabla$ ,  $\operatorname{div}$  und  $\operatorname{rot}$ :

- Für  $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}$  gilt  $\operatorname{rot}(\nabla\varphi) = \vec{0}$ .
- Für  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  gilt  $\operatorname{div}(\operatorname{rot} f) = 0$ .
- Für  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$  gilt

$$\begin{aligned} \Delta \varphi &:= \operatorname{div}(\nabla\varphi) = \operatorname{div}(\operatorname{grad} \varphi) = \operatorname{div} \begin{bmatrix} \frac{\partial \varphi}{\partial x_1} & \frac{\partial \varphi}{\partial x_2} & \cdots & \frac{\partial \varphi}{\partial x_n} \end{bmatrix} \\ &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \frac{\partial \varphi}{\partial x_j} = \sum_{j=1}^n \frac{\partial^2 \varphi}{\partial x_j^2}, \end{aligned}$$

und wir nennen  $\Delta := \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$  den *Laplace*-Operator.

**Definition 4.6 (Flächen im  $\mathbb{R}^3$ ):**

Sei  $D \subset \mathbb{R}^2$  kompakt,  $G \supset D$  offen, und sei  $\Phi: G \rightarrow \mathbb{R}^3$  stetig differenzierbar, sodass die Jacobi-Matrix von  $\Phi$ , das heißt

$$J_{\Phi}(u, v) = \begin{bmatrix} \frac{\partial \Phi_1}{\partial u}(u, v) & \frac{\partial \Phi_1}{\partial v}(u, v) \\ \frac{\partial \Phi_2}{\partial u}(u, v) & \frac{\partial \Phi_2}{\partial v}(u, v) \\ \frac{\partial \Phi_3}{\partial u}(u, v) & \frac{\partial \Phi_3}{\partial v}(u, v) \end{bmatrix},$$

für alle  $(u, v) \in G$  vollen Rang hat, also  $\operatorname{rang} J_{\Phi}(u, v) = 2$ .

Dann heißt  $\Phi$  *reguläre Fläche (über  $D$ )*, und

$$\mathcal{F} := \Phi(D) = \left\{ \Phi(u, v) : \begin{bmatrix} u \\ v \end{bmatrix} \in D \right\}$$

heißt *reguläres Flächenstück*.

**Beispiel 4.7:**

Wir wollen von einer Kugel (deren Mittelpunkt im Ursprung liegt) mit Radius  $R$  die obere und untere Halbkugelfläche beschreiben. Dazu wählen wir das Parametergebiet

$$G := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 : u^2 + v^2 < R^2 \right\},$$

was eine offene Teilmenge des  $\mathbb{R}^2$  ist. Für die Halbkugelflächen wählen wir zwei Funktionen  $\Phi_+$  und  $\Phi_-$ ,

$$\Phi_{\pm}(u, v) = \begin{bmatrix} u \\ v \\ \pm \sqrt{R^2 - u^2 - v^2} \end{bmatrix},$$

und diese sind stetig differenzierbar auf  $G$ . Wir haben die Jacobi-Matrizen

$$J_{\Phi_{\pm}}(u, v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\mp u}{\sqrt{R^2 - u^2 - v^2}} & \frac{\mp v}{\sqrt{R^2 - u^2 - v^2}} \end{bmatrix},$$

was für alle  $[u, v]^T \in G$  immer den Rang zwei besitzt.

Auf dem Rand von  $G$ , also für  $u^2 + v^2 = R^2$ , bekämen wir aber in der letzten Zeile von  $J_{\Phi_{\pm}}$  zwei Divisionen durch Null. Weil  $G$  eine offene Menge des  $\mathbb{R}^2$  ist, gehört der Rand von  $G$  aber sowieso nicht zu  $G$ . Wir wollen einen kleinen Sicherheitsabstand  $\varepsilon$  zum Rand von  $G$  einhalten. Deshalb definieren wir

$$D := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 : u^2 + v^2 \leq (R - \varepsilon)^2 \right\}, \quad \varepsilon > 0.$$

Die Funktionen  $\Phi_{\pm}$  sind reguläre Flächen über  $D$ , und die zugehörigen Flächenstücke  $\mathcal{F}_+ := \Phi_+(D)$  sowie  $\mathcal{F}_- := \Phi_-(D)$  sind dann (abgesehen von einer sehr schmalen Bauchbinde entlang des Äquators) die obere Halbkugelfläche und die untere Halbkugelfläche.

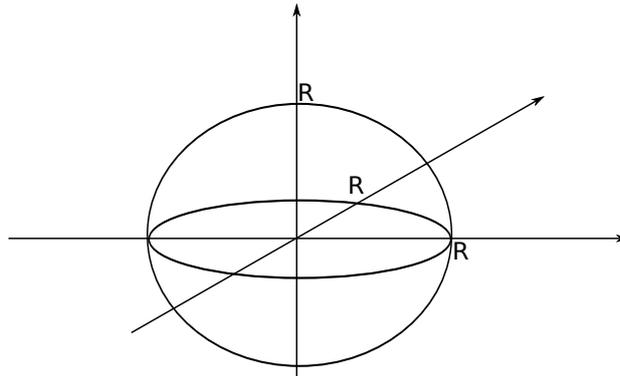


Abbildung 4.25: Die Kugel mit Radius  $R$ .

