

Höhere Mathematik II

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

Sommersemester 2013

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

1	Mehrdimensionale Differentialrechnung	5
1	Der Raum \mathbb{R}^n ; Skalarfelder; Vektorfelder	5
2	Folgen im \mathbb{R}^n ; Stetigkeit	9
3	Partielle Ableitungen; Gradient; höhere Ableitungen	11
4	Totale Differenzierbarkeit	17
2	Matrixrechnung, Determinanten und Eigenwerte	21
5	Erinnerungen	21
6	Determinanten	21
7	Eigenwerttheorie	27
3	Extremwertprobleme in mehreren Variablen	33
8	Notwendiges Kriterium und hinreichendes Kriterium	33
9	Extremwerte unter Nebenbedingungen	35
4	Gewöhnliche Differentialgleichungen	41
10	Erste Beispiele und Begriffe	41
11	Getrennte Variablen	42
12	Lineare Differentialgleichungen erster Ordnung	44
13	Systeme linearer Differentialgleichungen erster Ordnung	46
14	Der harmonische Oszillator	49
5	Integration von Funktionen mehrerer Variabler	51
15	Zweidimensionale Integrale	51
16	Dreidimensionale Integrale	53
17	Kurvenintegrale	55

Kapitel 1

Mehrdimensionale Differentialrechnung

1 Der Raum \mathbb{R}^n ; Skalarfelder; Vektorfelder

Wir betrachten in diesem Kapitel jetzt Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Beispiel 1.1. (a) lineare Abbildungen wie in HM I:

$$f(x) = Ax, \quad A \in \mathbb{R}^{m \times n},$$

(b) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^2 + y^2$,

(c) $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ wie etwa Temperatur, Druck, Konzentration (im Raum),

(d) $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ wie etwa Windgeschwindigkeit, elektrostatisches Kraftfeld, Gravitationskraftfeld (im Raum),

(e) die VAN DER WAALSsche Zustandsgleichung eines Gases: seien p der Druck, t die Temperatur und v das Molvolumen, R die allgemeine Gaskonstante und a, b gasspezifische Konstanten. Dann ist

$$p = p(t, v) = \frac{R \cdot t}{v - b} - \frac{a}{v^2}$$

und $p: [t_{kl}, t_{gr}] \times [v_{kl}, v_{gr}] \rightarrow \mathbb{R}$ bzw. präziser $\rightarrow (0, \infty)$. Hierbei sind t_{kl} und t_{gr} die kleinste bzw. größte jemals uns interessierende Temperatur, analog v_{kl} und v_{gr} .

Bemerkung 1.2. Eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ heißt Skalarfeld und eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (mit $m \geq 2$) nennt man Vektorfeld. Allgemein redet man von Feldern.

Beispiel 1.3. Das Vektorfeld $f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$ kann veranschaulicht werden wie in Abbildung 1.1. Beispiele für Skalarfelder finden sich in Abbildungen 1.2–1.6.

Im weiteren Kapitelverlauf bestehen unsere **Ziele** darin, folgende Begriffe zu verstehen:

- Stetigkeit von Feldern,
- Ableitungen von Feldern,
- Extremwertaufgaben für Skalarfelder (und wie man sie löst),
- Integration.

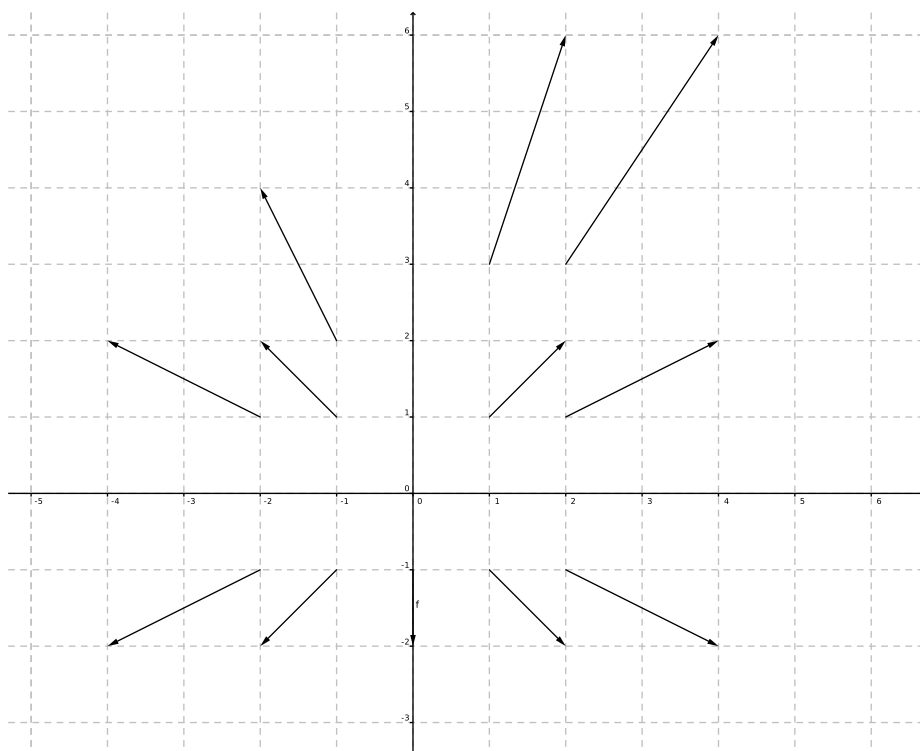


Abbildung 1.1: Das Vektorfeld $f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$. An jedem Punkt mit den Koordinaten $\begin{pmatrix} x \\ y \end{pmatrix}$ wird ein Pfeil mit den Koordinaten $\begin{pmatrix} x \\ y \end{pmatrix}$ angeheftet.

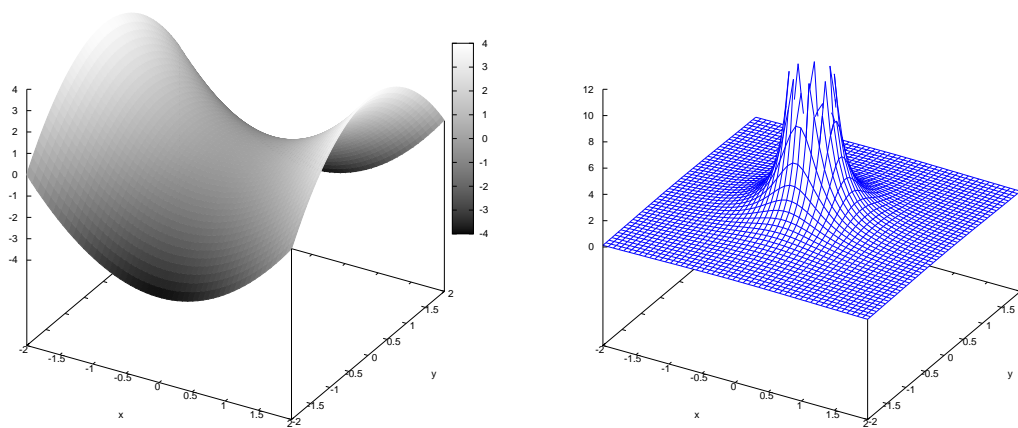


Abbildung 1.2: Links: die skalare Funktion $f(x, y) = x^2 - y^2$ über dem Quadrat $[-2, 2] \times [-2, 2]$. Die Höhe ist zusätzlich farblich kodiert: je heller die Graphik ist, umso größer ist dort $f(x, y)$. Rechts: die skalare Funktion $f(x, y) = 1/(x^2 + y^2)$ über dem Quadrat $[-2, 2] \times [-2, 2]$. Man sieht die Polstelle für $(x, y)^\top = (0, 0)^\top$. Vergleiche auch Beispiel 2.8.

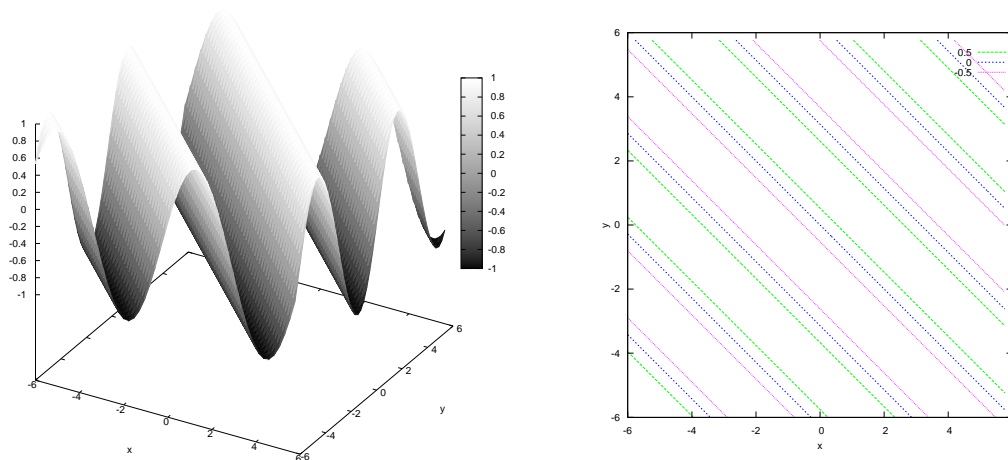


Abbildung 1.3: Links: die skalare Funktion $f(x, y) = \sin(x + y)$ über dem Quadrat $[-6, 6] \times [-6, 6]$. Die Höhe ist zusätzlich farblich kodiert. Rechts ist ein Höhenlinienplot derselben Funktion: entlang einer jede einfarbigen Linie ist f jeweils konstant.

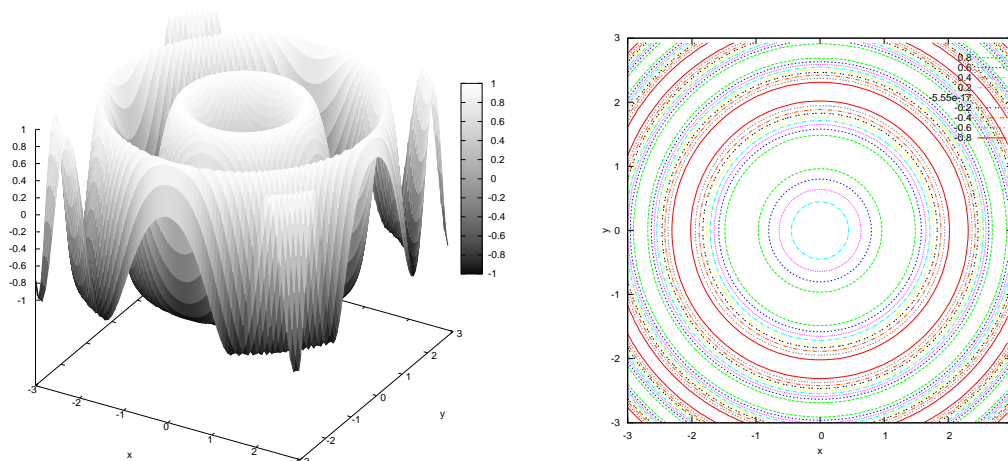


Abbildung 1.4: Links: die skalare Funktion $f(x, y) = \sin(x^2 + y^2)$ über dem Quadrat $[-3, 3] \times [-3, 3]$. Rechts ist der zugehörige Höhenlinienplot. Man beachte, daß die „Jahresringe“ nach außen immer enger aneinanderrücken.

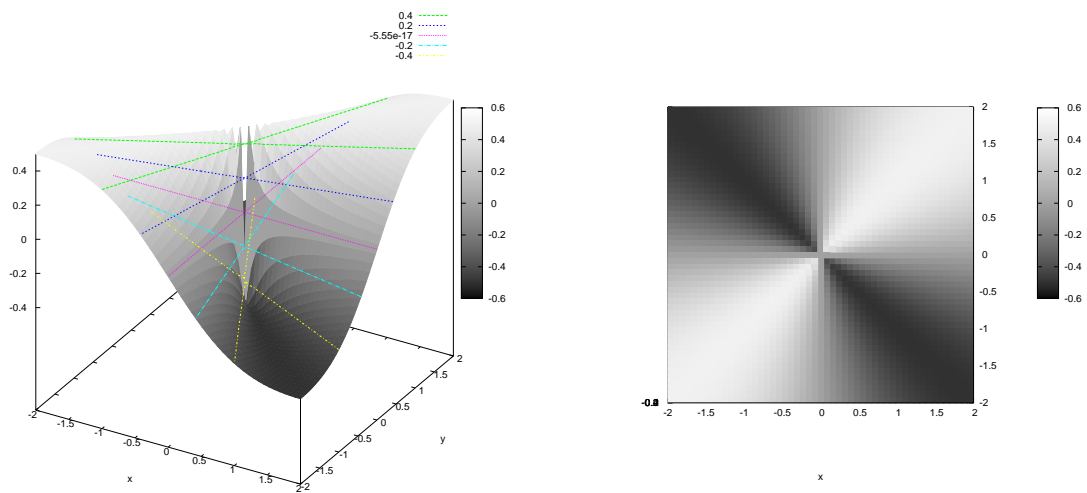


Abbildung 1.5: Links: die skalare Funktion $f(x,y) = \frac{xy}{x^2+y^2}$ über dem Quadrat $[-2,2] \times [-2,2]$. Die farbigen Linien sind Geraden, die in den Graphen eingezeichnet sind, und dort hat f jeweils den gleichen Wert. Rechts sieht man den Graphen derselben Funktion in der Draufsicht. Diese Funktion ist unstetig im Nullpunkt. Siehe auch die nächste Abbildung und Beispiel 2.11.

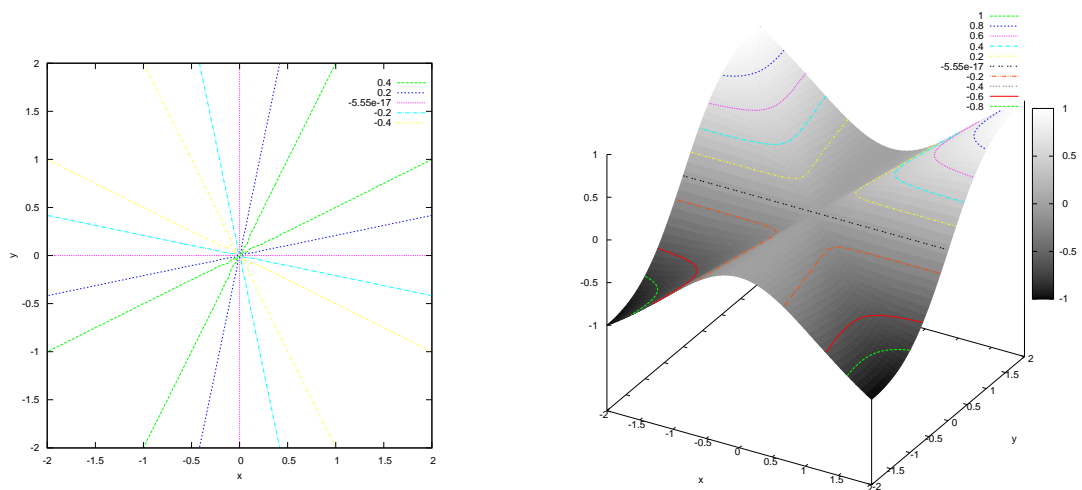


Abbildung 1.6: Links ist der Höhenlinienplot zur Funktion $f(x,y) = \frac{xy}{x^2+y^2}$. Rechts ist der Graph zur stetigen Funktion $f(x,y) = \frac{x^2y}{x^2+y^2}$. Siehe auch Beispiel 2.12.

2 Folgen im \mathbb{R}^n ; Stetigkeit

Für die versprochene Stetigkeit benötigen wir den Begriff einer Folge (und den dazugehörigen Konvergenzbegriff).

Erinnerung: Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ heißt *stetig* in $x^* \in \mathbb{R}$, falls für alle Folgen $(x_j)_{j \in \mathbb{N}}$ in \mathbb{R} mit $\lim_{j \rightarrow \infty} x_j = x^*$ gilt, daß $f(x_j) \rightarrow f(x^*)$ gilt für $j \rightarrow \infty$, oder anders geschrieben:

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

Definition 2.1. Eine Abbildung $a: \mathbb{N} \rightarrow \mathbb{R}^n$ heißt Folge im \mathbb{R}^n . Wir schreiben a_j anstatt $a(j)$, wobei $j \in \mathbb{N}$, oder noch besser $a^{(1)}$, $a^{(2)}$, $a^{(3)}$, da z.B. $a_2^{(3)}$ die zweite Komponente des dritten Folgenglieds $a^{(3)} \in \mathbb{R}^n$ ist.

Beispiel 2.2. (a) Sei $a^{(j)} = \left(\frac{1}{j}, \frac{1}{j^2}\right)$ für $j \in \mathbb{N}$. Das ergibt die Folge

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/3 \\ 1/9 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/16 \end{pmatrix}, \dots$$

(b) Sei $a^{(j)} = (2j, 3j, 4j)^\top$ für $j \in \mathbb{N}$. Das ergibt die Folge

$$\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix}, \begin{pmatrix} 6 \\ 9 \\ 12 \end{pmatrix}, \begin{pmatrix} 8 \\ 12 \\ 16 \end{pmatrix}, \dots$$

(c) Sei $a^{(j)} = (j, 1/j)^\top$ für $j \in \mathbb{N}$. Das ergibt die Folge

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1/3 \end{pmatrix}, \begin{pmatrix} 4 \\ 1/4 \end{pmatrix}, \dots$$

(d) Sei eine Folge $(a^{(j)})_{j \in \mathbb{N}}$ gegeben durch das Bildungsgesetz

$$(a^{(j)})_{j \in \mathbb{N}} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \dots \right).$$

Das Prinzip hinter dieser Folge kann man durch Veranschaulichung in einem Koordinatensystem erraten.

Bemerkung 2.3. Sei $(a^{(j)})_{j \in \mathbb{N}}$ eine Folge im \mathbb{R}^n . Für jedes $k = 1, \dots, n$ ist dann $(a_k^{(j)})_{j \in \mathbb{N}}$ eine Folge in \mathbb{R} , die k -te Koordinatenfolge genannt wird. Im Unterbeispiel (b) von Beispiel 2.2 ist $(a_3^{(j)})_{j \in \mathbb{N}} = (4, 8, 12, 16, \dots)$, und im Unterbeispiel (c) ist $(a_1^{(j)})_{j \in \mathbb{N}} = (1, 2, 3, \dots)$.

Definition 2.4. Sei $(a^{(j)})_{j \in \mathbb{N}}$ eine Folge im \mathbb{R}^n .

(a) Die Folge $(a^{(j)})_{j \in \mathbb{N}}$ heißt beschränkt genau dann, wenn jede ihrer Koordinatenfolgen $(a_k^{(j)})_{j \in \mathbb{N}}$ beschränkt ist, für $k = 1, \dots, n$.

(b) Wenn jede Koordinatenfolge $(a_k^{(j)})_{j \in \mathbb{N}}$ konvergent ist mit Grenzwert $a_k^* := \lim_{j \rightarrow \infty} a_k^{(j)}$, dann nennen wir die Folge $(a^{(j)})_{j \in \mathbb{N}}$ konvergent, und es ist

$$\lim_{j \rightarrow \infty} a^{(j)} = \begin{pmatrix} \lim_{j \rightarrow \infty} a_1^{(j)} \\ \vdots \\ \lim_{j \rightarrow \infty} a_n^{(j)} \end{pmatrix} = \begin{pmatrix} a_1^* \\ \vdots \\ a_n^* \end{pmatrix} =: a^*.$$

Beispiel 2.5. Wir illustrieren diese Begriffe anhand Beispiel 2.2.

(a) Diese Folge ist beschränkt und konvergent mit Grenzwert $a^* = \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

(b) Diese Folge ist weder beschränkt noch konvergent.

(c) Diese Folge ist weder beschränkt noch konvergent, da ihre erste Koordinatenfolge weder beschränkt noch konvergent ist. Es reicht nicht aus, daß die zweite Koordinatenfolge konvergiert.

(d) Diese Folge ist beschränkt, aber sie konvergiert nicht.

Satz 2.6. Sei $(a^{(j)})_{j \in \mathbb{N}}$ eine Folge im \mathbb{R}^n . Dann gilt:

- (a) Die Folge $(a^{(j)})_{j \in \mathbb{N}}$ ist beschränkt genau dann, wenn es eine Zahl $M > 0$ gibt mit $|a^{(j)}| \leq M$ für jedes $j \in \mathbb{N}$. (Wir erinnern an die Schreibweise $|a^{(j)}| := \sqrt{(a_1^{(j)})^2 + \dots + (a_n^{(j)})^2}$.)
- (b) Die Folge $(a^{(j)})_{j \in \mathbb{N}}$ ist konvergent mit Grenzwert $a^* \in \mathbb{R}^n$ genau dann, wenn es zu jedem positiven ε ein $J_\varepsilon \in \mathbb{N}$ gibt mit $|a^{(j)} - a^*| \leq \varepsilon$ für jedes $j \geq J_\varepsilon$ (vgl. HM I).
- (c) Wenn die Folge $(a^{(j)})_{j \in \mathbb{N}}$ konvergiert, dann ist sie auch beschränkt. Die Umkehrung ist im Allgemeinen falsch.

Definition 2.7. Sei $D \subseteq \mathbb{R}^n$ eine Menge, und sei $f: D \rightarrow \mathbb{R}^m$ eine Funktion. Sei $x^* \in D$ ein Punkt.