Jetzt können wir Flächenstücke beschreiben, und als nächstes schauen wir uns deren Tangentialebenen an. Siehe auch Bild 3.15 in Kapitel 2 für eine etwas speziellere Situation, in der wir schon mal eine Tangentialebene beschrieben hatten.

**Definition 4.8 (Tangentialebene und Normalenvektor einer Fläche):**

Sei  $D \subset \mathbb{R}^2$  kompakt,  $G \supset D$  offen,  $\Phi: G \rightarrow \mathbb{R}^3$  eine reguläre Fläche. Dann heißt für  $[u_0, v_0]^T \in D$  die Ebene

$$T(u_0, v_0) := \left\{ \Phi(u_0, v_0) + \lambda \frac{\partial \Phi}{\partial u}(u_0, v_0) + \mu \frac{\partial \Phi}{\partial v}(u_0, v_0) : \lambda, \mu \in \mathbb{R} \right\}$$

Tangentialebene an die Fläche  $\Phi$  im Punkt  $\Phi(u_0, v_0)$ .

Weiterhin heißt

$$\vec{n}(u_0, v_0) := \frac{\frac{\partial \Phi}{\partial u}(u_0, v_0) \times \frac{\partial \Phi}{\partial v}(u_0, v_0)}{\left\| \frac{\partial \Phi}{\partial u}(u_0, v_0) \times \frac{\partial \Phi}{\partial v}(u_0, v_0) \right\|}$$

Normalenvektor an die Fläche  $\Phi$  im Punkt  $\Phi(u_0, v_0)$ .

Man beachte: bei der Beschreibung der Tangentialebene haben wir den Aufpunkt  $\Phi(u_0, v_0)$  genommen, und dort haben wir die beiden Vektoren  $\lambda \partial_u \Phi$  und  $\mu \partial_v \Phi$  angehängt. Diese beiden Vektoren zeigen in unterschiedliche Richtungen des  $\mathbb{R}^3$ , denn sie sind Vielfache der beiden Spalten der Jacobi-Matrix  $J_\Phi$ , und diese Spalten sind linear unabhängig, weil wir vorausgesetzt hatten, dass die Jacobi-Matrix  $J_\Phi$  den Rang zwei hat (und der Rang ist die Anzahl der linear unabhängigen Spalten). Also ist  $T(u_0, v_0)$  tatsächlich die Beschreibung einer Ebene.

Wenn wir das Kreuzprodukt  $\partial_u \Phi \times \partial_v \Phi$  der beiden die Tangentialebene aufspannenden Vektoren bilden, dann entsteht (gemäß der Eigenschaften des Kreuzprodukts) ein Produktvektor, der auf jedem der beiden Faktoren  $\partial_u \Phi(u_0, v_0)$  und  $\partial_v \Phi(u_0, v_0)$  senkrecht steht. Also steht  $\vec{n}(u_0, v_0)$  tatsächlich auf der Tangentialebene  $T(u_0, v_0)$  senkrecht. Das rechtfertigt den Namen *Normalenvektor*.

**Definition 4.9 (Flächeninhalt):**

Sei  $G \subseteq \mathbb{R}^2$  offen,  $D \subset G$  kompakt, und  $\Phi: G \rightarrow \mathbb{R}^3$  sei eine reguläre Fläche über  $D$ . Dann ist der Flächeninhalt von  $\Phi(D)$  gegeben durch

$$\int_D \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d(u, v). \quad (\spadesuit)$$

Das probieren wir anhand einer Kugeloberfläche aus, siehe Beispiel 4.7:

**Beispiel 4.10:**

Wir haben

$$\Phi_\pm(u, v) = \begin{bmatrix} u \\ v \\ \pm \sqrt{R^2 - u^2 - v^2} \end{bmatrix}, \quad J_{\Phi_\pm}(u, v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \mp u & \mp v \\ \sqrt{R^2 - u^2 - v^2} & \sqrt{R^2 - u^2 - v^2} \end{bmatrix},$$

und das Kreuzprodukt der beiden Spalten von  $J_\Phi$  berechnen wir so:

$$\frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) = \begin{bmatrix} 0 \pm \frac{u}{\sqrt{R^2 - u^2 - v^2}} \\ 0 \pm \frac{v}{\sqrt{R^2 - u^2 - v^2}} \\ 1 - 0 \end{bmatrix} = \frac{\pm 1}{\sqrt{R^2 - u^2 - v^2}} \begin{bmatrix} u \\ v \\ \pm \sqrt{R^2 - u^2 - v^2} \end{bmatrix}.$$

Davon berechnen wir jetzt die pythagoräische Länge:

$$\left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| = \frac{1}{\sqrt{R^2 - u^2 - v^2}} \sqrt{u^2 + v^2 + (R^2 - u^2 - v^2)} = \frac{R}{\sqrt{R^2 - u^2 - v^2}}.$$

Wenn wir das Kreuzprodukt auf Länge Eins bringen, dann entsteht der Normalenvektor:

$$\vec{n}(u, v) = \frac{\frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v)}{\left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\|} = \frac{1}{R} \begin{bmatrix} u \\ v \\ \pm \sqrt{R^2 - u^2 - v^2} \end{bmatrix}.$$

Wir berechnen jetzt den Flächeninhalt der Halbkugelfläche, wobei wir  $D := \{[u, v]^\top \in \mathbb{R}^2 : u^2 + v^2 \leq R^2\}$  nehmen:

$$\begin{aligned} \int_D \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d(u, v) &= \int_D \frac{R}{\sqrt{R^2 - u^2 - v^2}} d(u, v) && \left| \text{nimm Polarkoordinaten} \right. \\ &= \int_{\varphi=0}^{2\pi} \int_{r=0}^R \frac{R}{\sqrt{R^2 - r^2}} r dr d\varphi \\ &= 2\pi \int_{r=0}^R \frac{R}{\sqrt{R^2 - r^2}} r dr && \left| \text{substituiere } t := R^2 - r^2, \frac{dt}{dr} = -2r \implies r dr = -\frac{1}{2} dt \right. \\ &= 2\pi \cdot \int_{t=R^2}^{t=0} \frac{R}{\sqrt{t}} \cdot \frac{-1}{2} dt \\ &= -R\pi \cdot \int_{t=R^2}^{t=0} \frac{1}{\sqrt{t}} dt \\ &= +R\pi \cdot \int_{t=0}^{t=R^2} \frac{1}{\sqrt{t}} dt \\ &= R\pi \cdot 2\sqrt{t} \Big|_{t=0}^{t=R^2} = 2\pi R^2. \end{aligned}$$

Die Gesamtkugeloberfläche ist also  $4\pi R^2$ , wie zu erwarten.

Wir führen eine Schreibweise ein:

$$d\sigma := \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d(u, v).$$

Dann können wir das Integral ( $\spadesuit$ ) auch als  $\int_D 1 d\sigma$  schreiben. Das wollen wir jetzt etwas verallgemeinern:

**Definition 4.11 (Oberflächenintegral):**

Sei  $\Phi$  eine reguläre Fläche über  $D$  und  $\mathcal{F} := \Phi(D)$  das zugehörige reguläre Flächenstück. Sei  $f: \mathcal{F} \rightarrow \mathbb{R}$  eine stetige skalare Funktion. Dann definieren wir

$$\int_{\mathcal{F}} f(\vec{x}) d\sigma(\vec{x}) := \int_D f(\Phi(u, v)) \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d(u, v),$$

und dieses Integral nennen wir *Oberflächenintegral von  $f$  über der Fläche  $\mathcal{F}$* . Beachte, dass wir genau den Flächeninhalt bekommen, falls  $f = 1$ .

**Beispiel 4.12:**

Wir nehmen  $f(x, y, z) = z$  und  $\Phi: D \rightarrow \mathbb{R}^3$  mit

$$D := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} : u^2 + v^2 \leq R^2 \right\}, \quad \Phi(u, v) = \begin{bmatrix} u \\ v \\ \sqrt{R^2 - u^2 - v^2} \end{bmatrix},$$

wie oben. Dann berechnen wir

$$\begin{aligned} \int_{\mathcal{F}} f(x, y, z) d\sigma &= \int_D f(u, v, \sqrt{R^2 - u^2 - v^2}) \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| d(u, v) && \left| \text{siehe oben} \right. \\ &= \int_D \sqrt{R^2 - u^2 - v^2} \cdot \frac{R}{\sqrt{R^2 - u^2 - v^2}} d(u, v) = R \int_D d(u, v) && \left| \text{Polarkoordinaten} \right. \\ &= R \int_{\varphi=0}^{2\pi} \int_{r=0}^R r dr d\varphi = 2\pi R \cdot \frac{1}{2} r^2 \Big|_{r=0}^{r=R} = \pi R^2. \end{aligned}$$

Der nächste Satz ist ein berühmtes Theorem, das uns die Möglichkeit gibt, bestimmte Integrale über dreidimensionale Mengen umzuwandeln in Oberflächenintegrale über den Rand dieser Mengen.