- (a) Es heißt f stetig in x^* genau dann, wenn für jede Folge $(x^{(j)})_{j \in \mathbb{N}}$ mit $\lim_{j \rightarrow \infty} x^{(j)} = x^*$ und $x^{(j)} \in D$ (für jedes $j \in \mathbb{N}$) gilt, daß $\lim_{j \rightarrow \infty} f(x^{(j)}) = f(\lim_{j \rightarrow \infty} x^{(j)}) = f(x^*)$.
- (b) Die Funktion f heißt stetig in D genau dann, wenn f stetig in x^* ist, für alle $x^* \in D$.

Beispiel 2.8. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y) := \begin{cases} \frac{1}{x^2 + y^2} & : (x, y)^\top \neq (0, 0)^\top, \\ 0 & : (x, y)^\top = (0, 0)^\top. \end{cases}$$

Wir wollen diese Funktion auf Stetigkeit in einem Punkt $(x^*, y^*)^\top$ untersuchen. Siehe auch Abbildung 1.2.

Fall 1: $(x^*, y^*)^\top \neq (0, 0)^\top$. Sei $\begin{pmatrix} x^{(j)} \\ y^{(j)} \end{pmatrix}$ eine Folge im \mathbb{R}^2 mit Grenzwert $(x^*, y^*)^\top$. Dann ist $\begin{pmatrix} x^{(j)} \\ y^{(j)} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ sobald j oberhalb einer gewissen Zahl N_0 liegt, und es ist $\lim_{j \rightarrow \infty} ((x^{(j)})^2 + (y^{(j)})^2) = (x^*)^2 + (y^*)^2 \neq 0$, also ist auch

$$\lim_{j \rightarrow \infty} f(x^{(j)}, y^{(j)}) = \lim_{j \rightarrow \infty} \frac{1}{(x^{(j)})^2 + (y^{(j)})^2} = \frac{1}{(x^*)^2 + (y^*)^2} = f(x^*, y^*),$$

also ist f stetig im Punkt $(x^*, y^*)^\top$.

Fall 2: $(x^*, y^*)^\top = (0, 0)^\top$. Wir wählen eine spezielle Folge, nämlich

$$\begin{pmatrix} x^{(j)} \\ y^{(j)} \end{pmatrix} := \begin{pmatrix} 1/j \\ 1/j \end{pmatrix},$$

und dann gilt $\begin{pmatrix} x^{(j)} \\ y^{(j)} \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ für $j \rightarrow \infty$, und es ist

$$f(x^{(j)}, y^{(j)}) = \frac{1}{\frac{1}{j^2} + \frac{1}{j^2}} = \frac{j^2}{2} \rightarrow +\infty \neq 0 = f(0, 0),$$

für $j \rightarrow \infty$. Also ist f nicht stetig im Punkt $(x^*, y^*)^\top = (0, 0)^\top$.

Bemerkung 2.9. Sei $D \subseteq \mathbb{R}^n$, $f: D \rightarrow \mathbb{R}^m$ und $x^* \in D$. Dann gilt:

f ist stetig in x^*

\iff für jede Folge $(x^{(j)})_{j \in \mathbb{N}}$ in D mit $x^{(j)} \rightarrow x^*$ (wobei $j \rightarrow \infty$) gilt $f(x^{(j)}) \rightarrow f(x^*)$ (wobei $j \rightarrow \infty$)

\iff für jede Folge $(x^{(j)})_{j \in \mathbb{N}}$ in D mit $x^{(j)} \rightarrow x^*$ (wobei $j \rightarrow \infty$) gilt $f(x^{(j)})_k \rightarrow f(x^*)_k$ für alle $k = 1, \dots, m$.

Deshalb konzentrieren wir uns im Folgenden auf Funktionen mit Werten in \mathbb{R}^1 anstatt \mathbb{R}^m .

Satz 2.10 (Baukastenprinzip). Es seien $D \subseteq \mathbb{R}^n$, $x^* \in D$ und $f, g: D \rightarrow \mathbb{R}$ beide stetig in x^* . Dann gilt:

- (a) $f + g$, $f - g$, $f \cdot g$ sind stetig in x^* .
- (b) Wenn $g(x^*) \neq 0$ ist, dann ist $\frac{f}{g}$ stetig in x^* .

(c) Ist $U \subset \mathbb{R}$ eine Umgebung von $y^* := f(x^*)$ und ist $h: U \rightarrow \mathbb{R}$ stetig, so ist die zusammengesetzte Funktion $h \circ f$ in x^* stetig.

Beispiel 2.11. (Siehe auch Abbildung 1.5.) Wir betrachten die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch

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

Es ist f stetig im Punkt $(x^*, y^*)^\top \neq (0, 0)^\top$, denn dort sind die beiden Funktionen

$$(x, y) \mapsto \frac{1}{x^2+y^2} \quad \text{und} \quad (x, y) \mapsto xy$$

stetig (siehe Beispiel 2.8), gemäß des Baukastenprinzips also auch ihr Produkt.

Jedoch ist f nicht stetig im Punkt $(x^*, y^*)^\top = (0, 0)^\top$. Um das zu zeigen, betrachten wir Punkte $(x_j, y_j)^\top = (\frac{1}{j}, \frac{1}{j})^\top$. Die Folge $(x_j, y_j)_{j \in \mathbb{N}}$ dieser Punkte konvergiert zum Grenzwert $(x^*, y^*)^\top = (0, 0)^\top$, allerdings ist

$$\lim_{j \rightarrow \infty} f(x_j, y_j) = \lim_{j \rightarrow \infty} \frac{\frac{1}{j} \cdot \frac{1}{j}}{\frac{1}{j^2} + \frac{1}{j^2}} = \lim_{j \rightarrow \infty} \frac{\frac{1}{j^2}}{\frac{2}{j^2}} = \lim_{j \rightarrow \infty} \frac{1}{2} = \frac{1}{2} \neq 0 = f(0, 0),$$

also ist f nicht stetig in $(0, 0)^\top$, wie angekündigt.

Beispiel 2.12. (Siehe auch den rechten Teil von Abbildung 1.6.) Wir wandeln den Zähler des letzten Beispiels ein wenig ab und betrachten die Funktion $g: \mathbb{R}^2 \rightarrow \mathbb{R}$, die gegeben ist durch

$$g(x, y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & : (x, y)^\top \neq (0, 0)^\top, \\ 0 & : (x, y)^\top = (0, 0)^\top. \end{cases}$$

Nach Baukastenprinzip und vorigem Beispiel ist g stetig an jedem Punkt außerhalb des Ursprungs. Aber g ist auch stetig im Ursprung $(0, 0)^\top$, denn sei $(x_j, y_j)_{j \in \mathbb{N}}^\top$ irgendeine Folge, die gegen $(0, 0)^\top$ konvergiert. Dann gilt $\lim_{j \rightarrow \infty} |y_j| = 0$ und es ist

$$0 \leq |g(x_j, y_j)| = \left| \frac{x_j^2 y_j}{x_j^2 + y_j^2} \right| = \frac{x_j^2 |y_j|}{x_j^2 + y_j^2} \leq \frac{x_j^2 |y_j|}{x_j^2} = |y_j|,$$

also $0 \leq \lim_{j \rightarrow \infty} |g(x_j, y_j)| \leq \lim_{j \rightarrow \infty} |y_j| = 0$, woraus wir $\lim_{j \rightarrow \infty} |g(x_j, y_j)| = 0$ folgern, also $\lim_{j \rightarrow \infty} g(x_j, y_j) = 0 = g(0, 0)$.

Definition 2.13. Sei $D \subseteq \mathbb{R}^n$, $f: D \rightarrow \mathbb{R}^m$ und $x^* \in D$. Wenn für alle Folgen $(x^{(j)})_{j \in \mathbb{N}}$ in D mit $x^{(j)} \rightarrow x^*$ (wobei $j \rightarrow \infty$) der Grenzwert $\lim_{j \rightarrow \infty} f(x^{(j)})$ existiert, und wenn er immer derselbe Grenzwert ist, dann schreibt man für diesen Grenzwert den Ausdruck $\lim_{x \rightarrow x^*} f(x)$.

Bemerkung 2.14. Dann gilt: die Funktion f ist stetig in x^* genau dann, wenn $\lim_{x \rightarrow x^*} f(x) = f(x^*)$.

Wie in HM I gilt:

Satz 2.15. Ist $D \subseteq \mathbb{R}^n$ abgeschlossen und beschränkt, und ist $f: D \rightarrow \mathbb{R}$ stetig in D , so nimmt f auf D sein Minimum und sein Maximum an. Das heißt, es gibt Punkte x_* und x^* in D mit

$$\min_{y \in D} f(y) = f(x_*) \leq f(x) \leq f(x^*) = \max_{y \in D} f(y), \quad \text{für alle } x \in D.$$

3 Partielle Ableitungen; Gradient; höhere Ableitungen

Unser **Ziel** ist es, den bisher bekannten Ableitungsbegriff auf Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ zu übertragen.

In HM I hatten wir vereinbart:

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

Jetzt gibt es aber zwei kleine Probleme:

technisches Problem: wir brauchen in allen Richtungen Platz im Definitionsgebiet D_f um den Punkt x_0 herum, um den Limes bilden zu können. Es darf x_0 nicht auf dem Rand von D_f liegen.

Hauptproblem: es gibt keine Division durch Vektoren (und es wird auch nie eine geben). Für $x, x_0 \in \mathbb{R}^n$ und $n \geq 2$ ist Division durch $x - x_0 \in \mathbb{R}^n$ nicht möglich !

Definition 3.1. (a) Für $x^* \in \mathbb{R}^n$ und $r > 0$ ist $K(x^*, r) := \{x \in \mathbb{R}^n : |x - x^*| < r\}$ die (offene) Kugel um x^* mit Radius r .

(b) Die Menge $D \subseteq \mathbb{R}^n$ heißt offen, wenn für jedes $x^* \in D$ ein Radius $\varepsilon > 0$ existiert mit $K(x^*, \varepsilon) \subseteq D$.

(c) Die Menge $D \subseteq \mathbb{R}^n$ heißt abgeschlossen, wenn die Komplementmenge $\mathbb{R}^n \setminus D$ offen ist.

Bemerkung 3.2. Offen bedeutet: jeder Punkt in D ist ganz von D umgeben; jeder Punkt von D liegt im Innern von D . Insbesondere enthalten offene Mengen keine Randpunkte.

Satz 3.3. Eine Menge $D \subseteq \mathbb{R}^n$ ist abgeschlossen genau dann, wenn für jede Folge $(x^{(j)})_{j \in \mathbb{N}}$ in \mathbb{R}^n , die in D verbleibt und einen Grenzwert besitzt, gilt, daß dieser Grenzwert in D liegt. Anders formuliert: wenn $x^{(j)} \in D$ für alle $j \in \mathbb{N}$ und wenn $\lim_{j \rightarrow \infty} x^{(j)}$ existiert, dann ist $\lim_{j \rightarrow \infty} x^{(j)} \in D$.

Anschaulich gesprochen: wenn eine Menge D abgeschlossen ist, dann ist der gesamte Rand von D ein Teilmenge von D .

Beispiel 3.4. Seien $a, b \in \mathbb{R}^n$ mit $a_k < b_k$ für alle $k = 1, \dots, n$.

(a) Die Menge $[a, b] := \{x \in \mathbb{R}^n : a_k \leq x_k \leq b_k \forall k = 1, \dots, n\}$ heißt abgeschlossenes Intervall in \mathbb{R}^n .

(b) Die Menge $(a, b) := \{x \in \mathbb{R}^n : a_k < x_k < b_k \forall k = 1, \dots, n\}$ heißt offenes Intervall in \mathbb{R}^n .

(c) Die Menge $\{x \in \mathbb{R}^n : a_k \leq x_k < b_k \forall k = 1, \dots, n\}$ ist weder offen noch abgeschlossen.

Warnung: Mengen sind keine Türen ! Die meisten Mengen sind weder offen noch abgeschlossen.

Damit sind alle Begriffe bereitgelegt für die Behandlung des technischen Problems, und wir kümmern jetzt um das Hauptproblem.

Beispiel 3.5. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x, y) = x \cdot e^y$, und wir wählen einen Punkt $(x^*, y^*)^\top \in \mathbb{R}^2$.

(a) Wir fixieren $y = y^*$ und lassen lediglich x beweglich. Wir erhalten dann eine Funktion $g: \mathbb{R} \rightarrow \mathbb{R}$ mit der Gleichung $g(x) = x \cdot e^{y^*}$. Die Ableitung (wie in HM I) von g ist

$$g'(x^*) = e^{y^*} \left(= \lim_{h \rightarrow 0} \frac{f(x^* + h, y^*) - f(x^*, y^*)}{h} \right).$$

(b) Wir fixieren $x = x^*$ und lassen lediglich y beweglich. Wir erhalten dann eine Funktion $k: \mathbb{R} \rightarrow \mathbb{R}$ mit der Gleichung $k(y) = x^* \cdot e^y$. Die Ableitung (wie in HM I) von k ist

$$k'(y^*) = x^* \cdot e^{y^*} \left(= \lim_{h \rightarrow 0} \frac{f(x^*, y^* + h) - f(x^*, y^*)}{h} \right).$$

Man nennt $g'(x^*) =: \frac{\partial f}{\partial x}(x^*, y^*)$ und $k'(y^*) =: \frac{\partial f}{\partial y}(x^*, y^*)$ die partielle Ableitung nach x bzw. nach y von f .

Allgemeiner vereinbaren wir:

Definition 3.6. Sei $D \subseteq \mathbb{R}^n$ eine offene Menge, sei $f: D \rightarrow \mathbb{R}^m$ eine Funktion, $k \in \{1, 2, \dots, n\}$, und e_k sei der k -te Einheitsvektor im \mathbb{R}^n .

(a) Sei $x^* \in D$. Wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x^* + h \cdot e_k) - f(x^*)}{h}$$

existiert, so heißt f im Punkt x^* partiell nach x_k differenzierbar. Der Grenzwert heißt partielle Ableitung von f nach x_k in x^* . Schreibweisen:

$$\frac{\partial f}{\partial x_k}(x^*) \quad \text{oder} \quad \partial_k f(x^*) \quad \text{oder} \quad f_{x_k}(x^*).$$

- (b) Sei $x^* \in D$. Ist f in x^* partiell nach x_k differenzierbar für jedes $k = 1, \dots, n$, so heißt f in x^* partiell differenzierbar.
- (c) Ist f für jedes $x^* \in D$ partiell differenzierbar, so heißt f partiell differenzierbar in D .
- (d) Ist f partiell differenzierbar in D und sind die Ableitungen $\partial_1 f, \dots, \partial_n f$ stetig im Punkt $x^* \in D$, so heißt f stetig partiell differenzierbar in x^* .

Bemerkung 3.7. (a) Partielles Differenzieren nach der Variablen x_k bedeutet: alle anderen Variablen werden als feste Parameter betrachtet. Wir leiten wie in HM I nach der einzigen noch beweglichen Variablen x_k ab.

(b) Das bedeutet: alle Ableitungsregeln aus HM I können direkt übernommen werden.

(c) Wenn wir eine Funktion $f: D \rightarrow \mathbb{R}^m$ haben, also

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix},$$

dann gilt

$$\frac{\partial f}{\partial x_k} = \begin{pmatrix} \partial_k f_1 \\ \partial_k f_2 \\ \vdots \\ \partial_k f_m \end{pmatrix},$$

da Limiten komponentenweise gebildet werden. Also betrachten jetzt meist nur noch Funktionen $f: D \rightarrow \mathbb{R}$.

Beispiel 3.8. (a) Sei $f(x, y) = x^3 y - 2x^2 y^2$ für $(x, y) \in \mathbb{R}^2 = D_f$. Dann folgt

$$\frac{\partial f}{\partial x} = 3x^2 y - 4xy^2, \quad \frac{\partial f}{\partial y} = x^3 - 4x^2 y.$$

(b) Sei $g(x, y, z) = xe^{xz+y^2}$ für $(x, y, z) \in \mathbb{R}^3 = D_g$. Dann

$$\begin{aligned} \frac{\partial g}{\partial x}(x, y, z) &= e^{xz+y^2} + xe^{xz+y^2} \cdot z = (1+xz)e^{xz+y^2}, \\ \frac{\partial g}{\partial y}(x, y, z) &= xe^{xz+y^2} \cdot 2y = 2xye^{xz+y^2}, \\ \frac{\partial g}{\partial z}(x, y, z) &= xe^{xz+y^2} \cdot x = x^2 e^{xz+y^2}. \end{aligned}$$

Eine anschauliche Interpretation lautet: $\partial_k f(x_1, \dots, x_k)$ bezeichnet die Änderungsrate von f , wenn die Variable x_k variiert wird und alle anderen Variablen konstant bleiben.

Beispiel 3.9 (van der Waals'sche Zustandsgleichung). Wir haben

$$p = p(t, v) = \frac{R \cdot t}{v - b} - \frac{a}{v^2}.$$

$\partial_v p$: die partielle Ableitung $\frac{\partial p}{\partial v}$ gibt an, mit welcher Rate der Druck p sich ändert, wenn das Volumen v sich ändert, aber die Temperatur t konstant bleibt („isotherme Zustandsänderung“):

$$\frac{\partial p}{\partial v}(t, v) = -\frac{Rt}{(v-b)^2} + \frac{2a}{v^3}.$$

Bei einem idealen Gas ist $a = b = 0$, und wir bekommen $\partial_v p = -\frac{Rt}{v^2} \leq 0$, was plausibel erscheint, denn bei Ausdehnung des Gefäßes sollte der Druck im Gefäß sinken; und wenn wir das Gefäß zusammendrücken, dann sollte der Druck darin steigen.

$\partial_t p$: die partielle Ableitung $\frac{\partial p}{\partial t}$ gibt an, mit welcher Rate der Druck p sich ändert, wenn die Temperatur t sich ändert, aber das Volumen v konstant bleibt („isochore Zustandsänderung“):

$$\frac{\partial p}{\partial t}(t, v) = \frac{R}{v - b}.$$

Bei einem idealen Gas bekommen wir $\partial_t p = \frac{R}{v} > 0$, was auch plausibel erscheint, denn das Heizen eines Gefäßes (wobei man die Ausdehnung des Gefäßes verhindert) steigert den Druck im Gefäß.

Wir kehren zurück zu Beispiel 2.11.

Beispiel 3.10. Wir betrachten die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch

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

Dann haben wir im Ursprung

$$\begin{aligned} \frac{\partial f}{\partial x}(0, 0) &= \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{\frac{h \cdot 0}{h^2} - 0}{h} = \lim_{h \rightarrow 0} 0 = 0, \\ \frac{\partial f}{\partial y}(0, 0) &= \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{\frac{0 \cdot h}{h^2} - 0}{h} = \lim_{h \rightarrow 0} 0 = 0, \end{aligned}$$

also ist f im Ursprung partiell differenzierbar. Aber aus Beispiel 2.11 wissen wir, daß f im Punkt $(0, 0)^\top$ nicht stetig ist! Die partielle Differenzierbarkeit ist also keine stärkere Eigenschaft als die Stetigkeit!

Beachte: die partiellen Ableitungen $\partial_x f$ und $\partial_y f$ sind nicht stetig, also ist f im Punkt $(0, 0)^\top$ nicht stetig partiell differenzierbar. Das sehen wir so: für $(x, y)^\top \neq (0, 0)^\top$ gilt

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^2 + y^2) - xy \cdot 2x}{(x^2 + y^2)^2} = \frac{xy^2 + y^3 - 2x^2y}{(x^2 + y^2)^2} = \frac{y^3 - x^2y}{(x^2 + y^2)^2}.$$

Wir betrachten eine Folge $(x_j, y_j)^\top = (1/j, 1/j^2)^\top$, für $j \in \mathbb{N}$. Dann haben wir die Konvergenz $(x_j, y_j)^\top \rightarrow (0, 0)^\top$ (wobei $j \rightarrow \infty$), und es ist

$$\frac{\partial f}{\partial x}(x_j, y_j) = \frac{(\frac{1}{j^6} - \frac{1}{j^4})j^4}{(\frac{1}{j^2} + \frac{1}{j^4})^2 j^4} = \frac{\frac{1}{j^2} - 1}{(1 + \frac{1}{j^2})^2} \rightarrow \frac{-1}{1} = -1 \neq 0 = \frac{\partial f}{\partial x}(0, 0).$$

Definition 3.11. Sei $D \subseteq \mathbb{R}^n$ eine offene Menge.

(a) Sei $f: D \rightarrow \mathbb{R}$ partiell differenzierbar in $x \in D$. Der Vektor

$$\nabla f(x) := \text{grad } f(x) := \left(\frac{\partial f}{\partial x_1}(x), \quad \frac{\partial f}{\partial x_2}(x), \quad \dots \quad \frac{\partial f}{\partial x_n}(x) \right)$$

heißt Gradient von f in x .

(b) Sei $f: D \rightarrow \mathbb{R}^m$ partiell differenzierbar in $x \in D$. Die Matrix

$$J_f(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

heißt Jacobi-Matrix von f in x .