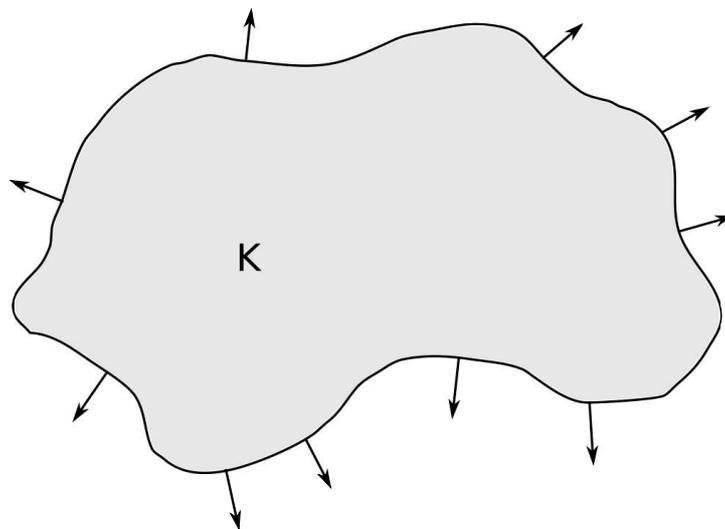


Abbildung 4.26: Eine Menge  $K$  mit einigen Normalenvektoren  $\vec{\nu}(\vec{x})$ . Diese Normalenvektoren stehen immer senkrecht auf dem Rand  $\partial K$ , und sie haben immer die Länge Eins.

**Satz 4.13 (Gaußscher Integralsatz):**

Es sei  $K \subset \mathbb{R}^3$  eine offene beschränkte Menge mit „genügend glattem“ Rand  $\partial K$  (also insbesondere kein Fraktal oder Ähnliches). Sei  $G \supset \overline{K}$  eine offene Menge, und sei  $f: G \rightarrow \mathbb{R}^3$  ein stetig differenzierbares Vektorfeld. Für jeden Punkt  $\vec{x} \in \partial K$  bezeichne  $\vec{\nu}(\vec{x})$  den nach Außen zeigenden Normalenvektor an die Fläche  $\partial K$ .

Dann ist

$$\underbrace{\int_K \operatorname{div} f(\vec{x}) \, d\vec{x}}_{\text{Gebietsintegral}} = \underbrace{\int_{\partial K} \langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x})}_{\text{Oberflächenintegral}}.$$

**Bemerkung 4.14 (Geometrische Interpretation):**

Wir interpretieren den Vektor  $f(\vec{x})$  als die vektorielle Fließgeschwindigkeit einer Flüssigkeit am Ortspunkt  $\vec{x}$ . Was bedeuten dann die beiden Integrale?

Wir beginnen mit dem Oberflächenintegral auf der rechten Seite. Wir nehmen einen Punkt  $\vec{x}$  auf der Oberfläche  $\partial K$ . Dann ist  $\langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle$  das Skalarprodukt der beiden Vektoren  $f(\vec{x})$  und  $\vec{\nu}(\vec{x})$ , also

$$\langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle = \|f(\vec{x})\| \cdot \|\vec{\nu}(\vec{x})\| \cdot \cos \angle (f(\vec{x}), \vec{\nu}(\vec{x})).$$

Wenn der Vektor  $f(\vec{x})$  aus der Menge  $K$  herauszeigt, dann fließt dort die Flüssigkeit aus  $K$  heraus, und der Winkel zwischen  $f(\vec{x})$  und  $\vec{\nu}(\vec{x})$  ist kleiner als  $\pi/2$ , der Cosinus also positiv. Um das Skalarprodukt besser zu verstehen, nehmen wir uns eine Menge  $D \subset \mathbb{R}^2$  und eine reguläre Fläche  $\Phi: D \rightarrow \mathbb{R}^3$ , die in einer Umgebung von  $\vec{x}$  den Rand  $\partial K$  beschreibt. Dann gilt

$$\vec{\nu}(\vec{x}) = \frac{\frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v)}{\left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\|},$$

also folgt dann

$$\begin{aligned} \langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x}) &= \langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \left\| \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\| \, d(u, v) \\ &= \underbrace{\left\langle f(\vec{x}), \frac{\partial \Phi}{\partial u}(u, v) \times \frac{\partial \Phi}{\partial v}(u, v) \right\rangle}_{\text{Spatprodukt}} \, d(u, v), \end{aligned}$$

und dieses Spatprodukt ist gleich dem Volumen des von den drei Vektoren  $f(\vec{x})$ ,  $\partial_u \Phi(u, v)$  und  $\partial_v \Phi(u, v)$  aufgespannten Parallelepipeds (Spats), also gleich dem Flüssigkeitsvolumen, das durch das Oberflächenelement  $d\sigma(\vec{x})$  strömt.

Das bedeutet insgesamt: das Oberflächenintegral im Gaußschen Satz entspricht dem Gesamtfluss des Vektorfelds  $f$  durch die Fläche  $\partial K$ . Dabei zählt der Fluss von innen nach außen positiv, und der Fluss von außen nach innen zählt negativ.

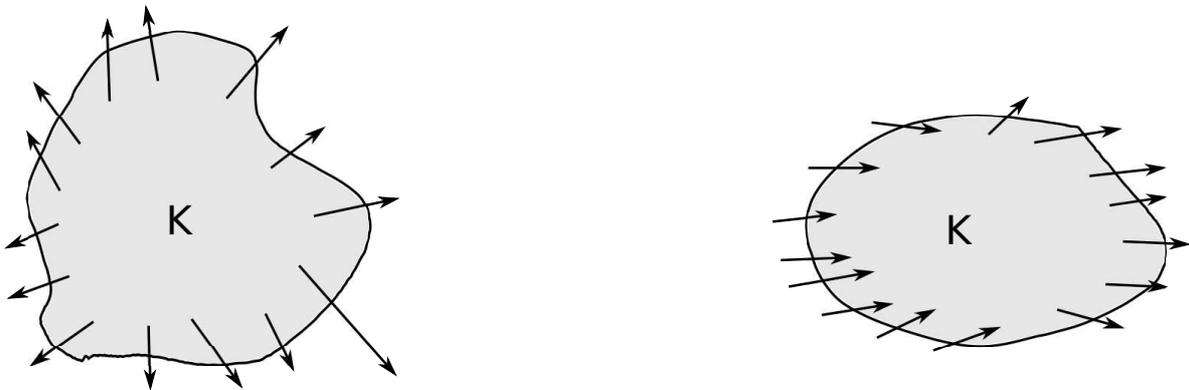


Abbildung 4.27: Links haben wir einen positiven Gesamtfluss über den Rand  $\partial K$ , rechts haben wir den Gesamtfluss gleich Null über  $\partial K$ .

Jetzt überlegen wir uns, was das Gebietsintegral auf der linken Seite des Integralsatzes von Gauß bedeutet. Wir erinnern daran, dass  $\operatorname{div} f(\vec{x})$  angibt, ob die Flüssigkeitsströmung im Punkt  $\vec{x}$  eine Quelle oder eine Senke besitzt. Dann ist  $\int_K \operatorname{div} f(\vec{x}) \, d\vec{x}$  also die Gesamtsumme des in den Quellen entstehenden Flusses, vermindert um die Gesamtsumme des in den Senken versickernden Flusses.

Der Integralsatz von Gauß besagt damit: der Gesamtfluss über den Rand  $\partial K$  ist gleich der Gesamtmenge des im Innern erzeugten oder verschwundenen Flusses.

**Beispiele 4.15:** (a) Gesucht ist  $\int_{\partial K} \langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x})$  für

$$f(x, y, z) = \begin{bmatrix} xy^2 \\ x^2y \\ y \end{bmatrix} \quad \text{und} \quad K = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x^2 + y^2 \leq 1, \quad -1 \leq z \leq 1 \right\}.$$

Wir integrieren also über einen aufrecht stehenden Zylinder. Der Gauß-Satz liefert

$$\begin{aligned} \int_{\partial K} \langle f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x}) &= \int_K \operatorname{div} f(\vec{x}) \, d\vec{x} = \int_K \left( \frac{\partial f_1}{\partial x_1}(\vec{x}) + \frac{\partial f_2}{\partial x_2}(\vec{x}) + \frac{\partial f_3}{\partial x_3}(\vec{x}) \right) \, d\vec{x} \\ &= \int_K (y^2 + x^2 + 0) \, d(x, y, z) \quad \Big| \quad \text{Zylinderkoordinaten} \\ &= \int_{z=-1}^1 \int_{\varphi=0}^{2\pi} \int_{r=0}^1 r^2 \cdot r \, dr \, d\varphi \, dz = 2 \cdot 2\pi \cdot \frac{1}{4} r^4 \Big|_{r=0}^{r=1} = 4\pi \cdot \frac{1}{4} = \pi. \end{aligned}$$

(b)

**Satz 4.16 (Integralsatz von Stokes):**

Sei  $D \subset \mathbb{R}^2$  kompakt,  $G \supset D$  offen und  $\Phi: G \rightarrow \mathbb{R}^3$  zweimal stetig differenzierbar und eine auf  $D$  reguläre Fläche. Sei  $\mathcal{F}$  das durch  $\Phi$  gegebene reguläre Flächenstück und  $R$  sein Rand. Weiter sei  $\gamma$  eine Kurve, die einmal gegen den Uhrzeigersinn entlang  $R$  läuft. Schließlich sei  $U \subseteq \mathbb{R}^3$  offen mit  $\mathcal{F} \subseteq U$  und  $f: U \rightarrow \mathbb{R}^3$  stetig differenzierbar. Dann gilt

$$\underbrace{\int_{\mathcal{F}} \langle \operatorname{rot} f, \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x})}_{\text{Oberflächenintegral}} = \underbrace{\int_{\gamma} f(\vec{x}) \, d\vec{x}}_{\text{Wegintegral}}.$$