Bemerkung 3.12. (a) Für $f: D \rightarrow \mathbb{R}^m$ gilt

$$J_f(x) = \begin{pmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \\ \nabla f_m(x) \end{pmatrix}.$$

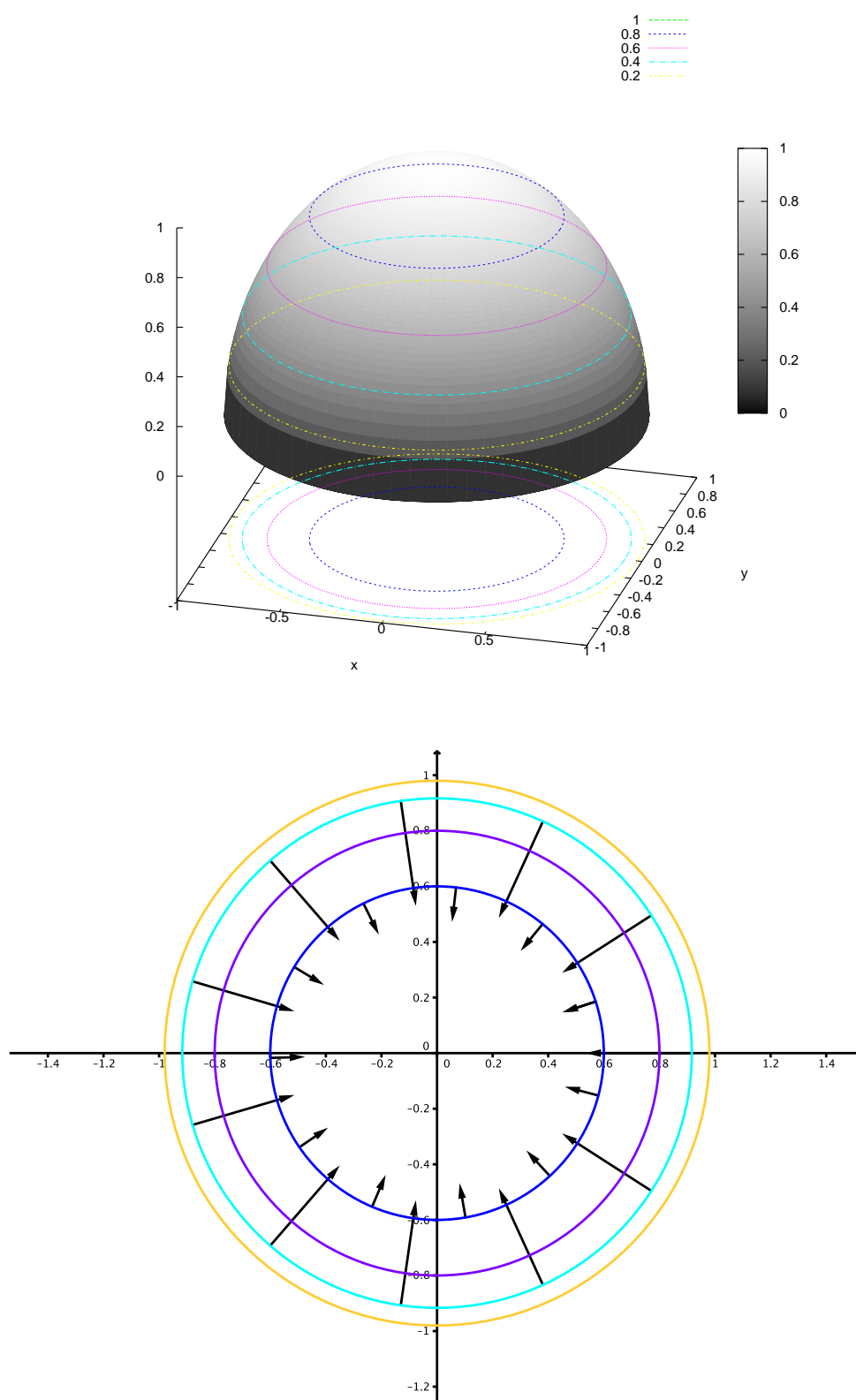


Abbildung 1.7: Oben: die skalare Funktion $f(x, y) = \sqrt{1 - x^2 - y^2}$. Die Höhenlinien sind zweifach eingetragen worden: einmal auf der Halbkugelfläche, und ein weiteres mal im Definitionsbereich $M \subset \mathbb{R}^2$. Unten sieht man die Höhenlinien in M und einige Gradientenvektoren. Die Gradientenvektoren 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.

(b) Wenn $f: D \rightarrow \mathbb{R}^1$ ist, dann ist ∇f eine spezielle $1 \times n$ -Jacobi-Matrix.

(c) Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}$ stetig partiell differenzierbar, und $x \in D$. Dann hat der Gradient eine anschauliche Bedeutung: $\nabla f(x) \in \mathbb{R}^n$ ist ein Vektor, dessen Richtung angibt, in welche Richtung f im Punkt x am stärksten anwächst. Sein Betrag ist diese größte Steigung.

Beispiel 3.13. Wir betrachten die obere Kuppe einer Kugel mit Radius 1. Siehe Bild 1.7. Es ist also $f: K(0,1) \rightarrow \mathbb{R}$, wobei $K(0,1) \subset \mathbb{R}^2$ die Kreisscheibe um den Ursprung mit Radius 1 ist, und f ist gegeben durch $f(x,y) = \sqrt{1-x^2-y^2}$. Dann haben wir

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y) \right) = \left(\frac{-2x}{2\sqrt{1-x^2-y^2}}, \frac{-2y}{2\sqrt{1-x^2-y^2}} \right) = \frac{1}{\sqrt{1-x^2-y^2}}(-x, -y).$$

Definition 3.14. Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}^n$ ein stetig partiell differenzierbares Vektorfeld. Dann heißt

$$\operatorname{div} f(x) := \sum_{k=1}^n \frac{\partial f_k}{\partial x_k}(x) = \frac{\partial f_1}{\partial x_1}(x) + \frac{\partial f_2}{\partial x_2}(x) + \cdots + \frac{\partial f_n}{\partial x_n}(x)$$

Divergenz von f im Punkt x .

Bemerkung 3.15. (a) Wenn man formal liest

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

(„Nabla-Operator“), so ist dann

$$\operatorname{div} f(x) = \langle \nabla, f \rangle(x) = \left\langle \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \right\rangle = \sum_{k=1}^n \frac{\partial f_k}{\partial x_k}.$$

(b) Anschauliche Bedeutung der Divergenz: Sei f ein Fluß von etwas (z.B. ein Materialfluß, ein elektrischer Strom, ein magnetisches Feld, ...). Dann beschreibt $\operatorname{div} f(x)$ die Menge der Größe, die im Punkt x verschwindet bzw. entsteht. Die Gleichung $\operatorname{div} f(x) = 0$ besagt dann, daß das fließende Material erhalten bleibt („alles was irgendwo reinfließt, fließt auch wieder raus“).

Wir kommen jetzt zu höheren Ableitungen.

Definition 3.16. Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}$ sei partiell differenzierbar in D . Wenn $\frac{\partial f}{\partial x_k}$ im Punkt $x \in D$ erneut partiell nach x_l differenzierbar ist, so heißt

$$\frac{\partial^2 f}{\partial x_k \partial x_l} := \frac{\partial}{\partial x_l} \left(\frac{\partial f}{\partial x_k} \right)$$

partielle Ableitung zweiter Ordnung nach x_k und x_l von f . Analog definiert man Ableitungen 3., 4., 5. Ordnung.

Beispiel 3.17. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch $f(x,y) = x^3y + xe^y$. Wir haben dann die partiellen Ableitungen erster Ordnung

$$f_x = \frac{\partial f}{\partial x}(x,y) = 3x^2y + e^y, \quad f_y = \frac{\partial f}{\partial y}(x,y) = x^3 + xe^y.$$

Und wir haben die Ableitungen zweiter Ordnung

$$\begin{aligned} f_{xx} &= \frac{\partial^2 f}{\partial x^2}(x,y) = 6xy, & f_{xy} &= \frac{\partial^2 f}{\partial x \partial y}(x,y) = 3x^2 + e^y, \\ f_{yx} &= \frac{\partial^2 f}{\partial y \partial x}(x,y) = 3x^2 + e^y, & f_{yy} &= \frac{\partial^2 f}{\partial y^2}(x,y) = xe^y. \end{aligned}$$

Wir beobachten $f_{xy} = f_{yx}$. Die partiellen Ableitungen dritter Ordnung sind bereits acht Stück:

$$\begin{aligned} f_{xxx} &= \frac{\partial^3 f}{\partial x^3}(x, y) = 6y, & f_{xxy} &= \frac{\partial^3 f}{\partial x^2 \partial y}(x, y) = 6x, \\ f_{yxx} &= \frac{\partial^3 f}{\partial y \partial x^2}(x, y) = 6x, & f_{yyx} &= \frac{\partial^3 f}{\partial y \partial x \partial y}(x, y) = e^y, \\ f_{xyx} &= \frac{\partial^3 f}{\partial x \partial y \partial x}(x, y) = 6x, & f_{xyy} &= \frac{\partial^3 f}{\partial x \partial y^2}(x, y) = e^y, \\ f_{yyx} &= \frac{\partial^3 f}{\partial y^2 \partial x}(x, y) = e^y, & f_{yyy} &= \frac{\partial^3 f}{\partial y^3}(x, y) = xe^y. \end{aligned}$$

Theorem 3.18 (Satz von Schwarz). Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}^m$ sei p -mal stetig partiell differenzierbar in D , das heißt:

- alle partiellen Ableitungen von f bis zur Ordnung p existieren in D ,
- und alle diese Ableitungen sind stetig in $x \in D$.

Dann ist bei den partiellen Ableitungen p -ter Ordnung die Reihenfolge der Differentiationen vertauschbar.

Warnung: Sind die partiellen Ableitungen nicht stetig, dann ist das im Allgemeinen falsch. Man betrachte z.B. die Funktion

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

Für diese Funktion ist $f_{xy}(0, 0) = -1$, aber $f_{yx}(0, 0) = +1$.

4 Totale Differenzierbarkeit

Wir bringen eine **neue Idee** ins Spiel: das Differenzieren hängt mit der linearen Approximation einer Funktion zusammen.

Wir erinnern an die HM I: eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ ist differenzierbar in einem Punkt $x_0 \in \mathbb{R}$ genau dann, wenn wir (für x in der Nähe von x_0) schreiben können

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + r(x),$$

wobei $r(x)$ ein Rest-Summand ist mit der Eigenschaft $\lim_{x \rightarrow x_0} \frac{|r(x)|}{|x - x_0|} = 0$, siehe Bild 1.8.

Diese Eigenschaft wollen wir jetzt übertragen auf Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Dann ist $f(x) \in \mathbb{R}^m$ und $x - x_0 \in \mathbb{R}^n$.

Definition 4.1. Sei $D \subset \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}^m$, und $x^* \in D$.

Die Funktion f heißt in x^* (total) differenzierbar, wenn eine Matrix $A \in \mathbb{R}^{m \times n}$ existiert, so daß

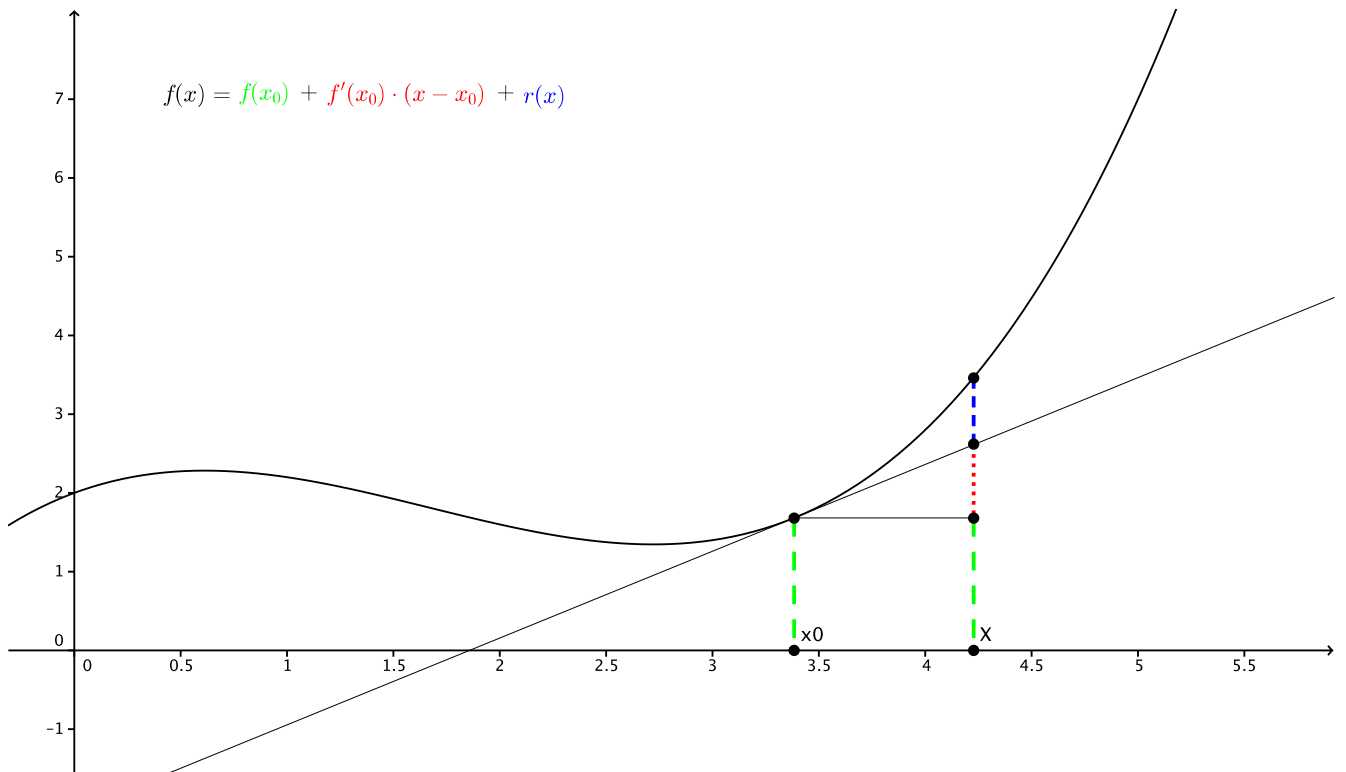
$$f(x) = f(x^*) + A(x - x^*) + r(x)$$

gilt mit $\lim_{x \rightarrow x^*} \frac{|r(x)|}{|x - x^*|} = 0$. Man schreibt dann $Df(x^*) := A$ oder $\frac{df}{dx}(x^*) := A$ oder $f'(x^*) := A$.

Bemerkung 4.2. Es gibt immer höchstens eine solche Matrix A . Obige Definition ist also sinnvoll.

Beispiel 4.3. Sei $f(x, y) = x^2 + y^2$. Wir suchen jetzt die Matrix A . Dazu schreiben wir

$$\begin{aligned} f(x, y) &= f(x^* + x - x^*, y^* + y - y^*) = \left(x^* + (x - x^*)\right)^2 + \left(y^* + (y - y^*)\right)^2 \\ &= (x^*)^2 + (y^*)^2 + 2x^*(x - x^*) + 2y^*(y - y^*) + (x - x^*)^2 + (y - y^*)^2 \\ &= f(x^*, y^*) + \begin{pmatrix} 2x^* & 2y^* \end{pmatrix} \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} + \left| \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right|^2. \end{aligned}$$

Abbildung 1.8: Wir zerlegen den Wert $f(x)$ in drei Teile.

Im letzten Summand erkennen wir $r(x, y)$. Das können wir so handhaben, denn

$$\frac{|r(x, y)|}{\left| \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right|} = \frac{\left| \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right|^2}{\left| \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right|} = \left| \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right| \rightarrow 0, \quad \text{wenn } (x, y)^\top \rightarrow (x^*, y^*)^\top.$$

Dann ist $A = (2x^*, 2y^*)$ eine Matrix mit einer Zeile und zwei Spalten, und die Funktion f ist total differenzierbar für alle $(x^*, y^*)^\top \in \mathbb{R}^2$.

Wir haben $Df(x^*, y^*) = (2x^*, 2y^*)$ und beobachten $Df = \nabla f$.

Das ist kein Zufall, wie wir gleich sehen werden.

Satz 4.4. Ist f total differenzierbar in $x^* \in D$, so ist f stetig in x^* .

Beweis. Wir rechnen $\lim_{x \rightarrow x^*} f(x)$ einfach aus:

$$\begin{aligned} \lim_{x \rightarrow x^*} f(x) &= \lim_{x \rightarrow x^*} \left(f(x^*) + A(x - x^*) + r(x) \right) \\ &= f(x^*) + A \cdot 0 + \lim_{x \rightarrow x^*} r(x) = f(x^*) + 0 = f(x^*), \end{aligned}$$

denn wir haben

$$\lim_{x \rightarrow x^*} |r(x)| = \lim_{x \rightarrow x^*} \frac{|r(x)|}{|x - x^*|} \cdot \lim_{x \rightarrow x^*} |x - x^*| = 0 \cdot 0 = 0.$$

□

Satz 4.5. Ist f total differenzierbar $x^* \in D$, so ist f partiell differenzierbar in $x^* \in D$ mit $Df(x^*) = J_f(x^*)$.

Beweis. Sei $k \in \{1, 2, \dots, n\}$. Dann ist die k -te Spalte von $J_f(x^*)$ gleich

$$\begin{aligned} \frac{\partial f}{\partial x_k}(x^*) &= \lim_{h \rightarrow 0} \frac{f(x^* + he_k) - f(x^*)}{h} && \left| \begin{array}{l} \text{benutze jetzt die Voraussetzung} \\ \\ \text{zweiter Summand ist 0, weil } f \text{ total diff'bar} \end{array} \right. \\ &= \lim_{h \rightarrow 0} \frac{f(x^*) + Df(x^*) \cdot (he_k) + r(x^* + he_k) - f(x^*)}{h} \\ &= \lim_{h \rightarrow 0} \left(\frac{h \cdot Df(x^*)e_k}{h} + \frac{r(x^* + he_k)}{h|e_k|} \right) \\ &= Df(x^*) \cdot e_k + \lim_{h \rightarrow 0} \frac{r(x^* + he_k)}{|x^* + he_k - x^*|} \\ &= Df(x^*) \cdot e_k, \end{aligned}$$

und das ist aber gleich der k -ten Spalte von $Df(x_0)$. \square

Bemerkung 4.6. Die Umkehrung ist falsch nach Satz 4.4 und Beispiel 3.10. Andererseits gilt jedoch:

Satz 4.7. Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}^m$, und $x^* \in D$. Ist f im Punkt x^* **stetig** partiell differenzierbar, so ist f im Punkt x^* total differenzierbar, und es gilt $Df(x^*) = J_f(x^*)$.

Wir haben damit eine **erfreuliche Zusammenfassung** für die Bestimmung der totalen Ableitung einer Funktion f :

- zuerst bestimmen wir die partiellen Ableitungen,
- dann schauen wir, ob diese stetig sind. Wenn ja, dann ist f total differenzierbar, und es ist $Df = J_f$.

Insbesondere erkennen wir:

- es sind keine neuen Differentiationsregeln nötig,
- beide Zugänge führen für schöne Funktionen auf den gleichen Ableitungsbegriff.

Bemerkung 4.8 (Kettenregel). Sei $D \subseteq \mathbb{R}^n$ offen, $g: D \rightarrow \mathbb{R}^m$ stetig differenzierbar, $G \subseteq g(D)$ offen und $f: G \rightarrow \mathbb{R}^d$ stetig differenzierbar. Dann gilt

$$D(f \circ g)(x) = Df(g(x)) \cdot Dg(x) \in \mathbb{R}^{d \times n},$$

wobei der \cdot die Multiplikation einer Matrix aus dem $\mathbb{R}^{d \times m}$ mit einer Matrix aus dem $\mathbb{R}^{m \times n}$ bezeichnet.

Beispiel 4.9. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x, y) = x^3y + xe^y$ (siehe Beispiel 3.17). Wir suchen

$$\frac{d}{dt}(f(t^2, t^3)).$$

Dazu geben wir der inneren Funktion einen Namen g , das heißt $g: \mathbb{R} \rightarrow \mathbb{R}^2$ sei gegeben durch $g(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}$. Dann ist nach Kettenregel

$$\frac{d}{dt}(f(t^2, t^3)) = \frac{d}{dt}(f \circ g)(t) = (Df \circ g)(t) = \nabla f(g(t)) \cdot Dg(t).$$

Aus Beispiel 3.17 wissen wir $\nabla f(x, y) = (3x^2y + e^y, x^3 + xe^y)$, also folgt dann

$$\begin{aligned} \nabla f(g(t)) \cdot Dg(t) &= \left(3(t^2)^2t^3 + e^{t^3}, (t^2)^3 + t^2e^{t^3} \right) \cdot \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix} \\ &= (3t^7 + e^{t^3}) \cdot 2t + (t^6 + t^2e^{t^3}) \cdot 3t^2 = 6t^8 + 2te^{t^3} + 3t^8 + 3t^4e^{t^3} \\ &= 9t^8 + t(2 + 3t^3)e^{t^3}. \end{aligned}$$

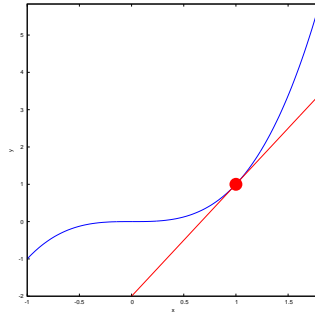


Abbildung 1.9: Die Tangente an den Graph der Funktion $y = f(x)$, genommen im Punkt $(x_0, f(x_0))^T$, hört auf die Gleichung $y = f(x_0) + f'(x_0) \cdot (x - x_0)$. Die Tangente ist diejenige Gerade, die den Graphen der Funktion f in der Nähe von x_0 bestmöglich annähert. Eine Funktion f ist genau dann im Punkt x_0 differenzierbar, wenn es dort eine eindeutig bestimmte Tangente gibt.

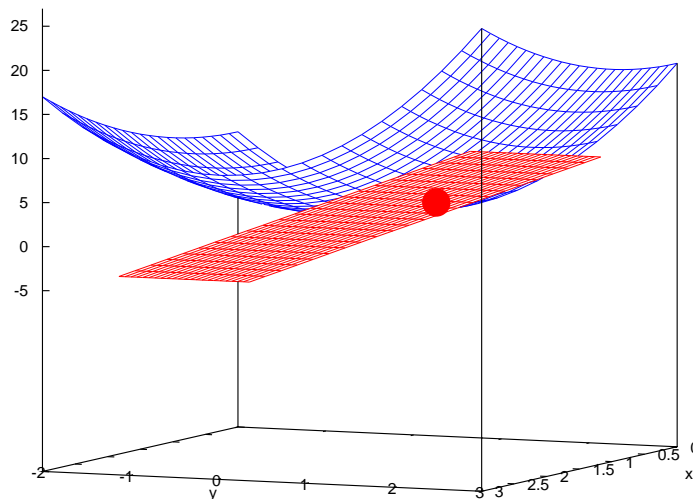


Abbildung 1.10: Die Tangentialebene an den Graph der Funktion $y = f(\vec{x})$, genommen im Punkt $(\vec{x}_0^T, f(\vec{x}_0))^T$, hört auf die Gleichung $y = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$. Die Tangentialebene ist diejenige Ebene, die den Graphen der Funktion f in der 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 !

Kapitel 2

Matrixrechnung, Determinanten und Eigenwerte

5 Erinnerungen

Wir erinnern an die **Matrixmultiplikation**:

Wenn $A \in \mathbb{R}^{d \times m}$ und $B \in \mathbb{R}^{m \times n}$, so ist $A \cdot B \in \mathbb{R}^{d \times n}$. Die Berechnung von $A \cdot B$ erfolgt wie in diesem Beispiel:

$$\begin{pmatrix} 3 & 1 & 0 \\ 2 & -1 & 5 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 + 1 \cdot 2 + 0 \cdot 3 & 3 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 \\ 2 \cdot 1 - 1 \cdot 2 + 5 \cdot 3 & 2 \cdot 1 - 1 \cdot 1 + 5 \cdot 1 \\ 1 \cdot 1 + 0 \cdot 2 + 1 \cdot 3 & 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 16 & 6 \\ 4 & 2 \end{pmatrix}.$$

In Zukunft wird der Multiplikationspunkt \cdot oft weggelassen.

Und wir erinnern an die **inversen Matrizen**: Eine Matrix $A \in \mathbb{R}^{n \times n}$ (also mit quadratischen Abmessungen) heißt invertierbar, falls eine Matrix $A^{-1} \in \mathbb{R}^{n \times n}$ existiert mit $AA^{-1} = I_n$ und $A^{-1}A = I_n$.

Wir berechnen die inverse Matrix über das Gaußverfahren: in der linken Hälfte steht die Matrix A , in der rechten Hälfte beginnen wir mit der Einheitsmatrix, anschließend führen wir Zeilenumformungen durch:

$$\begin{array}{l} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \quad \left| \begin{array}{l} \\ \\ \text{addiere } Z_1 \text{ zu } Z_3 \end{array} \right. \\ \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & 1 & 0 & 1 \end{array} \right) \quad \left| \begin{array}{l} \\ \\ \text{addiere } -2 \cdot Z_2 \text{ zu } Z_1 \text{ und zu } Z_3 \end{array} \right. \\ \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & -2 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right) \quad \left| \begin{array}{l} \\ \\ \text{addiere } -1 \cdot Z_3 \text{ zu } Z_2; \text{ addiere } Z_3 \text{ zu } Z_1 \end{array} \right. \\ \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -4 & 1 \\ 0 & 1 & 0 & -1 & 3 & -1 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right). \end{array}$$

Wenn in der linken Hälfte die Einheitsmatrix verwirklicht wurde, dann steht in der rechten Hälfte die inverse Matrix A^{-1} .

6 Determinanten

Wann ist eine Matrix $A \in \mathbb{R}^{2 \times 2}$ invertierbar? Die Matrix schreiben wir als

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Wenn a und c beide gleich Null wären, dann hätte A den Rang eins, und damit gäbe es keine inverse Matrix zu A . Also können a und c nicht beide gleich Null sein.

Fall 1: sei $a \neq 0$. Dann rechnen wir so:

$$\begin{aligned} & \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \quad \left| \quad Z_2 \text{ mal } a. \text{ Beachte } a \neq 0 \right. \\ \rightsquigarrow & \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ ca & da & 0 & a \end{array} \right) \quad \left| \quad c \cdot Z_1 \text{ von } Z_2 \text{ abziehen} \right. \\ \rightsquigarrow & \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad-bc & -c & a \end{array} \right). \end{aligned}$$

Wenn $ad - bc = 0$ sein sollte, dann bräche der Invertierungsversuch an dieser Stelle zusammen, weil wir in der linken Hälfte an der Stelle rechts unten keine Eins bauen können, ohne die Null links unten wieder zu zerstören. Sei jetzt also $ad - bc \neq 0$ angenommen, und wir setzen die Rechnung fort:

$$\begin{aligned} & \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad-bc & -c & a \end{array} \right) \quad \left| \quad \frac{b}{ad-bc} \cdot Z_2 \text{ von } Z_1 \text{ abziehen} \right. \\ \rightsquigarrow & \left(\begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{ad-bc} & -\frac{ab}{ad-bc} \\ 0 & ad-bc & -c & a \end{array} \right) \quad \left| \quad \text{vereinfache} \right. \\ \rightsquigarrow & \left(\begin{array}{cc|cc} a & 0 & \frac{ad}{ad-bc} & -\frac{ab}{ad-bc} \\ 0 & ad-bc & -c & a \end{array} \right) \quad \left| \quad Z_1 \text{ durch } a, Z_2 \text{ durch } ad-bc \right. \\ \rightsquigarrow & \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right). \end{aligned}$$

In der rechten Hälfte steht jetzt die inverse Matrix.

Fall 2: sei $c \neq 0$. Dann haben wir eine ähnliche Rechnung und das gleiche Ergebnis.

Wir haben also folgenden Satz bewiesen:

Satz 6.1. Sei eine Matrix $A \in \mathbb{R}^{2 \times 2}$ gegeben,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dann ist diese Matrix invertierbar genau dann, wenn $ad - bc \neq 0$ ist; und in diesem Fall lautet die inverse Matrix

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

Definition 6.2. (a) Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, so heißt die Zahl $\det(A) := ad - bc$ Determinante von A , geschrieben als

$$\det(A) \quad \text{oder} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

(b) Sei $A \in \mathbb{R}^{n \times n}$. Für $j, k = 1, \dots, n$ entsteht $A_{jk} \in \mathbb{R}^{(n-1) \times (n-1)}$ aus A durch Streichen der j -ten Zeile und der k -ten Spalte. Zum Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}, \quad A_{13} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}.$$

(c) Ist $A \in \mathbb{R}^{n \times n}$ und $n \geq 3$, so definieren wir rekursiv

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

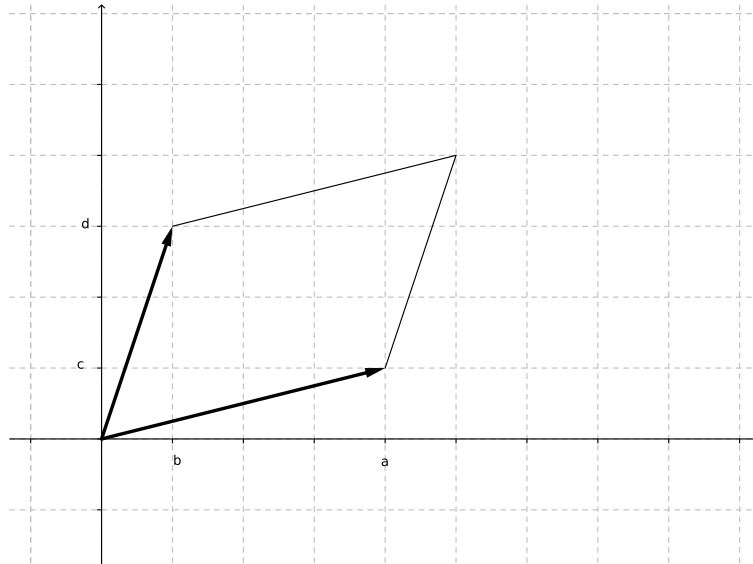


Abbildung 2.1: Die beiden Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ leben in der Ebene und spannen dort ein Parallelogramm auf. Dessen Flächeninhalt ist der Betrag der Determinante $\left| \begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right|$. Wenn diese Determinante Null ist, dann ist das Parallelogramm „platt“, und die beiden Vektoren sind linear abhängig. Siehe Satz 6.4.

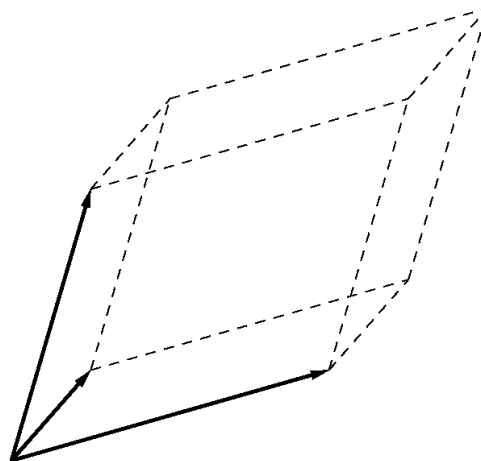


Abbildung 2.2: Die drei Vektoren (interpretiert als Spalten mit je drei Einträgen) leben im dreidimensionalen Raum und spannen dort einen Spat auf. Dessen Volumen ist der Betrag der Determinante zu derjenigen 3×3 -Matrix, die durch Nebeneinanderstellen der drei Spaltenvektoren entsteht. Wenn diese Determinante Null ist, dann ist der Spat „platt“, die drei Vektoren sind in einer Ebene, sind also linear abhängig. Siehe Satz 6.4.

Beispiel 6.3. *Wir betrachten*

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

Dann bekommen wir

$$\begin{aligned} \det A &= (-1)^2 \cdot 3 \cdot \begin{vmatrix} 4 & 2 \\ -1 & 3 \end{vmatrix} + (-1)^3 \cdot 0 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + (-1)^4 \cdot (-1) \cdot \begin{vmatrix} 1 & 4 \\ 2 & -1 \end{vmatrix} \\ &= 3 \cdot (12 + 2) - 0 - (-1 - 8) = 42 - (-9) = 51. \end{aligned}$$

Dieses Rechenverfahren heißt „Entwickeln nach der ersten Zeile“.

Satz 6.4. *Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist invertierbar genau dann, wenn $\det(A) \neq 0$.*

Satz 6.5 (Rechenregeln). Seien $A, B \in \mathbb{R}^{n \times n}$ und $\lambda \in \mathbb{R}$. Dann gilt:

(a) $\det(AB) = \det(A) \cdot \det(B)$,

(b) $\det(\lambda A) = \lambda^n \det(A)$,

(c) wenn A invertierbar ist, so gilt $\det(A^{-1}) = \frac{1}{\det(A)}$,

(d) $\det(A^\top) = \det(A)$,

(e) wenn A eine untere Dreiecksmatrix ist, das heißt

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix},$$

so gilt $\det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdot \dots \cdot a_{nn}$. Analog für obere Dreiecksmatrizen.

Beweis. (e): Wir rechnen die Determinante aus durch wiederholtes Entwickeln nach der ersten Zeile:

$$\begin{aligned} \det(A) &= (-1)^2 a_{11} \begin{vmatrix} a_{22} & 0 & 0 & \dots & 0 \\ a_{32} & a_{33} & 0 & \dots & 0 \\ a_{42} & a_{43} & a_{44} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & a_{n4} & \dots & a_{nn} \end{vmatrix} \\ &= a_{11} \cdot (-1)^2 a_{22} \begin{vmatrix} a_{33} & 0 & \dots & 0 \\ a_{43} & a_{44} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n3} & a_{n4} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} \cdot \dots \cdot a_{nn}. \end{aligned}$$

Die Aussage über obere Dreiecksmatrizen folgt aus (d).

(c): Nach (e) gilt $\det(I_n) = 1$, also ist

$$1 = \det(I_n) = \det(A \cdot A^{-1}) \stackrel{(a)}{=} \det(A) \cdot \det(A^{-1}),$$

woraus dann die Behauptung $\det(A^{-1}) = \frac{1}{\det(A)}$ folgt.

□

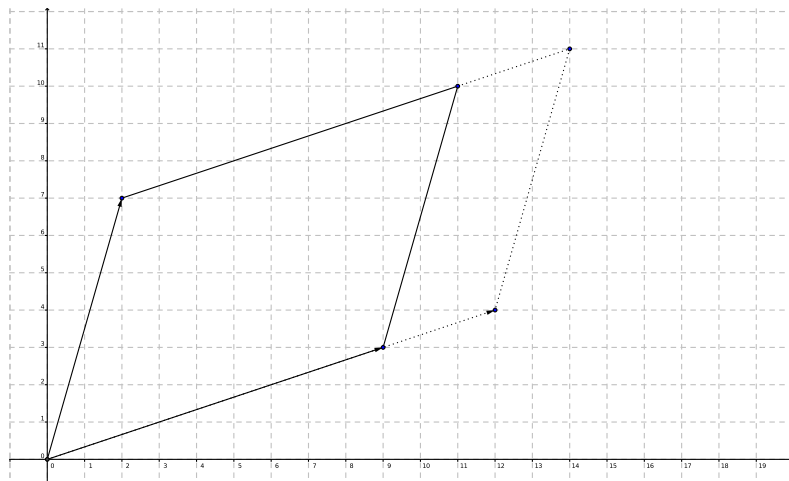


Abbildung 2.3: Es ist $\det(\frac{1}{3}\vec{a}, \vec{b}) = \frac{1}{4} \det(\vec{a}, \vec{b})$. Siehe Bild 2.1 und Satz 6.6 (b) sowie Bemerkung 6.7 (b).

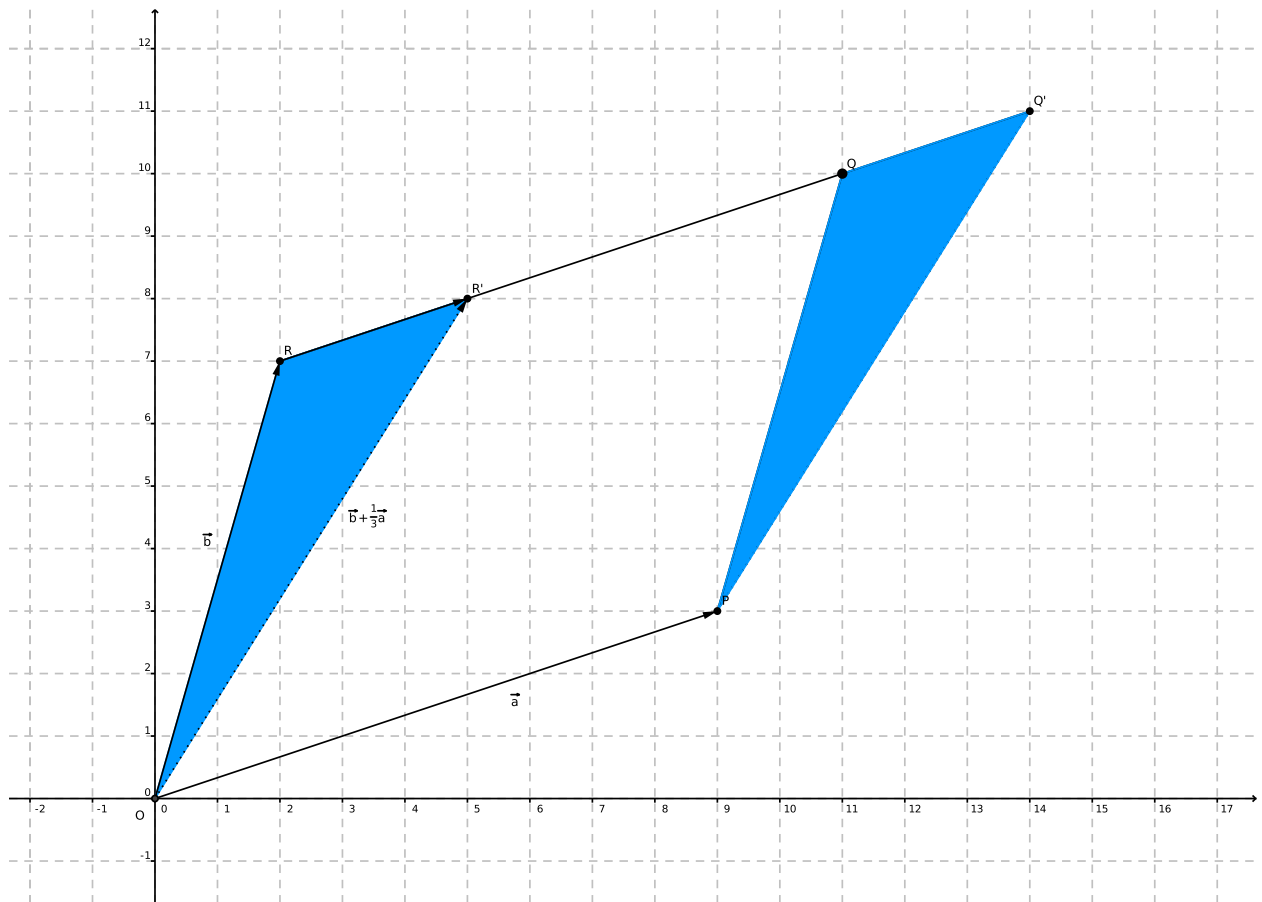


Abbildung 2.4: Die zweikomponentigen Spaltenvektoren \vec{a} und \vec{b} können zu einer 2×2 -Matrix nebeneinandergestellt werden, und deren Determinante $\det(\vec{a}, \vec{b})$ ist der orientierte Flächeninhalt des Parallelogramms $OPQR$. Die blauen Dreiecke sind kongruent, also haben die Parallelogramme $OPQR$ und $OPQR'$ denselben Flächeninhalt, und deshalb ist $\det(\vec{a}, \vec{b}) = \det(\vec{a}, \vec{b} + \frac{1}{3}\vec{a})$. Siehe Satz 6.6 (c) sowie Bemerkung 6.7 (b).

Satz 6.6 (Determinanten und Elementarumformungen). Sei $A \in \mathbb{R}^{n \times n}$ und a_k die k -te Zeile von A , für $k = 1, \dots, n$. Dann gilt für alle $\lambda \in \mathbb{R}$ und alle $j, k = 1, \dots, n$:

- (a) Das Vertauschen zweier Zeilen multipliziert die Determinante mit -1 .
 (b) Das Multiplizieren einer Zeile mit λ multipliziert die Determinante mit λ :

$$\begin{vmatrix} a_1 \\ \vdots \\ a_{j-1} \\ \lambda a_j \\ a_{j+1} \\ \vdots \\ a_n \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_1 \\ \vdots \\ a_{j-1} \\ a_j \\ a_{j+1} \\ \vdots \\ a_n \end{vmatrix} = \lambda \cdot \det(A).$$

- (c) Das Addieren des λ -Fachen der k -ten Zeile auf die j -Zeile ändert den Wert der Determinante nicht, wenn $k \neq j$:

$$\begin{vmatrix} a_1 \\ \vdots \\ a_{j-1} \\ a_j + \lambda a_k \\ a_{j+1} \\ \vdots \\ a_n \end{vmatrix} = \begin{vmatrix} a_1 \\ \vdots \\ a_{j-1} \\ a_j \\ a_{j+1} \\ \vdots \\ a_n \end{vmatrix} = \det(A).$$

Bemerkung 6.7. (a) Zu Teil (b) vergleiche man Satz 6.5, Teil (b).

(b) Wegen Teil (d) aus Satz 6.5 gilt Satz 6.6 genauso auch für Spaltenumformungen.