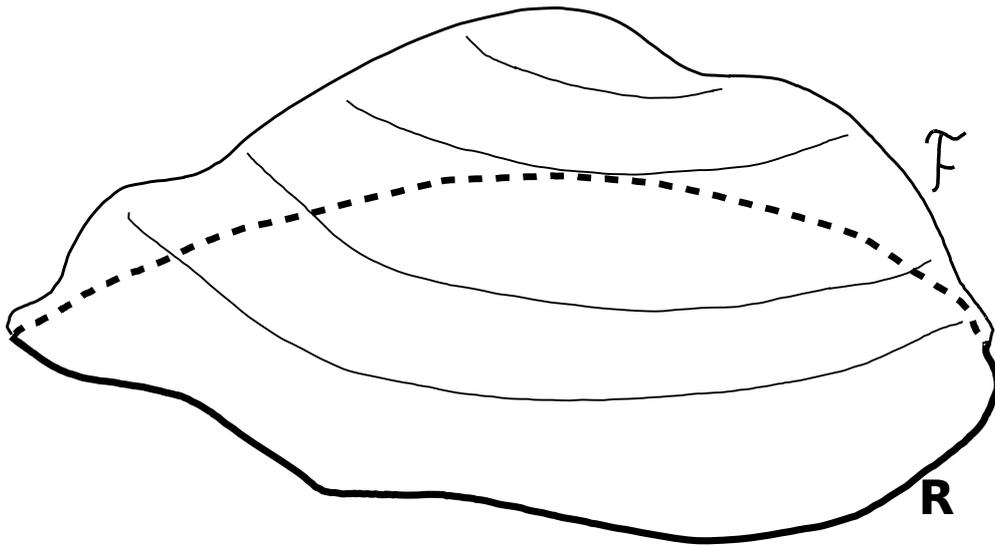


Abbildung 4.28: Ein Flächenstück  $\mathcal{F}$  mit seinem Rand  $R$  (untere fette Linie).

**Beispiel 4.17:**

Wir betrachten erneut  $\mathcal{F}$  als obere Halbkugel mit Radius  $R$ . Wie üblich sei also

$$G = \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 : u^2 + v^2 < R^2 \right\},$$

und die Fläche darüber sei

$$\Phi(u, v) = \begin{bmatrix} u \\ v \\ \sqrt{R^2 - u^2 - v^2} \end{bmatrix}.$$

Wir nehmen als Vektorfeld

$$f(x, y, z) = \begin{bmatrix} -y \\ x \\ 1 \end{bmatrix},$$

und darauf wenden wir den Satz von Stokes an. Es ist also

$$\operatorname{rot} f(\vec{x}) = \begin{bmatrix} \frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \\ \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \\ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \end{bmatrix} = \begin{bmatrix} 0 - 0 \\ 0 - 0 \\ 1 + 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Wir rechnen jetzt das Oberflächenintegral aus:

$$\begin{aligned} \int_{\mathcal{F}} \langle \operatorname{rot} f(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, d\sigma(\vec{x}) &= \int_G \langle \operatorname{rot} f(\Phi(u, v)), \vec{\nu}(\Phi(u, v)) \rangle \|\partial_u \Phi(u, v) \times \partial_v \Phi(u, v)\| \, d(u, v) \\ &= \int_G \left\langle \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \frac{\partial_u \Phi(u, v) \times \partial_v \Phi(u, v)}{\|\partial_u \Phi(u, v) \times \partial_v \Phi(u, v)\|} \right\rangle \|\partial_u \Phi(u, v) \times \partial_v \Phi(u, v)\| \, d(u, v) \\ &= \int_G \left\langle \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} u/\sqrt{R^2 - u^2 - v^2} \\ v/\sqrt{R^2 - u^2 - v^2} \\ 1 \end{bmatrix} \right\rangle \, d(u, v) \\ &= \int_G 2 \, d(u, v) = \int_{\varphi=0}^{2\pi} \int_{r=0}^R 2r \, dr \, d\varphi = 2\pi \cdot r^2 \Big|_{r=0}^{r=R} = 2\pi R^2. \end{aligned}$$

Und jetzt kommen wir zur Berechnung des Wegintegrals: Der Rand von  $\mathcal{F}$  ist der Äquator und demnach gegeben durch

$$\gamma(t) = \begin{bmatrix} R \cos(t) \\ R \sin(t) \\ 0 \end{bmatrix}, \quad t \in [0, 2\pi],$$

und somit folgt dann

$$\begin{aligned} \int_{\gamma} f(\vec{x}) \, d\vec{x} &= \int_{t=0}^{2\pi} \langle f(\gamma(t)), \gamma'(t) \rangle \, dt \quad \Big| \quad f(x, y, z) = [-y, x, 1]^T \\ &= \int_{t=0}^{2\pi} \left\langle \begin{bmatrix} -R \sin(t) \\ R \cos(t) \\ 0 \end{bmatrix}, \begin{bmatrix} -R \sin(t) \\ R \cos(t) \\ 0 \end{bmatrix} \right\rangle \, dt \\ &= \int_{t=0}^{2\pi} (R^2 \sin^2(t) + R^2 \cos^2(t)) \, dt = 2\pi R^2. \end{aligned}$$

Das wollten wir haben.