(c) Wegen Satz 6.6 gilt allgemeiner als in Definition 6.2, Teil (c):

$$\det(A) = \sum_{k=1}^n (-1)^{k+j} a_{jk} \det(A_{jk}), \quad \forall j = 1, \dots, n \quad (\text{Entwickeln nach der } j\text{-ten Zeile}),$$

$$\det(A) = \sum_{j=1}^n (-1)^{k+j} a_{jk} \det(A_{jk}), \quad \forall k = 1, \dots, n \quad (\text{Entwickeln nach der } k\text{-ten Spalte}).$$

Beispiel 6.8.

$$\begin{vmatrix} 0 & 2 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ -1 & 3 & -4 & 3 & 4 \\ -1 & 2 & 3 & 0 & 0 \\ 1 & 0 & 1 & 1 & -1 \end{vmatrix} \stackrel{\text{Entw. } Z_2}{=} (-1)^6 \cdot (-1) \begin{vmatrix} 0 & 2 & -1 & 1 \\ -1 & 3 & -4 & 4 \\ -1 & 2 & 3 & 0 \\ 1 & 0 & 1 & -1 \end{vmatrix} \quad \left| \begin{array}{l} S_4 \text{ auf } S_3, S_2 \end{array} \right.$$

$$= \begin{vmatrix} 1 & 2 & 0 & 1 \\ 3 & 3 & 0 & 4 \\ -1 & 2 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \stackrel{\text{Entw. } Z_4}{=} (-1)^8 \cdot (-1) \cdot \begin{vmatrix} 1 & 2 & 0 \\ 3 & 3 & 0 \\ -1 & 2 & 3 \end{vmatrix} \stackrel{\text{Entw. } S_3}{=} (-1)^6 \cdot 3 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 3 \end{vmatrix}$$

$$= -3 \cdot (3 - 6) = 9.$$

Die nächste Determinante bestimmen wir mittels zweier Zeilenvertauschungen, um Dreiecksform zu erzielen:

$$\begin{vmatrix} 0 & -1 & 3 & 1 \\ 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{vmatrix} = (-1) \cdot (-1) \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 4 \end{vmatrix} = 1 \cdot (-1) \cdot 1 \cdot 4 = -4.$$

Und noch ein Beispiel

$$\begin{vmatrix} 3 & 1 & -2 & 0 \\ 4 & 2 & 7 & 2 \\ -3 & 2 & 1 & -1 \\ 4 & 5 & 6 & 1 \end{vmatrix} \quad \left| \begin{array}{l} Z_4 \text{ auf } Z_3; (-2) \cdot Z_4 \text{ auf } Z_2 \end{array} \right.$$

$$= \begin{vmatrix} 3 & 1 & -2 & 0 \\ -4 & -8 & -5 & 0 \\ 1 & 7 & 7 & 0 \\ 4 & 5 & 6 & 1 \end{vmatrix} \quad \left| \begin{array}{l} 4 \cdot Z_3 \text{ auf } Z_2; (-3) \cdot Z_3 \text{ auf } Z_1 \end{array} \right.$$

$$= \begin{vmatrix} 0 & -20 & -23 & 0 \\ 0 & 20 & 23 & 0 \\ 1 & 7 & 7 & 0 \\ 4 & 5 & 6 & 1 \end{vmatrix} \quad \left| \begin{array}{l} Z_2 \text{ auf } Z_1 \end{array} \right.$$

$$= \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 20 & 23 & 0 \\ 1 & 7 & 7 & 0 \\ 4 & 5 & 6 & 1 \end{vmatrix} = 0,$$

da offensichtlich der Rang dieser Matrix kleiner ist als 4. Also kann die Determinante nur gleich Null sein.

Bemerkung 6.9. Für $A \in \mathbb{R}^{3 \times 3}$ gilt die Regel von Sarrus:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - afh - bdi.$$

Diese Regel gibt es aber nur für 3×3 -Matrizen !

7 Eigenwerttheorie

Erinnerung: zu jeder Matrix $A \in \mathbb{R}^{n \times n}$ gehört eine lineare Abbildung ϕ_A ,

$$\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \phi_A(x) := Ax.$$

Beispiel 7.1. Wir betrachten ausgezeichnete Richtungen von linearen Abbildungen: Spiegelung an der Diagonalen des ersten Quadranten im \mathbb{R}^2 ; Drehung im \mathbb{R}^3 ; Schrägprojektion auf eine Leinwand.

Siehe Bild 2.5, 2.6, 2.7.

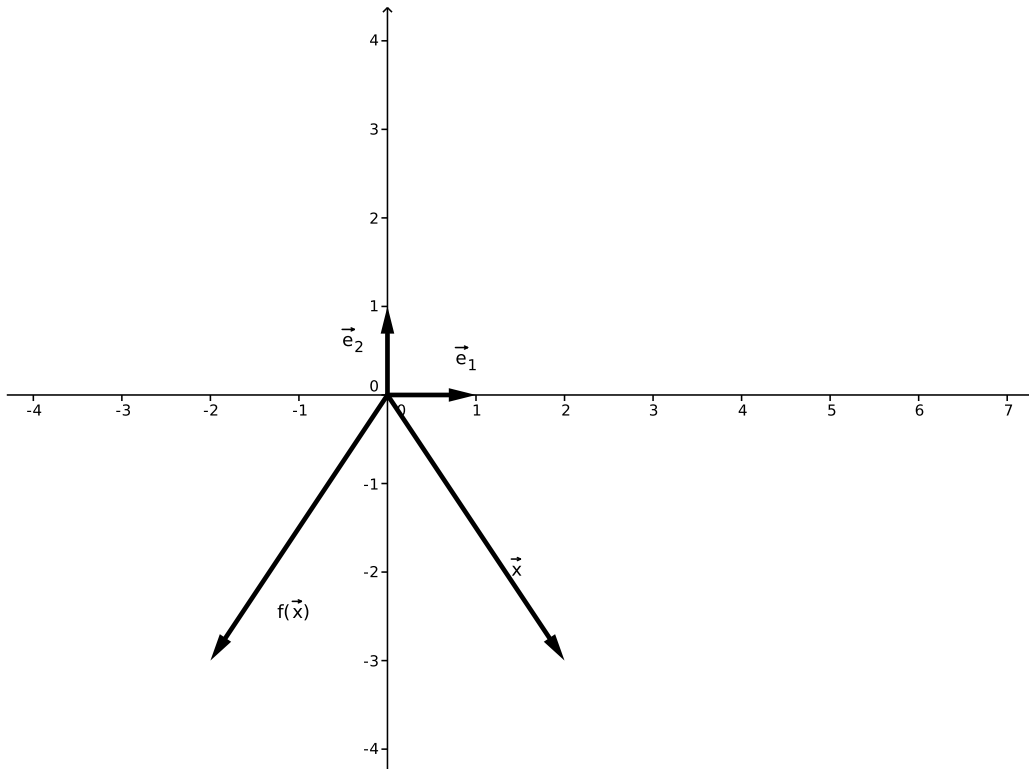


Abbildung 2.5: Die Spiegelung an der y -Achse.

Definition 7.2. Sei $A \in \mathbb{R}^{n \times n}$. Eine Zahl $\lambda \in \mathbb{R}$ heißt Eigenwert von A , wenn es einen Vektor $x \in \mathbb{R}^n$ mit $x \neq 0$ gibt, sodaß $Ax = \lambda x$ gilt. Dieser Vektor x heißt dann Eigenvektor von A .

Bemerkung 7.3. (a) Seien x und y Eigenvektoren von A zum selben Eigenwert λ . Dann gilt:

- $x + y$ ist auch ein Eigenvektor von A zum Eigenwert λ (wenn $x + y \neq 0$),
- αx ist auch ein Eigenvektor von A zum Eigenwert λ , wenn $\alpha \in \mathbb{R}$ und $\alpha \neq 0$.

Denn wir haben $A(x + y) = Ax + Ay = \lambda x + \lambda y = \lambda(x + y)$ und auch $A(\alpha x) = \alpha Ax = \alpha \lambda x = \lambda(\alpha x)$.

(b) Sei $A \in \mathbb{R}^{n \times n}$ und λ ein Eigenwert von A . Dann bedeutet Teil (a), daß

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

ein Untervektorraum des \mathbb{R}^n ist.

Definition 7.4. Die Menge U_λ heißt Eigenraum von A zum Eigenwert λ .

Beispiel 7.5 (Spiegelung an Winkelhalbierender). Wir haben $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, denn die Spalten der Abbildungsmatrix sind die Koordinaten der Bilder der Basisvektoren. Die Zeichnung in Bild 2.7 suggeriert einige Eigenwerte und Eigenvektoren:

$$\lambda_1 = +1 \quad \text{mit Eigenvektor} \quad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\lambda_2 = -1 \quad \text{mit Eigenvektor} \quad y = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

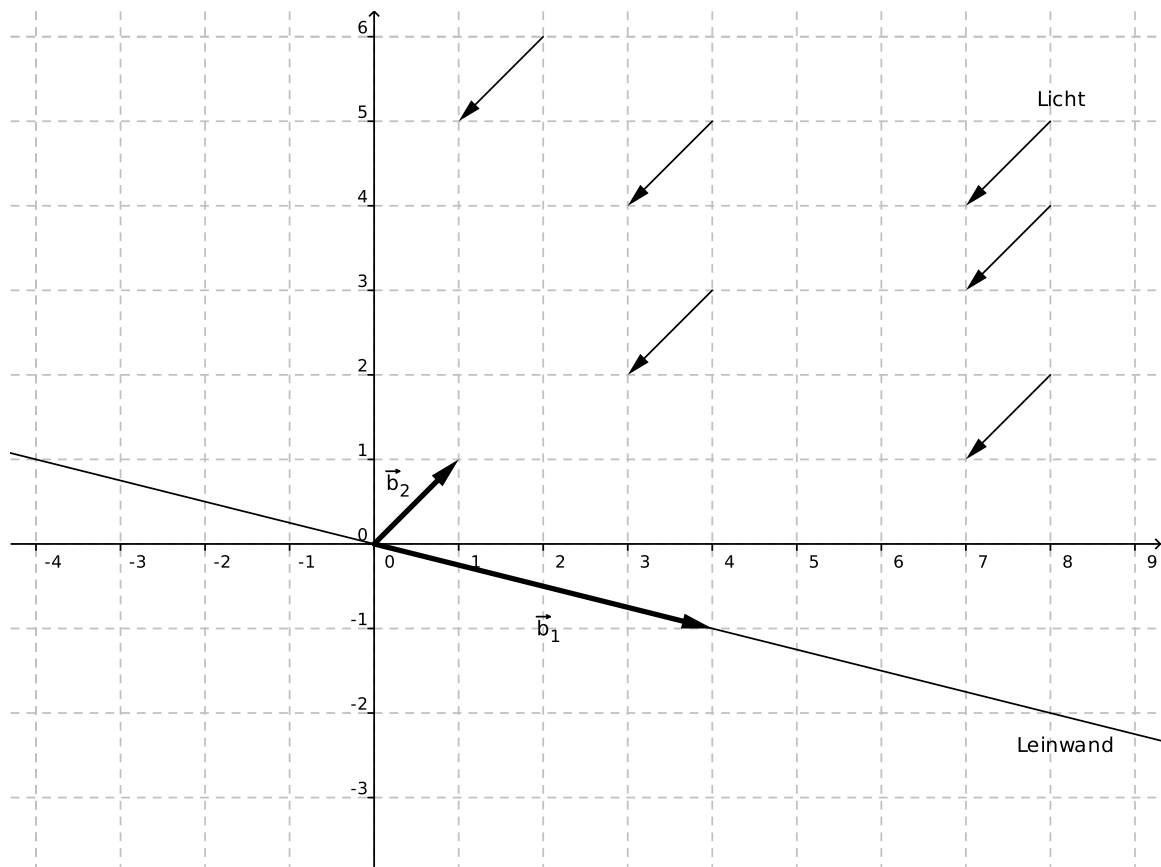


Abbildung 2.6: Licht fällt auf eine Leinwand.

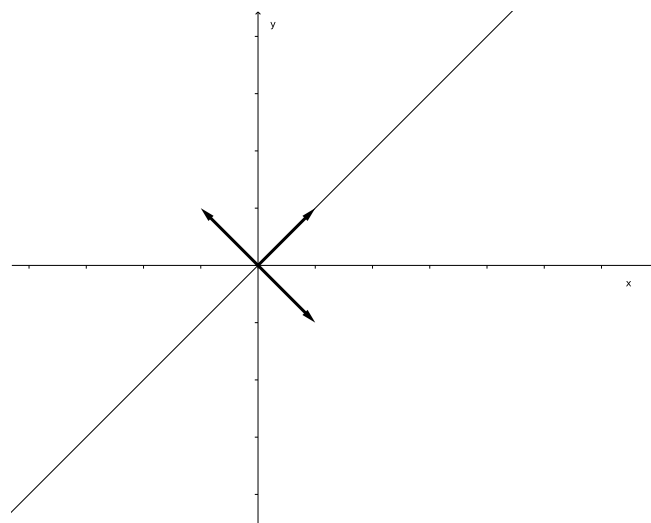


Abbildung 2.7: Die Spiegelung an der Winkelhalbierenden.

Tatsächlich ist

$$Ax = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot x,$$

$$Ay = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \cdot y.$$

Für die **Bestimmung von Eigenwerten und Eigenvektoren** überlegen wir uns: Sei $A \in \mathbb{R}^{n \times n}$. Dann gilt:

λ ist Eigenwert von A

\Leftrightarrow Das lineare Gleichungssystem $Ax = \lambda x$ hat eine Lösung $x \neq 0$.

\Leftrightarrow Das lineare Gleichungssystem $Ax = (\lambda I_n) \cdot x$ hat eine Lösung $x \neq 0$.

\Leftrightarrow Das lineare Gleichungssystem $(A - \lambda I_n)x = 0$ hat eine Lösung $x \neq 0$.

\Leftrightarrow Die Matrix $A - \lambda I_n$ ist nicht invertierbar (vgl. HM I).

\Leftrightarrow Es ist $\det(A - \lambda I_n) = 0$.

Wenn wir $\det(A - \lambda I_n)$ entwickeln, dann erhalten wir

$$\det(A - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - \lambda \end{vmatrix} = (\text{Rechnungen}) \\ = c_0 + c_1 \lambda + c_2 \lambda^2 + \dots + c_{n-1} \lambda^{n-1} + (-1)^n \lambda^n$$

mit gewissen Koeffizienten $c_0, c_1, \dots, c_{n-1} \in \mathbb{R}$.

Satz 7.6. Sei $A \in \mathbb{R}^{n \times n}$. Dann sind die Eigenwerte von A genau die (womöglich komplexen) Nullstellen des Polynoms $p(\lambda) := \det(A - \lambda I_n)$.

Das Polynom p heißt *charakteristisches Polynom* von A .

Beispiel 7.7. (a) Wir suchen die Eigenwerte und Eigenvektoren von

$$A = \begin{pmatrix} -1 & -3 & -3 \\ -2 & -1 & -2 \\ 2 & 3 & 4 \end{pmatrix}.$$

Dazu stellen wir das charakteristische Polynom auf:

$$\det(A - \lambda I_3) = \begin{vmatrix} -1 - \lambda & -3 & -3 \\ -2 & -1 - \lambda & -2 \\ 2 & 3 & 4 - \lambda \end{vmatrix} = \begin{vmatrix} -1 - \lambda & -3 & -3 \\ 0 & 2 - \lambda & 2 - \lambda \\ 2 & 3 & 4 - \lambda \end{vmatrix} \\ = (2 - \lambda) \begin{vmatrix} -1 - \lambda & -3 & -3 \\ 0 & 1 & 1 \\ 2 & 3 & 4 - \lambda \end{vmatrix} = (2 - \lambda) \begin{vmatrix} -1 - \lambda & -3 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 1 - \lambda \end{vmatrix} \\ = (2 - \lambda)(1 - \lambda) \begin{vmatrix} -1 - \lambda & -3 \\ 0 & 1 \end{vmatrix} = (2 - \lambda)(1 - \lambda)(-1 - \lambda).$$

Die Eigenwerte sind also $\lambda_1 = 2$, $\lambda_2 = 1$ und $\lambda_3 = -1$.

Wir ermitteln jetzt die Eigenvektoren zu $\lambda_1 = 2$, indem wir das LGS $(A - 2I_3)x = 0$ lösen. Hierfür rechnen wir wie folgt:

$$\begin{pmatrix} -3 & -3 & -3 & \left| & 0 \right. \\ -2 & -3 & -2 & \left| & 0 \right. \\ 2 & 3 & 2 & \left| & 0 \right. \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 0 & -1 & \left| & 0 \right. \\ -2 & -3 & -2 & \left| & 0 \right. \\ 0 & 0 & 0 & \left| & 0 \right. \end{pmatrix} \\ \rightsquigarrow \begin{pmatrix} -1 & 0 & -1 & \left| & 0 \right. \\ 0 & -3 & 0 & \left| & 0 \right. \\ 0 & 0 & 0 & \left| & 0 \right. \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & \left| & 0 \right. \\ 0 & 1 & 0 & \left| & 0 \right. \\ 0 & 0 & 0 & \left| & 0 \right. \end{pmatrix},$$

und die Lösungen zu diesem System sind $x = \alpha(1, 0, -1)^\top$ für $\alpha \in \mathbb{R}$. Alle Eigenvektoren zu $\lambda_1 = 2$ sind also gegeben durch

$$\alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \alpha \in \mathbb{R} \setminus \{0\}.$$

Entsprechend finden wir die Eigenvektoren zu $\lambda_2 = 1$ und $\lambda_3 = -1$ als $\alpha(0, 1, -1)^\top$ bzw. $\alpha(1, 1, -1)^\top$, wobei jeweils $\alpha \in \mathbb{R} \setminus \{0\}$.

(b) Ganz reelle Matrizen können komplexe Eigenwerte haben! Wir betrachten z.B. die Matrix

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

Dann haben wir

$$\det(A - \lambda I_2) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1,$$

also bekommen wir die Eigenwerte $\lambda_1 = i$ und $\lambda_2 = -i$. Um die Eigenvektoren zu $\lambda_1 = i$ zu finden, lösen wir das LGS $(A - iI_2)x = 0$:

$$\left(\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} -i & -1 & 0 \\ 0 & 0 & 0 \end{array} \right),$$

was die Lösungen $\alpha \begin{pmatrix} 1 \\ -i \end{pmatrix}$ mit $\alpha \in \mathbb{C}$ hat. Entsprechend finden wir die Eigenvektoren zu $\lambda_2 = -i$:

$$\left(\begin{array}{cc|c} i & -1 & 0 \\ 1 & i & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} i & -1 & 0 \\ 0 & 0 & 0 \end{array} \right),$$

mit den Lösungen $\alpha \begin{pmatrix} 1 \\ i \end{pmatrix}$, wobei $\alpha \in \mathbb{C}$.

Bemerkung 7.8. Ist $\alpha \in \mathbb{R}^{n \times n}$ eine obere (oder untere) Dreiecksmatrix, so sind die Diagonaleinträge gerade die Eigenwerte, denn

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} - \lambda \end{vmatrix} = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda).$$

Satz 7.9 (lineare Unabhängigkeit von Eigenvektoren). Sei $A \in \mathbb{R}^{n \times n}$ und seien $\lambda_1, \dots, \lambda_n$ paarweise verschiedene Eigenwerte von A mit Eigenvektoren x_1, \dots, x_n . Dann sind die Vektoren x_1, \dots, x_n linear unabhängig.

Definition 7.10. Sei $A = (a_{jk})_{j,k=1,\dots,n} \in \mathbb{R}^{n \times n}$. Dann heißt

$$\text{spur}(A) := \sum_{k=1}^n a_{kk} = a_{11} + a_{22} + \dots + a_{nn}$$

die Spur von A .

Beispiel 7.11. Es ist (vgl. Beispiel 7.7)

$$\text{spur} \begin{pmatrix} -1 & -3 & -3 \\ -2 & -1 & -2 \\ 2 & 3 & 4 \end{pmatrix} = (-1) + (-1) + 4 = 2.$$

Satz 7.12. Sei $A \in \mathbb{R}^{n \times n}$ und seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A . Dann gilt

$$(a) \text{ spur}(A) = \sum_{j=1}^n \lambda_j = \lambda_1 + \lambda_2 + \dots + \lambda_n.$$

$$(b) \det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n.$$

Beispiel 7.13. Für A aus Beispiel 7.7 hatten wir $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = -1$. Tatsächlich ist

$$2 = \text{spur}(A) = 2 + 1 + (-1) = \lambda_1 + \lambda_2 + \lambda_3,$$

$$\det(A) = \begin{vmatrix} -1 & -3 & -3 \\ -2 & -1 & -2 \\ 2 & 3 & 4 \end{vmatrix} = 4 + 12 + 18 - 6 - 24 - 6 = -2 = 2 \cdot 1 \cdot (-1) = \lambda_1 \cdot \lambda_2 \cdot \lambda_3.$$

Beispiel 7.14. Wir suchen (mit kurzem Rechenweg) die Eigenwerte der Matrix

$$A = \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}.$$

Es ist $\text{spur}(A) = 1$ und $\det(A) = -12 + 10 = -2$, also $\lambda_1 + \lambda_2 = 1$ und $\lambda_1 \lambda_2 = -2$. Wenn wir dies jetzt auflösen würden, dann kämen wir auf eine quadratische Gleichung, was keine Beschleunigung der Rechnung ergäbe. Stattdessen raten wir mit einem scharfen Blick und finden $\lambda_1 = -1$ und $\lambda_2 = 2$.

Definition 7.15. Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt symmetrisch genau dann, wenn $A^\top = A$.

Beispiel 7.16. Folgende Matrix ist symmetrisch:

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

Wir vermerken, daß die Eigenwerte einer symmetrischen Matrix immer reell sind.

Definition 7.17. Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ heißt

- (a) positiv definit (bzw. negativ definit), falls alle Eigenwerte von A echt größer als Null (bzw. echt kleiner als Null) sind.
- (b) positiv semidefinit (bzw. negativ semidefinit), falls alle Eigenwerte von A größer oder gleich Null (bzw. kleiner oder gleich Null) sind.
- (c) indefinit, wenn keiner der beiden vorigen Punkte zutrifft.

Satz 7.18. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Die folgenden Bedingungen sind äquivalent:

- (a) A ist positiv definit (bzw. negativ definit),
- (b) $\langle Ax, x \rangle > 0$ (bzw. $\langle Ax, x \rangle < 0$) für jeden Vektor $x \in \mathbb{R}^n \setminus \{0\}$,
- (c)

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix} > 0 \quad \forall k = 1, \dots, n \quad (\text{positive Untermioren})$$

bzw. wenn

$$(-1)^k \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix} > 0 \quad \forall k = 1, \dots, n \quad (\text{abwechselnde Vorzeichen der Untermioren}).$$

Beispiel 7.19. Wir betrachten die symmetrischen Matrizen

$$A = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 30 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} -2 & 3 & 0 \\ 3 & -5 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Dann sind die Untermioren von A :

$$|1| = 1 > 0, \quad \begin{vmatrix} 1 & -2 \\ -2 & 5 \end{vmatrix} = 5 - 4 = 1 > 0, \quad \begin{vmatrix} 1 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 30 \end{vmatrix} = 150 - 20 - 120 = 10 > 0,$$

also ist A positiv definit.

Die Untermioren von B lauten

$$|-2| = -2 < 0, \quad \begin{vmatrix} -2 & 3 \\ 3 & -5 \end{vmatrix} = 10 - 9 = 1 > 0, \quad \begin{vmatrix} -2 & 3 & 0 \\ 3 & -5 & 0 \\ 0 & 0 & -1 \end{vmatrix} = (-1) \begin{vmatrix} -2 & 3 \\ 3 & -5 \end{vmatrix} = -1 > 0,$$

also ist B negativ definit.

Die Matrix aus Beispiel 7.7 ist indefinit, da sie die Eigenwerte $+2$, $+1$, -1 besitzt.

Kapitel 3

Extremwertprobleme in mehreren Variablen

8 Notwendiges Kriterium und hinreichendes Kriterium

Aus der HM I wissen wir: sei $f: \mathbb{R} \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann gilt:

- wenn f ein lokales Extremum in x_0 hat, so ist $f'(x_0) = 0$
- wenn $f'(x_0) = 0$ ist und $f''(x_0) \neq 0$ ist, so besitzt f in x_0 ein lokales Extremum. Dieses ist ein Maximum falls $f''(x_0) < 0$ ist, und es ist ein Minimum, falls $f''(x_0) > 0$ ist.

Jetzt betrachten wir eine schön differenzierbare Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ und fragen nach deren Extremstellen.

Definition 8.1. Sei $U \subseteq \mathbb{R}^n$ und $f: U \rightarrow \mathbb{R}$. Dann hat f in $x_0 \in U$ ein

globales Maximum (bzw. globales Minimum), wenn $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) für jedes $x \in U$ gilt,

lokales Maximum (bzw. lokales Minimum), wenn es ein $\delta > 0$ gibt mit $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) für jedes $x \in U \cap K(x_0, \delta)$,

lokales (bzw. globales) Extremum, falls f dort ein lokales (bzw. globales) Minimum oder Maximum hat.

Satz 8.2. Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ stetig partiell differenzierbar. Dann gilt:

$$x_0 \in U \text{ ist lokale Extremstelle von } f \implies \nabla f(x_0) = 0.$$

Bemerkung 8.3. Wie in HM I gilt auch hier: die Umkehrung ist falsch.

Definition 8.4. Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar. Die Matrix

$$H_f(x) = \left(\frac{\partial^2}{\partial x_j \partial x_k} f(x) \right)_{j,k=1}^n = \begin{pmatrix} f_{x_1 x_1}(x) & f_{x_1 x_2}(x) & \cdots & f_{x_1 x_n}(x) \\ f_{x_2 x_1}(x) & f_{x_2 x_2}(x) & \cdots & f_{x_2 x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(x) & f_{x_n x_2}(x) & \cdots & f_{x_n x_n}(x) \end{pmatrix}$$

heißt Hesse-Matrix von f .

Dies ist die Matrix von allen zweiten partiellen Ableitungen.

Bemerkung 8.5. (a) Beachte: nach dem Satz von Schwarz (Theorem 3.18) ist $H_f(x)$ immer eine symmetrische Matrix aus $\mathbb{R}^{n \times n}$.

(b) Es ist H_f die Jacobi-Matrix von $\nabla f: U \rightarrow \mathbb{R}^n$.

Beispiel 8.6. Sei $f(x, y) = x^2y^2 - 3xy$. Dann haben wir

$$\nabla f(x, y) = (2xy^2 - 3y, 2x^2y - 3x), \quad H_f(f, x) = \begin{pmatrix} 2y^2 & 4xy - 3 \\ 4xy - 3 & 2x^2 \end{pmatrix}.$$

Satz 8.7. Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $\nabla f(x_0) = 0$ für ein $x_0 \in U$. Ist

- $H_f(x_0)$ positiv definit, so ist x_0 lokale Minimumstelle von f ,
- $H_f(x_0)$ negativ definit, so ist x_0 lokale Maximumstelle von f ,
- indefinit, so ist x_0 keine lokale Extremstelle von f , sondern ein Sattelpunkt.

Beispiel 8.8. Für die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = (x^2 + 2y^2)e^{-x^2 - y^2}$ suchen wir die kritischen Punkte; also alle diejenigen (x, y) mit $\nabla f(x, y) = 0$. Dazu bestimmen wir ∇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) 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:

$$(0, 0), \quad (0, 1), \quad (0, -1), \quad (1, 0), \quad (-1, 0).$$

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

$$H_f(x, y) = 2e^{-x^2 - y^2} \begin{pmatrix} 1 - 9x^2 - 2y^2 + 4x^4 + 8x^2y^2 & -2xy(3 - x^2 - 2y^2) \\ -2xy(3 - x^2 - 2y^2) & 2 - x^2 - 10y^2 + 2x^2y^2 + 4y^4 \end{pmatrix}.$$

Wir setzen darin die kritischen Stellen ein:

$$H_f(0, 0) = 2 \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{mit Eigenwerten 2 und 4, also hat } f \text{ lokales Minimum in } (0, 0),$$

$$H_f(0, 1) = \frac{2}{e} \begin{pmatrix} 1 - 2 & 0 \\ 0 & 2 - 10 + 4 \end{pmatrix} = \begin{pmatrix} -\frac{2}{e} & 0 \\ 0 & -\frac{8}{e} \end{pmatrix} \quad \text{mit EW } -\frac{2}{e} \text{ und } -\frac{8}{e}, \text{ also hat } f \text{ lok. Max. in } (0, 1),$$

$$H_f(0, -1) = H_f(0, 1), \quad \text{also ebenso lokales Maximum in } (0, -1),$$

$$H_f(1, 0) = \frac{2}{e} \begin{pmatrix} 1 - 9 + 4 & 0 \\ 0 & 2 - 1 \end{pmatrix} = \begin{pmatrix} -\frac{8}{e} & 0 \\ 0 & \frac{2}{e} \end{pmatrix} \quad \text{mit EW } -\frac{8}{e} \text{ und } +\frac{2}{e}, \text{ also einem Sattelpunkt in } (1, 0),$$

$$H_f(-1, 0) = H_f(1, 0), \quad \text{also ebenso ein Sattelpunkt in } (-1, 0).$$

Warnung: Für Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}$ kann mehr passieren als für Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$. Zum Beispiel ist es jetzt möglich, dass eine Funktion zwei Hochpunkte hat, aber keinen Tiefpunkt.

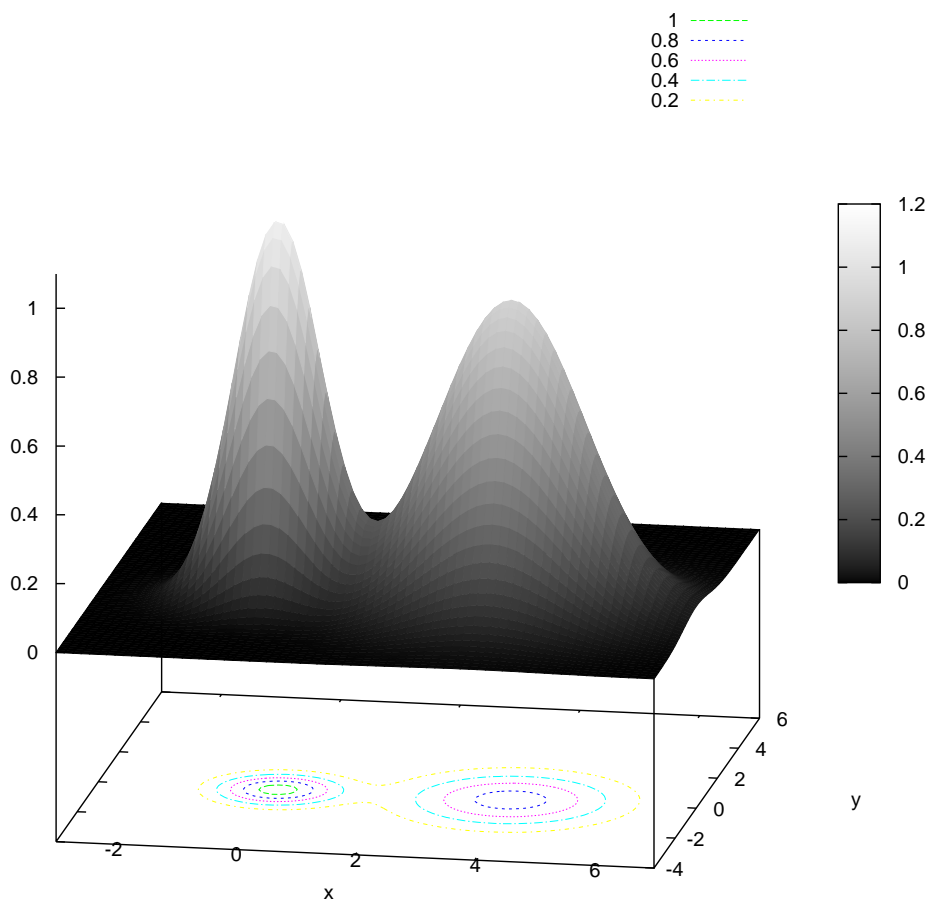


Abbildung 3.1: Achtung: diese Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch die Funktionsgleichung $f(x, y) = 1.1e^{-x^2-y^2} + 0.9e^{-(x-3.9)^2/2-y^2/3}$ hat zwei Maxima, aber dazwischen kein Minimum (denn sie hat überhaupt kein Minimum auf dem \mathbb{R}^2). Funktionen im \mathbb{R}^2 können sich also anders verhalten als Funktionen im \mathbb{R}^1 .

9 Extremwerte unter Nebenbedingungen

Problem: es sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ sei stetig differenzierbar, und sei $g: U \rightarrow \mathbb{R}$ gegeben. Wir setzen

$$M := \{x \in U : g(x) = 0\}.$$

Gesucht: die Extremwerte von f auf M .

Beispiel 9.1. Wir betrachten eine quaderförmige Schachtel ohne Deckel. Die Kantenlängen der Grundfläche seien x und y ; die Höhe sei z . Wir wollen die Oberfläche

$$O(x, y, z) = xy + 2xz + 2yz$$

minimieren bei vorgegebenem Volumen

$$V(x, y, z) = xyz.$$

Hier ist $g(x, y, z) = xyz - V$.

Die Vorstellung dahinter ist: ein Kohlenofen hat unten einen Aschekasten. Dieser Blechkasten soll hergestellt werden unter minimalem Blechverbrauch, bei vorgegebenem Fassungsvermögen. (Es sollte klar sein, dass Ofenbauer die Aschekästen nach anderen Kriterien konzipieren und unsere Interpretation lediglich die Vorstellung erleichtern soll.)

Eine **erste Methode** besteht in der Elimination einer Variablen, also $z = \frac{V}{xy}$. Dann ist

$$O(x, y, z) = xy + 2x \frac{V}{xy} + 2y \frac{V}{xy} = xy + \frac{2V}{y} + \frac{2V}{x} =: f(x, y).$$

Wir untersuchen f auf Extremalstellen:

$$\begin{aligned} \nabla f(x, y) &= \left(y - \frac{2V}{x^2}, x - \frac{2V}{y^2} \right) \stackrel{!}{=} (0, 0), \\ \implies y &= \frac{2V}{x^2} \quad \text{und} \quad x = \frac{2V}{y^2} \\ \implies y &= \frac{2Vy^4}{(2V)^2} \quad \text{und} \quad x = \frac{2V}{y^2} \\ \implies y &= \frac{y^4}{2V} \quad \text{und} \quad x = \frac{2V}{y^2} \\ \implies y &= \sqrt[3]{2V} \quad \text{und} \quad x = \frac{2V}{(\sqrt[3]{2V})^2} = (2V)^{1-2/3} = (2V)^{1/3} = \sqrt[3]{2V}. \end{aligned}$$

Ist $(x_0, y_0) = (\sqrt[3]{2V}, \sqrt[3]{2V})$ tatsächlich eine Minimumstelle? Dazu schreiben wir uns die Hesse-Matrix auf:

$$H_f(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix} = \begin{pmatrix} \frac{4V}{x^3} & 1 \\ 1 & \frac{4V}{y^3} \end{pmatrix}.$$

Somit ist

$$H_f(x_0, y_0) = \begin{pmatrix} \frac{4V}{2V} & 1 \\ 1 & \frac{4V}{2V} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Die Untermioren sind

$$|2| = 2 > 0, \quad \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 4 - 1 = 3 > 0$$

also ist $H_f(x_0, y_0)$ positiv definit, und f hat in (x_0, y_0) eine Minimumstelle.

Die minimale Oberfläche wird für $x = y = \sqrt[3]{2V}$ und $z = \frac{V}{xy} = \frac{1}{2} \sqrt[3]{2V}$ angenommen.

Beispiel 9.2. Gesucht sei jetzt die Extremstelle von $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^3 - y^2$ unter der Nebenbedingung $x^2 + y^2 = 1$. Es ist also $g(x, y) = x^2 + y^2 - 1$. Siehe Bild 3.2.

Hier ist es ebenfalls möglich, die Gleichung $g(x, y) = 0$ nach einer der beiden Variablen umzustellen (es bietet sich y an) und dies dann in f einzusetzen (sodass bloß noch die Variable x übrigbleibt) und anschließend weiterzurechnen. Es gibt allerdings noch eine zweite Methode, die wir jetzt vorbereiten wollen.

Wir stellen nun eine **zweite Methode** vor, deren Idee in Bild 3.3 erklärt wird.

Theorem 9.3 (Methode von Lagrange). Seien $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}$ und $g: D \rightarrow \mathbb{R}$ stetig differenzierbar.

Wenn die Funktion f unter der Nebenbedingung $f(x) = 0$ ein relatives Extremum in $x_0 \in D$ hat, und wenn dort $\nabla g(x_0) \neq 0$ gilt, dann existiert ein $\lambda \in \mathbb{R}$ mit $\nabla f(x_0) + \lambda \nabla g(x_0) = 0$.

Die **Anwendung** geschieht wie folgt: Gegeben seien die Funktionen f und g . Wir definieren die Lagrangefunktion

$$L(x, \lambda) := f(x) + \lambda \cdot g(x).$$

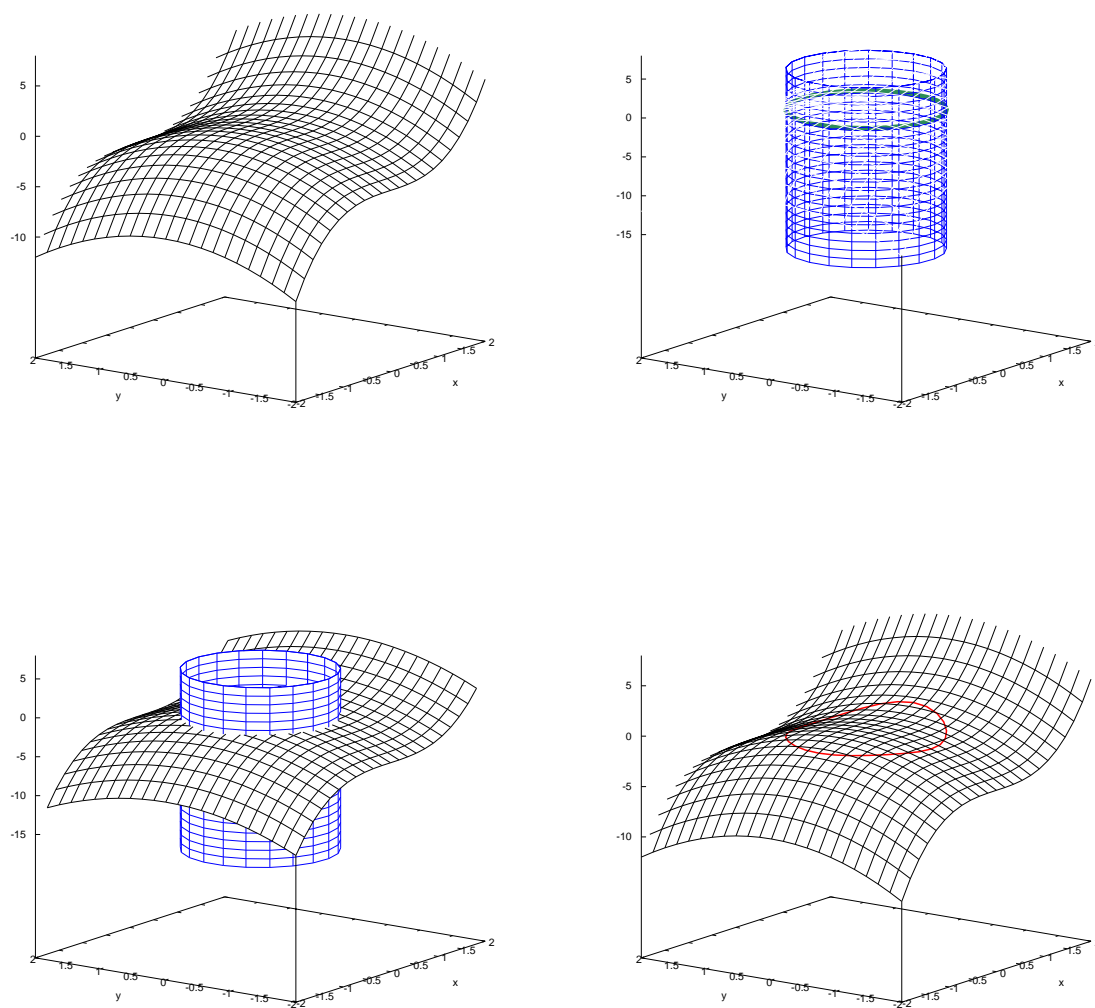


Abbildung 3.2: Links oben: die durch die Funktion $z = f(x, y) = x^3 - y^2$ beschriebene Fläche. Rechts oben: die Nebenbedingung $x^2 + y^2 = 1$ beschreibt in der xy -Ebene einen Kreis (grün). Wenn wir diesen entlang der z -Achse parallel verschieben, entsteht ein blauer Zylinder. Links unten: wir schneiden 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 9.2 (a).

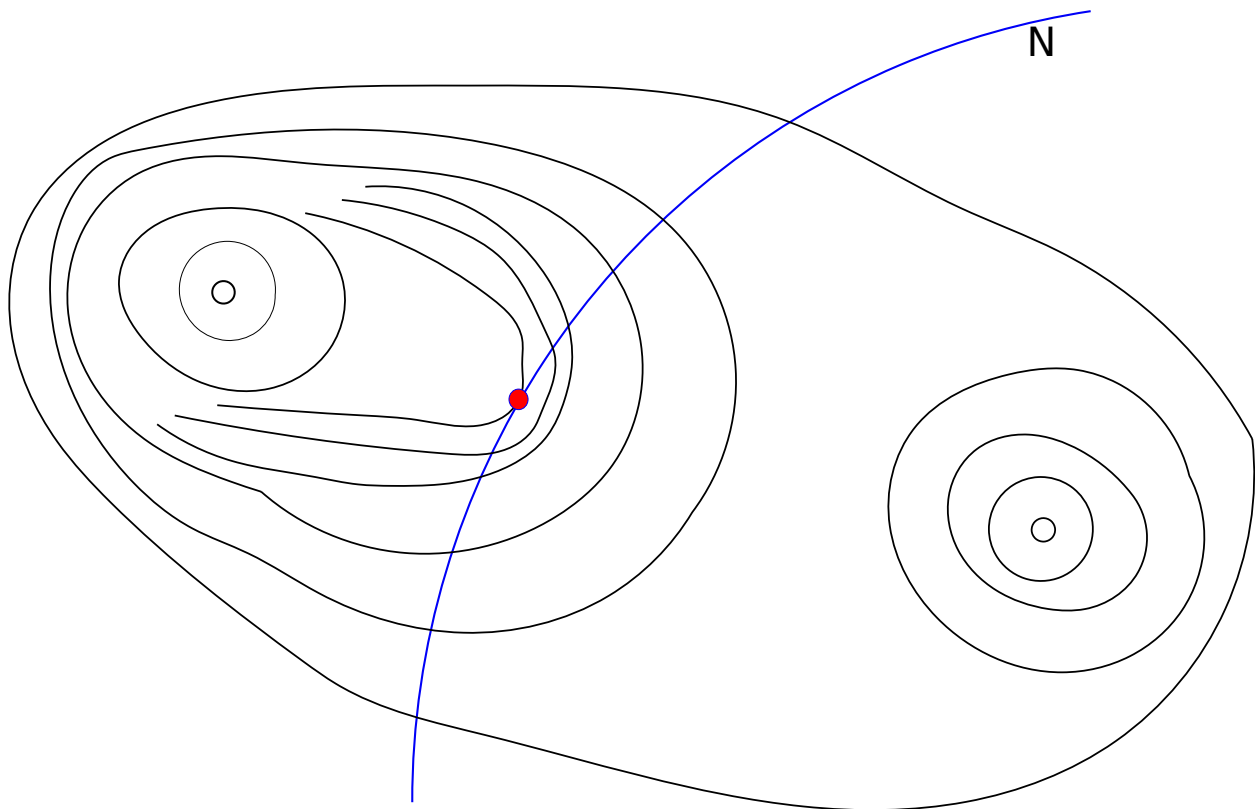


Abbildung 3.3: Eine handgemalte Skizze zur Suche nach Maxima der Funktion $f(\vec{x})$ für $\vec{x} \in \mathbb{R}^2$ unter der Nebenbedingung $g(\vec{x}) = 0$. Die Menge aller derjenigen $\vec{x} \in \mathbb{R}^2$ mit $g(\vec{x}) = 0$ ist gerade die Menge N (blaue Kurve), und dies ist die Höhenlinie 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. Nun erinnern wir uns aber an Folgendes: die Gradientenvektoren stehen senkrecht auf den Höhenlinien. Also gilt $\nabla f(\vec{x}_0) \parallel \nabla g(\vec{x}_0)$, wenn \vec{x}_0 der rote Punkt ist. Die Parallelität der beiden Gradientenvektoren schreiben wir als $\nabla f(\vec{x}_0) = -\lambda \nabla g(\vec{x}_0)$, bzw. als $\nabla f(\vec{x}_0) + \lambda \nabla g(\vec{x}_0) = \vec{0}$. Das ist aber gerade die Lagrange-Bedingung.