# Index

- Abbildung
  - lineare, 28, 37
- Abbildungsmatrix, 29, 31, 44
- abgeschlossene Kugel, 62
- abgeschlossene Menge, 62
- Ableitung
  - partielle, 63
  - totale, 70
  - zweiter Ordnung, 68
- Addition von Matrizen, 22
- Affensattel, 77
- antisymmetrische Matrix, 25
- Arbeitsintegral, 95
- Aufpunkt, 21
  
- Basis, 10
- Basiswechsel, 34
- Basiswechselmatrizen, 36
- Baukastenprinzip, 60
- beschränkte Menge, 62
- bijektiv, 38
  
- charakteristisches Polynom, 46
- Cramer'sche Regel, 44
  
- Definitionsbereich, 55
- Determinante, 39
  - Rechenregeln, 41
- Diagonalisierbarkeit, 45
- Diagonalmatrizen, 36
- differenzierbar
  - partiell, 63, 67
  - stetig partiell, 63
  - total, 70
- differenzierbare Funktion, 63
- Dimension, 11
- Dimensionsformel, 20
- Divergenz
  - eines Vektorfelds, 108
- Divergenz von Folgen, 59
- Drehung, 28, 33, 50
- Dreiecksmatrix, 40
  
- Eigenraum, 45
- Eigenvektor, 44
- Eigenwert, 44
- einfach zusammenhängend, 99
- Einheitsmatrix, 27
- Elementarumformungen, 17, 26
- Ellipse, 51
  
- Extremum
  - globales, 77
  - hinreichende Bedingung, 78
  - lokales, 77
  - notwendige Bedingung, 77
- Extremwert, 77
  
- Feld
  - konservatives, 96
- Fläche, 110
- Flächeninhalt, 112
  - Parallelogramm, 40
- Flächenstück, 110
- Folgenkonvergenz, 59
  
- Gauß-Algorithmus, 19, 26, 37, 42, 48
- Gaußglockenfunktion, 87
- Gaußscher Integralsatz, 114
- Gebietsintegral, 101
- geschlossene Kurve, 92
- Gradient, 65
- Gradientenfeld, 96, 109
- Grenzwertsätze, 59
  
- Hauptachsentransformation, 50
- Hesse-Matrix, 68
- hinreichende Bedingung
  - für Extremum, 78
- Hyperbel, 51
- Häufungspunkt, 62
- Höhenlinien, 56, 65
- Hülle
  - lineare, 7
  
- identische Abbildung, 56
- indefinit, 78
- inertierbar, 27
- injektiv, 38
- innerer Punkt, 62
- Integrabilitätsbedingung, 98, 109
- Integral
  - über ein Gebiet, 101
  - über ein Intervall, 100
- Integralsatz
  - von Gauß, 114
  - von Stokes, 115
- Integrationsreihenfolge, 100
- Intervall
  - im  $\mathbb{R}^n$ , 100
- inverse Matrix, 27, 38

- Invertierbarkeitskriterium, 28
- Jacobi-Matrix, 67
- Kegelschnitt, 49
- Kegelvolumen, 107
- Kern, 20, 21, 28, 38
- Kettenregel, 67
- kompakte Menge, 62
- komplexe Vektorräume, 6
- konservatives Feld, 96
- Konvergenz von Folgen, 59
- Koordinaten, 5
  - bezüglich einer Basis, 11, 29
  - bezüglich einer ONB, 12
- Koordinatenfunktion, 55
- koordinatenweise Konvergenz, 59
- Kraftfeld, 90
- Kriterien für Untervektorräume, 6
- Kriterium
  - für Invertierbarkeit, 28
- kritischer Punkt, 77
- Kronecker-Delta, 12
- Kugelkoordinaten, 108
- Kugelvolumen, 108
- Kurve, 91
  - geschlossene, 92
  - Länge, 94
  - Tangente daran, 92
- Kurvenintegral, 90
- LGS, 12, 16
  - homogen, 20, 21
  - inhomogen, 20, 21
  - Lösungsstruktur, 20
  - Matrixform, 16
- lineare Abbildung, 28
  - Verknüpfung, 37
- lineare Abhängigkeit, 9
- lineare Gleichungssysteme, 12
- lineare Hülle, 7
- lineare Unabhängigkeit, 9
- linearer Teilraum, 6
- lineares Gleichungssystem
  - überbestimmt, 16
  - quadratisch, 16
  - unterbestimmt, 16
- Linearkombination, 7
- Länge einer Kurve, 94
- Magnetfeld, 110
- Matrix, 21
  - Addition, 22
  - antisymmetrische, 25
  - diagonalisierbare, 45, 49
  - indefinite, 78
  - inverse, 27, 38
  - invertierbare, 27
  - negativ definite, 78
  - nicht diagonalisierbare, 48
  - orthogonale, 38, 49
  - positiv definite, 78
  - quadratische, 22
  - Rang, 20
  - reguläre, 27
  - singuläre, 27
  - Skalarmultiplikation, 22
  - symmetrische, 25, 49
  - transponierte, 25
  - zu einer Abbildung, 29, 31
- Matrix-Vektor-Multiplikation, 24
- Matrixgleichungssysteme, 26
- Matrixmultiplikation, 23
- Matrixprodukt
  - nicht kommutativ, 25
- Matrizen
  - diagonale, 36
  - Rechenregeln, 23
- Maximum
  - globales, 77
  - lokales, 77
- Maximumprinzip, 62
- mechanische Arbeit, 90, 95
- Menge
  - abgeschlossene, 62
  - beschränkte, 62
  - kompakte, 62
  - offene, 62
  - projizierbare, 102
- Minimum
  - globales, 77
  - lokales, 77
- Multiplikation
  - von Matrix und Vektor, 24
  - von Matrizen, 23
- Nabla, 65
- negativ definit, 78
- negativ semi-definit, 78
- Niveaulinien, 56, 65
- Norm, 58
- Normalenvektor, 111
- Normalform, 50
- notwendige Bedingung für Extremum, 77
- Nullmatrix, 22
- Oberflächenintegral, 113
- offene Kugel, 62
- offene Menge, 62
- ONB, 12, 50
- Orthogonalbasen, 12
- orthogonale Matrix, 38, 49
- Orthonormalbasen, 12
- Ortsvektor, 5
- Parabel, 51
- Parallelepiped, 41
- Parallelogramm