Eine notwendige Bedingung für das Vorliegen eines Extremums ist dann

$$\begin{aligned}\frac{\partial}{\partial x_1} L(x, \lambda) &\stackrel{!}{=} 0, \\ \frac{\partial}{\partial x_2} L(x, \lambda) &\stackrel{!}{=} 0, \\ &\dots \\ \frac{\partial}{\partial x_n} L(x, \lambda) &\stackrel{!}{=} 0, \\ \frac{\partial}{\partial \lambda} L(x, \lambda) &\stackrel{!}{=} 0.\end{aligned}$$

Dies sind $n + 1$ Gleichungen für die $n + 1$ Unbekannten x_1, \dots, x_n und λ .

Beispiel 9.4. Wir suchen die Extremstellen von $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^3 + 3y$, unter der Nebenbedingung $x^2 + 4y^2 = 1$. Es ist also $g(x, y) = x^2 + 4y^2 - 1$.

$$L(x, y) = f(x, y) + \lambda g(x, y) = x^3 + 3y + \lambda(x^2 + 4y^2 - 1)$$

Wir haben $\nabla g(x, y) = (2x, 8y) = (0, 0)$ genau dann, wenn $x = y = 0$; also ist (wie gewünscht) $\nabla g(x, y) \neq (0, 0)$ für alle diejenigen x, y , die die Nebenbedingung $g(x, y) = 0$ erfüllen.

Die notwendigen Bedingungen für das Extremum sind dann

$$(0, 0, 0) \stackrel{!}{=} \left(\frac{\partial L}{\partial x}(x, y, \lambda), \frac{\partial L}{\partial y}(x, y, \lambda), \frac{\partial L}{\partial \lambda}(x, y, \lambda) \right) = (3x^2 + 2\lambda, 3 + 8\lambda y, x^2 + 4y^2 - 1).$$

Das heißt:

$$3x^2 + 2\lambda = 0, \tag{I}$$

$$3 + 8\lambda y = 0, \tag{II}$$

$$x^2 + 4y^2 = 1, \tag{III}$$

Im **ersten Fall** sei $x = 0$. Dann folgt aus (III), daß $4y^2 = 1$, d.h. $y = \pm \frac{1}{2}$, und aus (II) bekommen wir dann $3 + 4\lambda = 0$, also $\lambda = -\frac{3}{4}$. Die kritischen Stellen sind $(0, \frac{1}{2})$ und $(0, -\frac{1}{2})$.

Im **zweiten Fall** sei $x \neq 0$. Aus (I) folgt dann $3x + 2\lambda = 0$ bzw. $\lambda = -3x/2$. Mit (II) entsteht $3 - 12xy = 0$, also $xy = 1/4$, also $y = 1/(4x)$. Damit gehen wir zu (III):

$$x^2 + 4 \cdot \frac{1}{16x^2} = 1 \implies x^4 + \frac{1}{4} = x^2 \implies x^4 - x^2 + \frac{1}{4} = 0.$$

Wir taufen $z = x^2$ und erhalten dann $z^2 - z + 1/4 = 0$. also

$$z_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{4}} = \frac{1}{2},$$

also $x = \pm 1/\sqrt{2}$. Es folgt dann mit $y = 1/(4x)$, daß

$$y = \frac{1}{\pm 4 \cdot \frac{1}{\sqrt{2}}} = \pm \frac{\sqrt{2}}{4} = \pm \frac{1}{2\sqrt{2}}.$$

Die kritischen Stellen im zweiten Fall sind somit

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}} \right), \quad \left(-\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}} \right).$$

Wir bestimmen die Funktionswerte an den kritischen Stellen:

$$\begin{aligned}f\left(0, \frac{1}{2}\right) &= \frac{3}{2}, & f\left(0, -\frac{1}{2}\right) &= -\frac{3}{2}, \\ f\left(\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}\right) &= \frac{1}{2\sqrt{2}} + \frac{3}{2\sqrt{2}} = \frac{4}{2\sqrt{2}} = \frac{2}{\sqrt{2}} = \sqrt{2}, \\ f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) &= -\frac{1}{2\sqrt{2}} - \frac{3}{2\sqrt{2}} = -\sqrt{2}.\end{aligned}$$

Der Maximalwert $\frac{3}{2}$ wird bei $(0, \frac{1}{2})$ angenommen; der Minimalwert $-\frac{3}{2}$ bei $(0, -\frac{1}{2})$.

Wir kehren zurück zur Aschenkastenaufgabe.

Beispiel 9.5. Zu maximieren ist $f(x, y, z) = xy + 2xz + 2zy$ unter der Nebenbedingung $g(x, y, z) = xyz - V \stackrel{!}{=} 0$.

Die Lagrangefunktion ist

$$L(x, y, z, \lambda) := f(x, y, z) + \lambda \cdot g(x, y, z) = (xy + 2xz + 2zy) + \lambda \cdot (xyz - V),$$

und jetzt setzen wir ihre Ableitungen gleich Null:

$$\begin{aligned} \frac{\partial L}{\partial x}(x, y, z) &= y + 2z + \lambda yz \stackrel{!}{=} 0, \\ \frac{\partial L}{\partial y}(x, y, z) &= x + 2z + \lambda xz \stackrel{!}{=} 0, \\ \frac{\partial L}{\partial z}(x, y, z) &= 2x + 2y + \lambda xy \stackrel{!}{=} 0, \\ \frac{\partial L}{\partial \lambda}(x, y, z) &= xyz - V \stackrel{!}{=} 0. \end{aligned}$$

Dies sind vier Gleichungen für vier Unbekannte (x, y, z, λ) , wobei am Ende die Zahl λ relativ unwichtig sein wird. Leider ist dieses Gleichungssystem nichtlinear, was man an den Produkten λyz , λxz , λxy und xyz erkennt. Wir schreiben die ersten drei Gleichungen anders:

$$\begin{array}{l|l} y + 2z = -\lambda yz & \cdot x \\ x + 2z = -\lambda xz & \cdot y \\ 2x + 2y = -\lambda xy & \cdot z, \end{array}$$

und es entsteht

$$xy + 2xz = -\lambda xyz \tag{I}$$

$$xy + 2yz = -\lambda xyz \tag{II}$$

$$2xz + 2yz = -\lambda xyz \tag{III}$$

Der Zweck dieser Multiplikation war, dass wir jetzt die unangenehmen rechten Seiten durch Subtraktion vernichten können:

$$\begin{aligned} (II) - (I): \quad 2yz - 2xz &= 0 & \implies & (y - x)z = 0, \\ (III) - (I): \quad 2yz - 2xy &= 0 & \implies & y(2z - x) = 0. \end{aligned}$$

Laut der ersten Gleichung ist $z = 0$ (was absurderweise auf einen plattgedrückten Kasten ohne Volumen führt, der die Nebenbedingung $xyz - V = 0$ niemals erfüllen kann) oder $y - x = 0$. Analog argumentieren wir mit der zweiten Gleichung: die Variante $y = 0$ ist absurd, also ist $2z - x = 0$.

Wir haben also $x = y = 2z$, und damit gehen wir in die Nebenbedingung hinein:

$$V = xyz = x \cdot x \cdot \frac{1}{2}x = \frac{x^3}{2},$$

woraus wir $x = \sqrt[3]{2V}$ gewinnen, und dann haben wir y und z sehr schnell.

Die Unbekannte λ könnten wir berechnen, aber daraus gewinnen wir keine Erkenntnis.

Kapitel 4

Gewöhnliche Differentialgleichungen

10 Erste Beispiele und Begriffe

Beispiel 10.1 (natürliches Wachstum). Wir denken uns eine Population von Individuen (bzw. Mikroorganismen oder ein chemischer Stoff). Sei $N(t)$ die Anzahl (in Stück) dieser Individuen zur Zeit t . Wir setzen voraus, dass die Geburtenrate und die Sterberate zeitlich konstant sind. Seien also

β die Anzahl der Geburten pro Individuum pro Zeiteinheit,

δ die Anzahl der Todesfälle pro Individuum pro Zeiteinheit.

Wir betrachten ein kurzes Zeitintervall $[t, t + \Delta t]$. Währenddessen gibt es dann ungefähr $\beta N(t)\Delta t$ Geburten und $\delta N(t)\Delta t$ Todesfälle. Hierbei haben wir vorausgesetzt, dass im betrachteten Zeitintervall die Anzahl N sich nicht erheblich ändert. Es folgt dann

$$N(t + \Delta t) = N(t) + \beta N(t)\Delta t - \delta N(t)\Delta t,$$
$$\frac{N(t + \Delta t) - N(t)}{\Delta t} = (\beta - \delta)N(t).$$

Jetzt schicken wir formal $\Delta t \rightarrow 0$ und erhalten damit

$$N'(t) = \lim_{\Delta t \rightarrow 0} \frac{N(t + \Delta t) - N(t)}{\Delta t} = (\beta - \delta)N(t).$$

Hierbei haben wir etwas gemogelt, denn zu Beginn war vorausgesetzt worden, dass $N(t)$ für jeden Zeitpunkt t eine ganze Zahl ist, also ist dann $N = N(t)$ eine Funktion, die stückweise konstant ist, also unstetig, also gar nicht differenzierbar. Trotzdem ist es sinnvoll, die Funktion N als reellwertig und stetig differenzierbar anzunehmen, denn z.B. im Falle von chemischen Reaktionen wird $N(t) \sim 10^{24}$ sein, und dann ist es ein vernachlässigbarer Fehler, N als kontinuierliche Funktion anzunehmen.

Wir erhalten also die Differentialgleichung $N'(t) = (\beta - \delta)N(t)$. Dies ist eine Differentialgleichung erster Ordnung, weil nur die erste Ableitung vorkommt.

Ein Beispiel für eine Differentialgleichung dritter Ordnung ist

$$f'''(t) - 2f'(t) = (f(t))^4.$$

Wir erinnern an HM I, Beispiel 5.3.4 b): alle Lösungen der Gleichung $f'(t) = kf(t)$ sind gegeben durch $f(t) = ce^{kt}$, wobei $c \in \mathbb{R}$ ein beliebig wählbarer Parameter ist. Die Gleichung $f(t) = ce^{kt}$ heißt auch *Darstellung der allgemeinen Lösung*.

Ist zum Zeitpunkt t_0 die Größe der Startpopulation y_0 gegeben, so heißt

$$\begin{cases} f'(t) = kf(t), \\ f(t_0) = y_0 \end{cases}$$

Anfangswertproblem. Wir können jetzt den noch beliebig wählbaren Parameter c so wählen, dass die Anfangsbedingung $f(t_0) = y_0$ erfüllt wird, und es ergibt sich dann die eindeutige Lösung

$$f(t) = y_0 e^{-kt_0} e^{kt}$$

für das Anfangswertproblem.

Beispiel 10.2. Wir betrachten einige typische Differentialgleichungen der Reaktionskinetik:

- (a) eine Zerfallsreaktion $A \xrightarrow{k} B + C$ und eine Isomerisierung $A \xrightarrow{k} B$. Hierbei ist k eine Reaktionsgeschwindigkeitskonstante. Wenn $c_A(t)$ die Konzentration der Molekülsorte A zur Zeit t angibt, dann lautet in beiden Fällen die Differentialgleichung

$$c'_A(t) = -kc_A(t).$$

- (b) die elementare Reaktion $2A \xrightarrow{k} B + C$. (In der Chemie spricht man von einer elementaren Reaktion, wenn sie nicht weiter in Einzelreaktionen zerlegt werden kann). Die Differentialgleichung für die Konzentration des Stoffs A ist

$$c'_A(t) = -kc_A^2(t).$$

- (c) eine elementare Reaktion $A + B \xrightarrow{k} P$. Dann erhalten wir das gekoppelte System von Differentialgleichungen

$$\begin{aligned} c'_A(t) &= -kc_A(t)c_B(t), \\ c'_B(t) &= -kc_A(t)c_B(t), \\ c'_P(t) &= kc_A(t)c_B(t). \end{aligned}$$

- (d) eine Folgereaktion $A \xrightarrow{k_1} B \xrightarrow{k_2} C$. Das zugehörige System von Differentialgleichungen lautet

$$\begin{aligned} c'_A(t) &= -k_1c_A(t), \\ c'_B(t) &= k_1c_A(t) - k_2c_B(t), \\ c'_C(t) &= k_2c_B(t). \end{aligned}$$

Dieses System ist linear, weil die gesuchten Funktionen c_A , c_B , c_C nicht etwa quadriert oder in gemischten Produkten auftauchen oder sonstwie, sondern nur auf lineare Weise. Wenn wir diese Funktionen zu einem Vektor y zusammenfassen, also $y(t) := (c_A(t), c_B(t), c_C(t))^T$, dann ist

$$y'(t) = \begin{pmatrix} c'_A(t) \\ c'_B(t) \\ c'_C(t) \end{pmatrix} = \begin{pmatrix} -k_1c_A(t) \\ k_1c_A(t) - k_2c_B(t) \\ k_2c_B(t) \end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{pmatrix} \begin{pmatrix} c_A(t) \\ c_B(t) \\ c_C(t) \end{pmatrix} =: Ay(t),$$

wobei A die genannte 3×3 -Matrix ist.

11 Getrennte Variablen

Definition 11.1. Eine Differentialgleichung erster Ordnung der Form $y'(t) = g(t) \cdot h(y(t))$ heißt Differentialgleichung mit getrennten Variablen.

Beispiel 11.2. (a) $y'(t) = -ky^2(t)$ (siehe Beispiel 10.2, Teil (b)). Hierbei haben wir $g(t) = -k$ und $h(y) = y^2$.

(b) $y'(t) = e^{y(t)} \sin(t)$. Hierbei ist $g(t) = \sin(t)$ und $h(y) = e^y$.

(c) $y'(t) = y^2(t) + t^2$ hat keine getrennten Variablen.

Die **Lösungsidee** ist: falls $h(y(t)) \neq 0$ ist, dann haben wir

$$\frac{y'(t)}{h(y(t))} = g(t),$$

und Integration dieser Gleichung bringt uns dann

$$\int g(t) dt = \int \frac{y'(t)}{h(y(t))} dt \stackrel{\text{Subst.}}{=} \int \frac{dy}{h(y)}. \quad (4.1)$$

Wenn wir die Integrale berechnen, dann finden wir alle Lösungen $y(t)$, für die $h(y(t)) \neq 0$ gilt. Der Fall $h(y(t)) = 0$ muss gesondert behandelt werden.

Beispiel 11.3. (a) Wir lösen $y'(t) = ky(t)$. Hier ist $k(t) = k$ und $h(y) = y$. Dann liefert (4.1) uns

$$\int k dt = \int \frac{dy}{y}, \quad \text{also} \quad kt + c_1 = \ln(|y|),$$

und daraus bekommen wir

$$|y| = e^{kt+c_1}, \quad \text{also} \quad y(t) = e^{c_1}e^{kt} \quad \text{oder} \quad y(t) = -e^{c_1}e^{kt}.$$

Beide Fälle lassen sich zusammenfassen zu $y(t) = ce^{kt}$ für ein $c \neq 0$. Andererseits ist $y(t) \equiv 0$ offensichtlich auch eine Lösung. Also sind alle Lösungen gegeben durch

$$y(t) = ce^{kt}, \quad c \in \mathbb{R} \quad \text{frei wählbar.}$$

(b) Das Anfangswertproblem

$$y'(t) = -ky^2(t), \quad y(0) = y_0$$

mit $k > 0$, $y_0 > 0$ soll gelöst werden. Hier ist $g(t) = -k$ und $h(y) = y^2$. Dann liefert (4.1) uns

$$\int (-k) dt = \int \frac{dy}{y^2}, \quad \text{also} \quad -kt + c = -\frac{1}{y}, \quad \text{also} \quad y(t) = \frac{1}{kt - c},$$

wobei $c \in \mathbb{R}$ beliebig wählbar ist. Die Anfangsbedingung soll auch noch gelten, also

$$y_0 = y_0 = \frac{1}{0 - c} = -\frac{1}{c}, \quad \text{also} \quad c = -\frac{1}{y_0}.$$

Wir erhalten dann insgesamt die Lösungsdarstellung

$$y(t) = \frac{1}{\frac{1}{y_0} + k \cdot t}.$$

(c) Wir betrachten die Elementarreaktion $A + B \xrightarrow{k} P$ und das dazugehörige System von Differentialgleichungen

$$c'_A(t) = -kc_A(t)c_B(t), \tag{I}$$

$$c'_B(t) = -kc_A(t)c_B(t), \tag{II}$$

$$c'_P(t) = kc_A(t)c_B(t), \tag{III}$$

Die Anfangsbedingungen seien $c_A(0) = a$, $c_B(0) = b$, $c_P(0) = 0$ wobei a und b positive gegebene Zahlen seien.

Die gemischten Produkte auf der rechten Seite machen das System nichtlinear, was Schwierigkeiten bringt. Da wir uns für die Herstellung des Produkts P interessieren, ist in erster Linie die letzte Zeile interessant.

Anschaulich ist klar, dass für die Herstellung eines Moleküls von P genau ein Molekül von A verbraucht wird (und ein Molekül von B). Wir vermuten also die beiden Identitäten

$$c_A(t) = a - c_P(t), \quad c_B(t) = b - c_P(t),$$

was sich umstellt zu $c_A(t) + c_P(t) = a$ und $c_B(t) + c_P(t) = b$. Diese beiden vermuteten Gleichungen können wir sogar beweisen, denn es ist

$$\left(c_A(t) + c_P(t) \right)' = c'_A(t) + c'_P(t) = -kc_A(t)c_B(t) + kc_A(t)c_B(t) = 0$$

für alle $t \geq 0$, also ist die Summenfunktion $c_A(t) + c_P(t)$ zeitlich konstant. Da zur Zeit 0 aber $c_A(0) + c_P(0) = a + 0 = a$ gilt, ist tatsächlich $c_A(t) + c_P(t) = a$ für alle $t \geq 0$. Analog beweist man die Gleichung für $c_B(t)$.

Damit gehen wir in (III) hinein und finden

$$c'_P(t) = kc_A(t)c_B(t) = k \cdot (a - c_P(t)) \cdot (b - c_P(t)).$$

Hier können wir die Variablen trennen durch die Setzung $g(t) = k$ und $h(c_P) = (a - c_P) \cdot (b - c_P)$. Aus (4.1) bekommen wir dann

$$\int \frac{dc_P}{(a - c_P)(b - c_P)} = \int k dt = kt + c$$

mit beliebig wählbarem $c \in \mathbb{R}$. Jetzt haben wir noch das Integral

$$\int \frac{dx}{(a - x)(b - x)}$$

zu bestimmen. Im **ersten Fall** sei $a = b$, und dann haben wir

$$\int \frac{dx}{(a - x)^2} = \frac{2}{a - x} + C.$$

Diesen Fall wollen wir ab jetzt vernachlässigen. Im **zweiten Fall** ist $a \neq b$. Für eine Partialbruchzerlegung machen wir den Ansatz

$$\frac{1}{(a - x)(b - x)} = \frac{A}{a - x} + \frac{B}{b - x}.$$

Wir multiplizieren die Nenner hoch und erhalten $1 = A(b - x) + B(a - x)$. Hierin setzen wir $x = a$ und bekommen $A = \frac{1}{b-a}$; und wir setzen $x = b$ und bekommen $B = \frac{1}{a-b}$. Also ist

$$\frac{1}{(a - x)(b - x)} = \frac{1}{a - b} \left(\frac{1}{b - x} - \frac{1}{a - x} \right).$$

Damit ergibt sich die Integration

$$\int \frac{dx}{(a - x)(b - x)} = \frac{1}{a - b} (-\ln |b - x| + \ln |a - x|) + d = \frac{1}{a - b} \ln \left| \frac{a - x}{b - x} \right| + d,$$

mit einer Integrationskonstanten d . Insgesamt entsteht dann

$$\frac{1}{a - b} \ln \left| \frac{a - c_P(t)}{b - c_P(t)} \right| = kt + C,$$

wobei $C = c - d$. Wir bestimmen C mit Hilfe des Anfangswerts zur Zeit $t = 0$:

$$C = \frac{1}{a - b} \ln \left| \frac{a - c_P(0)}{b - c_P(0)} \right| = \frac{1}{a - b} \ln \left| \frac{a}{b} \right|.$$

Also ist

$$\begin{aligned} kt &= \frac{1}{a - b} \left(\ln \left| \frac{a - c_P(t)}{b - c_P(t)} \right| - \ln \left| \frac{a}{b} \right| \right) = \frac{1}{a - b} \ln \left| \frac{b(a - c_P(t))}{a(b - c_P(t))} \right| \\ \implies (a - b)kt &= \ln \left| \frac{b(a - c_P(t))}{a(b - c_P(t))} \right| \implies e^{(a-b)kt} = \frac{b}{a} \left| \frac{a - c_P(t)}{b - c_P(t)} \right|. \end{aligned}$$

Anschaulich ist klar, dass $c_P(t) \leq a$ und $c_P(t) \leq b$ für alle $t \geq 0$ gelten muss, denn ein Molekül P entsteht aus einem Molekül A und einem Molekül B . Also folgt $a - c_P \geq 0$ und $b - c_P(t) \geq 0$, sodaß der Betrag weggelassen werden kann. Mathematisch müsste dies aber anhand der Lösungsdarstellung nachträglich verifiziert werden. Also ist

$$\frac{a}{b} e^{(a-b)kt} = \frac{a - c_P(t)}{b - c_P(t)} \implies \dots \implies c_P(t) = ab \frac{1 - e^{k(a-b)t}}{b - ae^{k(a-b)t}}.$$

12 Lineare Differentialgleichungen erster Ordnung

Lineare Differentialgleichungen erster Ordnung haben die allgemeine Form

$$y'(t) + a(t)y(t) = f(t), \tag{LDGL}$$

wobei a und f gegebene Funktionen sind, und $y(t)$ ist gesucht.