- Flächeninhalt, 40
- Parameterintegral, 87
- partiell differenzierbar, 63, 67
- partielle Ableitung, 63
- Partikulärlösung, 20
- Pfeilbild, 58
- Polarkoordinaten, 56, 67, 105
- Polynom
  - charakteristisches, 46
- positiv definit, 78
- positiv semi-definit, 78
- Potential, 96, 98
- Potentialfeld, 96
  - Stammfunktion, 97
- Projektion, 28, 33
- projizierbare Menge, 102
- Punkt
  - kritischer, 77
- quadratische Ergänzung, 50
- quadratische Form, 49
- quadratische Gleichung, 49
- quadratische Matrix, 22
- Quadrik, 49
- Quelle, 109
- quellenfrei, 109
- Rand einer Menge, 62
- Rang, 20, 28, 38
- Rechenregeln
  - für das Matrixprodukt, 24
  - für Matrizen, 23
  - für Determinanten, 41
  - für die inverse Matrix, 27
  - für die Transposition, 25
  - für Eigenwerte und Eigenvektoren, 47
  - für Vektoren, 5
- reguläre Matrix, 27
- Restglied, 73, 74
- Richtungsableitung, 63, 71
- Richtungsfeld, 58
- Rotation
  - eines Vektorfelds, 109
- Sattelpunkt, 77
- Satz von Schwarz, 69
- Schattenwurfabbildung, 33
- Scherung, 29
- Schraubenlinie, 92
- semi-definit
  - negativ, 78
  - positiv, 78
- Senke, 109
- singuläre Matrix, 27
- Skalarfeld, 55
- Skalarmultiplikation von Matrizen, 22
- Spat
  - Volumen, 41
- Spiegelung, 28, 32, 45
- Stammfunktion, 97
- Standardbasis, 10
- stetig partiell differenzierbar, 63
- Stetigkeit, 60
- Stokes-Integralsatz, 115
- Streckung, 29
- Substitutionsregel, 105
- surjektiv, 38
- symmetrische Matrix, 25, 49
- Tangente, 76
- Tangente an eine Kurve, 92
- Tangentialebene, 76, 111
- Taylorpolynom, 73
- Taylorscher Satz, 74
- Teilraum
  - linearer, 6
- topologische Begriffe, 62
- total differenzierbar, 70
- Transformation
  - der Basis, 34
- transponierte Matrix, 25
- triviale Untervektorräume, 7
- Umrechnungsformel
  - für Abbildungsmatrizen, 35
- Untermioren, 79
- Unterraum, 6
- Untervektorraum, 6, 9, 20, 45
- Untervektorraumkriterium, 6
- Untervektorräume
  - triviale, 7
- Vektoren
  - Rechenregeln, 5
- Vektorfeld, 55
  - Divergenz, 108
  - quellenfrei, 109
  - Rotation, 109
- Vektorraum, 5
- Verbindungsstrecke, 91
- Verkettungsprinzip, 60
- Verschiebungsfeld, 108
- Volumen
  - eines Spats, 41
- Volumen eines Kegels, 107
- Wechsel
  - der Basis, 34
- Wirbeldichte, 109
- Wärmeleitungsgleichung, 87
- Zeilenstufenform, 17–20, 26
- Zylinderkoordinaten, 107