Die Differentialgleichung

$$y'(t) + a(t)y(t) = 0$$

heißt die zu (LDGL) gehörige homogene Gleichung.

Bemerkung 12.1. Die homogene lineare Differentialgleichung $y'(t) + a(t)y(t) = 0$ kann geschrieben werden als $y'(t) = -a(t)y(t)$, hat also getrennte Variablen, mit $g(t) = -a(t)$ und $h(y) = y$.

Sei $A(t)$ eine Stammfunktion von $a(t)$. Dann haben wir

$$\begin{aligned} \int \frac{dy}{h(y)} = \int g(t) dt &\iff \int \frac{dy}{y} = \int -a(t) dt &\iff \ln |y| = -A(t) + C_1 \\ &\iff |y(t)| = e^{C_1} e^{-A(t)} &\implies \dots &\implies y(t) = C e^{-A(t)} \end{aligned}$$

für einen beliebig wählbaren Parameter $C \in \mathbb{R}$, mit einer Begründung wie in Abschnitt 11.

Satz 12.2. Ist y_s eine Lösung von (LDGL) (auch spezielle Lösung genannt), so erhält man alle Lösungen von (LDGL) durch

$$y(t) = y_s(t) + y_h(t),$$

wobei y_h die in Bemerkung 12.1 ermittelte Lösung der zugehörigen homogenen Gleichung ist.

Damit erhalten wir ein Lösungsverfahren von (LDGL):

- bestimme alle Lösungen der zugehörigen homogenen Gleichung (siehe Bemerkung 12.1),
- bestimme eine spezielle Lösung y_s ,
- dann bilden alle Lösungen von (LDGL) gerade die Menge $\{y_s + y_h : y_h \text{ löst die zugehörige homogene Gleichung}\}$.

Die eine spezielle Lösung y_s im zweiten Schritt kann man durch Raten ermitteln, oder man benutzt die Methode der **Variation der Konstanten**:

Wir erinnern daran, daß alle Lösungen der zugehörigen homogenen Gleichung die Gestalt $y_h(t) = C e^{-A(t)}$ haben, wobei $C \in \mathbb{R}$ eine beliebig wählbare Konstante ist.

Jetzt machen wir den Ansatz $y_s(t) = C(t)e^{-A(t)}$, das heißt, anstatt einer Konstanten C versuchen wir unser Glück mit einer sich variierenden Funktion $C(t)$ (was den Namen der Methode erklärt). Diesen Ansatz setzen wir in (LDGL) ein:

$$\begin{aligned} y'(t) &= C'(t)e^{-A(t)} + C(t)e^{-A(t)} \cdot (-A'(t)) \\ &= C'(t)e^{-A(t)} - a(t)C(t)e^{-A(t)} \\ &= C'(t)e^{-A(t)} - a(t)y_s(t), \\ f(t) &\stackrel{!}{=} y_s'(t) + a(t)y_s(t) = C'(t)e^{-A(t)}, \end{aligned}$$

also folgt $C'(t) = e^{A(t)} f(t)$, und jetzt finden wir $C(t)$ durch Integration:

$$C(t) = \int_0^t e^{A(\tau)} f(\tau) d\tau$$

und damit ist dann

$$y_s(t) = e^{-A(t)} \int_0^t e^{A(\tau)} f(\tau) d\tau.$$

Wir fassen zusammen: die Lösungsformel für (LDGL) ist

$$y(t) = y_s(t) + y_h(t) = e^{-A(t)} \int_0^t e^{A(\tau)} f(\tau) d\tau + c e^{-A(t)},$$

wobei A eine Stammfunktion von a ist und $c \in \mathbb{R}$ eine beliebig wählbare Konstante.

Beispiel 12.3. Wir betrachten die Folgereaktion $A \xrightarrow{k_1} B \xrightarrow{k_2} C$ und das zugehörige Anfangswertproblem

$$\begin{aligned} c_A'(t) &= -k_1 c_A(t), & c_A(0) &= a > 0, \\ c_B'(t) &= k_1 c_A(t) - k_2 c_B(t), & c_B(0) &= 0, \\ c_C'(t) &= k_2 c_B(t), & c_C(0) &= 0. \end{aligned}$$

Gesucht ist die Funktion c_C .

Im **ersten Schritt** lösen wir

$$c'_A(t) = -k_1 c_A(t), \quad c_A(0) = a.$$

Dies ist eine lineare homogene Differentialgleichung mit Anfangsbedingung, und die Lösung ist $c_A(t) = ae^{-k_1 t}$.

Im **zweiten Schritt** lösen wir

$$c'_B(t) = k_1 a e^{-k_1 t} - k_2 c_B(t), \quad c_B(0) = 0.$$

Die Differentialgleichung können wir schreiben als $c'_B(t) + k_2 c_B(t) = k_1 a e^{-k_1 t}$, was von der Form (LDGL) ist mit $a(t) = k_2$ und $f(t) = k_1 a e^{-k_1 t}$. Dann ist $A(t) = k_2 t$ eine Stammfunktion zu $a(t)$, und wir erhalten alle Lösungen gemäß der Formel

$$\begin{aligned} c_B(t) &= e^{-k_2 t} \int_0^t e^{k_2 \tau} k_1 a e^{-k_1 \tau} d\tau + c e^{-k_2 t} && (c \in \mathbb{R}) \\ &= k_1 a e^{-k_2 t} \int_0^t e^{(k_2 - k_1)\tau} d\tau + c e^{-k_2 t} && (\text{nimm } k_1 \neq k_2 \text{ an}) \\ &= k_1 a e^{-k_2 t} \frac{1}{k_2 - k_1} e^{(k_2 - k_1)\tau} \Big|_0^t + c e^{-k_2 t} \\ &= k_1 a e^{-k_2 t} \frac{1}{k_2 - k_1} \left(e^{(k_2 - k_1)t} - 1 \right) + c e^{-k_2 t} = \frac{k_1 a}{k_2 - k_1} e^{-k_1 t} + \underbrace{\left(c - \frac{k_1 a}{k_2 - k_1} \right)}_{=: c'} \cdot e^{-k_2 t} \\ &= \frac{k_1 a}{k_2 - k_1} e^{-k_1 t} + c' e^{-k_2 t}. \end{aligned}$$

Zur Bestimmung von c' erinnern wir an die Anfangsbedingung $c_B(0) = 0$:

$$0 = c_B(0) = \frac{k_1 a}{k_2 - k_1} + c' \implies c' = -\frac{k_1 a}{k_2 - k_1}.$$

Also ist

$$c_B(t) = \frac{k_1 a}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}).$$

Und im **dritten Schritt** kommen wir zur Bestimmung von c_C :

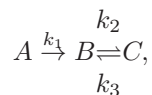
$$c'_C(t) = k_2 c_B(t) = \frac{k_1 k_2 a}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}), \quad c_C(0) = 0.$$

Wir brauchen nur noch zu integrieren und kommen dann nach kurzer Rechnung auf

$$c_C(t) = a \left(1 + \frac{k_1 e^{-k_2 t} - k_2 e^{-k_1 t}}{k_2 - k_1} \right).$$

13 Systeme linearer Differentialgleichungen erster Ordnung

Beispiel 13.1. Wir betrachten die Reaktion



die vom linearen System

$$\begin{aligned} c'_A(t) &= -k_1 c_A(t), \\ c'_B(t) &= k_1 c_A(t) - k_2 c_B(t) + k_3 c_C(t), \\ c'_C(t) &= k_2 c_B(t) - k_3 c_C(t) \end{aligned}$$

beschrieben wird. Wir stapeln die gesuchten Funktionen zu einer Vektorfunktion $y(t)$ auf,

$$y(t) := \begin{pmatrix} c_A(t) \\ c_B(t) \\ c_C(t) \end{pmatrix},$$

und dann erhalten wir

$$y'(t) = \begin{pmatrix} c'_A(t) \\ c'_B(t) \\ c'_C(t) \end{pmatrix} = \begin{pmatrix} -k_1 c_A(t) \\ k_1 c_A(t) - k_2 c_B(t) + k_3 c_C(t) \\ k_2 c_B(t) - k_3 c_C(t) \end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & k_3 \\ 0 & k_2 & k_3 \end{pmatrix} \begin{pmatrix} c_A(t) \\ c_B(t) \\ c_C(t) \end{pmatrix}.$$

Also ist $y'(t) = Ay(t)$ mit

$$A = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & k_3 \\ 0 & k_2 & k_3 \end{pmatrix}.$$

Zum Finden der Lösung wählen wir folgenden **Ansatz**: für einen konstanten Vektor $c \in \mathbb{C}^n$ betrachten wir eine Funktion $y(t) = ce^{\lambda t}$, wobei $\lambda \in \mathbb{C}$ ein Parameter ist, denn wir später wählen werden. Einsetzen liefert

$$y'(t) = \frac{d}{dt} ce^{\lambda t} = c\lambda e^{\lambda t} = \lambda y(t) \stackrel{!}{=} Ay(t) \implies Ace^{\lambda t} = \lambda ce^{\lambda t},$$

also $Ac = \lambda c$. Also ist $y(t) = ce^{\lambda t}$ eine Lösung von $y' = Ay$, wenn c ein Eigenvektor von A ist und λ der zugehörige Eigenwert.

Definition 13.2. Sei $A \in \mathbb{R}^{n \times n}$. Eine Menge $\{y_1(t), \dots, y_n(t)\}$ von Vektor-Funktionen heißt Fundamentalsystem von $y' = Ay$, falls

- $y_1(t), \dots, y_n(t)$ lineare unabhängige Vektoren sind und
- $y_1(t), \dots, y_n(t)$ Lösungen sind von $y'(t) = Ay(t)$.

Satz 13.3. Sei $A \in \mathbb{R}^{n \times n}$ und $\{y_1(t), \dots, y_n(t)\}$ ein Fundamentalsystem von $y'(t) = Ay(t)$. Dann sind alle Lösungen des Systems $y'(t) = Ay(t)$ von der Form

$$y(t) = \alpha_1 y_1(t) + \dots + \alpha_n y_n(t),$$

wobei $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ beliebig wählbare Konstanten sind.

Deshalb verfolgen wir jetzt das Ziel, ein Fundamentalsystem zu finden. Wir haben bereits gesehen: wenn λ ein Eigenwert von A ist mit dem Eigenvektor c , so ist $ce^{\lambda t}$ eine Lösung.

Satz 13.4 (Finden von Fundamentalsystemen). (a) Sind $\lambda_1 \neq \lambda_2$ Eigenwerte von A , so sind die zugehörigen Eigenvektoren linear unabhängig. Allgemeiner: hat A n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$ mit Eigenvektoren c_1, \dots, c_n , so ist

$$\{c_1 e^{\lambda_1 t}, c_2 e^{\lambda_2 t}, \dots, c_n e^{\lambda_n t}\}$$

ein Fundamentalsystem zum Differentialgleichungssystem $y' = Ay$.

(b) Gibt es zu einem Eigenwert λ von A mehrere linear unabhängige Eigenvektoren c_1, \dots, c_k , so sind $y_1(t) = c_1 e^{\lambda t}, \dots, y_k(t) = c_k e^{\lambda t}$ linear unabhängige Lösungen.

(c) Ist λ ein Eigenwert von A mit $\lambda \notin \mathbb{R}$ und $c \in \mathbb{C}^n$ als zugehörigem Eigenvektor, so sind $y_1(t) = \Re(ce^{\lambda t})$ und $y_2(t) = \Im(ce^{\lambda t})$ linear unabhängige reelle Lösungen von $y' = Ay$.

Beispiel 13.5. Bei uns ist

$$A = \begin{pmatrix} -k_1 & 0 & 0 \\ k_2 & -k_2 & k_3 \\ 0 & k_2 & -k_3 \end{pmatrix}.$$

Damit wir etwas konkreter rechnen können, nehmen wir $k_1 = 1, k_2 = 2$ und $k_3 = 3$ an, also

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

Wir bestimmen die Eigenwerte von A :

$$\begin{aligned} \det(A - \lambda I_3) &= \begin{vmatrix} -1 - \lambda & 0 & 0 \\ 1 & -2 - \lambda & 3 \\ 0 & 2 & -3 - \lambda \end{vmatrix} = (-1 - \lambda) \begin{vmatrix} -2 - \lambda & 3 \\ 2 & -3 - \lambda \end{vmatrix} \\ &= -(1 + \lambda) \begin{vmatrix} -2 - \lambda & 3 \\ -\lambda & -\lambda \end{vmatrix} = (1 + \lambda) \lambda \begin{vmatrix} -2 - \lambda & 3 \\ 1 & 1 \end{vmatrix} \\ &= (1 + \lambda) \lambda \begin{vmatrix} -5 - \lambda & 0 \\ 1 & 1 \end{vmatrix} = -(1 + \lambda) \lambda (5 + \lambda). \end{aligned}$$

Also besitzt A die Eigenwerte $\lambda_1 = 0$, $\lambda_2 = -1$ und $\lambda_3 = -5$.

Wir bestimmen die Eigenvektoren zu $\lambda_1 = 0$:

$$\left(\begin{array}{ccc|c} -1 & 0 & 0 & 0 \\ 1 & -2 & 3 & 0 \\ 0 & 2 & -3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 0 \end{array} \right),$$

woraus wir $x_1 = 0$ und $3x_3 = 2x_2$ gewinnen, also $c_1 = (0, 3, 2)^\top$.

Die Bestimmung der Eigenvektoren zu $\lambda_2 = -1$ geschieht wie folgt:

$$\left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & 0 \\ 0 & 2 & -2 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & 0 \\ 0 & 1 & -1 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \end{array} \right).$$

Wenn wir $x_3 = t$ setzen, dann ist $x_1 = -2t$ und $x_2 = +t$, also ist $c_2 = (-2, 1, 1)^\top$.

Und ein Eigenvektor c_3 zu λ_3 bestimmt sich zu $c_3 = (0, 1, -1)^\top$.

Nach Satz 13.4 erhalten wir dann das Fundamentalsystem

$$\{c_1 e^{\lambda_1 t}, c_2 e^{\lambda_2 t}, c_3 e^{\lambda_3 t}\} = \left\{ \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} e^{-t}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} e^{-5t} \right\}.$$

Die allgemeine Lösung ist dann (mit beliebig wählbaren Parametern $\alpha_1, \alpha_2, \alpha_3$)

$$y(t) = \alpha_1 \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} e^{-t} + \alpha_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} e^{-5t}.$$

Zur Bestimmung von $\alpha_1, \alpha_2, \alpha_3$ erinnern wir an die Anfangsbedingungen $c_A(0) = a$, $c_B(0) = 0$, $c_C(0) = 0$:

$$\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} = y(0) = \begin{pmatrix} c_A(0) \\ c_B(0) \\ c_C(0) \end{pmatrix} = \alpha_1 \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

also erhalten wir das System

$$\begin{aligned} -2\alpha_2 &= a, \\ 3\alpha_1 - \alpha_2 + \alpha_3 &= 0, \\ 2\alpha_1 + \alpha_2 - \alpha_3 &= 0. \end{aligned}$$

Dieses System lösen wir und erhalten $\alpha_1 = \frac{a}{5}$, $\alpha_2 = -\frac{a}{2}$, $\alpha_3 = -\frac{a}{10}$, sodaß wir endlich die Lösung finden:

$$\begin{pmatrix} c_A(t) \\ c_B(t) \\ c_C(t) \end{pmatrix} = y(t) = \frac{a}{5} \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} - \frac{a}{2} e^{-t} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} - \frac{a}{10} e^{-5t} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Es folgt insgesamt

$$\begin{aligned} c_A(t) &= a e^{-t}, \\ c_B(t) &= a \left(\frac{3}{5} - \frac{1}{2} e^{-t} - \frac{1}{10} e^{-5t} \right), \\ c_C(t) &= a \left(\frac{2}{5} - \frac{1}{2} e^{-t} + \frac{1}{10} e^{-5t} \right). \end{aligned}$$

Die Substanz A verschwindet exponentiell schnell, und wir haben das Langzeitverhalten

$$\lim_{t \rightarrow \infty} c_A(t) = 0, \quad \lim_{t \rightarrow \infty} c_B(t) = \frac{3}{5}a, \quad \lim_{t \rightarrow \infty} c_C(t) = \frac{2}{5}a.$$

14 Der harmonische Oszillator

Wir betrachten einen Federschwinger. Die sich bewegende Masse sei m ; die Federkonstante sei c ; die Auslenkung aus der Ruhelage zur Zeit t sei $x(t)$. Dann ist $x'(t)$ die Geschwindigkeit der Masse zur Zeit t , und $x''(t)$ ist die Beschleunigung der Masse zur Zeit t .

Je nachdem, welche physikalischen Effekte wir berücksichtigen oder vernachlässigen, erhalten wir dann aus Newtons drittem Axiom eine der Differentialgleichungen

$$\begin{aligned} mx''(t) &= -cx(t) && \text{ungedämpfter Fall (harmonischer Oszillator),} \\ mx''(t) &= -cx(t) - dx'(t) && \text{Dämpfung proportional zur Geschwindigkeit,} \\ mx''(t) &= -cx(t) - dx'(t) - F_{ext}(t) && \text{zusätzlich mit äußerer Kraft.} \end{aligned}$$

Es entsteht also eine Differentialgleichung vom Typ

$$y''(t) + ay'(t) + by(t) = f(t),$$

die wir *lineare Differentialgleichung zweiter Ordnung* nennen wollen. Solche Differentialgleichungen sind sehr häufig in den Anwendungen; sie treten z.B. ebenfalls auf bei der Beschreibung von Wechselströmen in elektrischen Schwingkreisen, die aus Spule, Kondensator und ohmschem Widerstand bestehen.

Die zugehörige *homogene Gleichung* lautet

$$y''(t) + ay'(t) + by(t) = 0.$$

Für die homogene Gleichung machen wir den Lösungsansatz $y(t) = e^{\lambda t}$ mit einem noch nicht bestimmten $\lambda \in \mathbb{C}$. Der Hintergedanke bei diesem Ansatz ist, dass dann y'' , y' und y alle „vom gleichen Typ“ sind.

Es gilt $y''(t) = \lambda^2 e^{\lambda t}$ und $y'(t) = \lambda e^{\lambda t}$, also ergibt sich im Falle der homogenen Gleichung

$$0 \stackrel{!}{=} y''(t) + ay'(t) + by(t) = \lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + be^{\lambda t},$$

wegen $e^{\lambda t} \neq 0$ für alle $t \in \mathbb{R}$ also

$$\lambda^2 + a\lambda + b \stackrel{!}{=} 0$$

als Bestimmungsgleichung für $\lambda \in \mathbb{C}$. Es folgt demnach

$$\lambda_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}.$$

Wir setzen $D := \frac{a^2}{4} - b$ zur Abkürzung.

Wir beginnen mit dem **Spezialfall** $a = 0$ und $b = \frac{c}{m}$, also dem harmonischen Oszillator. Unser Ansatz führt uns auf

$$y(t) = e^{\pm i\sqrt{\frac{c}{m}}t} = \cos\left(t\sqrt{\frac{c}{m}}\right) \pm i \sin\left(t\sqrt{\frac{c}{m}}\right).$$

Die reellen Lösungen sind

$$\cos\left(t\sqrt{\frac{c}{m}}\right), \quad \sin\left(t\sqrt{\frac{c}{m}}\right),$$

und die allgemeine Lösung ist

$$y(t) = c_1 \cos\left(t\sqrt{\frac{c}{m}}\right) + c_2 \sin\left(t\sqrt{\frac{c}{m}}\right).$$

Wir haben es also mit einer Schwingung zu tun, deren Kreisfrequenz $\sqrt{\frac{c}{m}}$ beträgt. Die Konstanten c_1 und c_2 bestimmen sich aus den Anfangsbedingungen. Wenn z.B. $y(0) = 1$ und $y'(0) = 0$ sein sollte, dann folgt

$$1 = c_1 \cos(0) + c_2 \sin(0) = c_1,$$

$$0 = -c_1 \sqrt{\frac{c}{m}} \sin\left(0 \cdot \sqrt{\frac{c}{m}}\right) + c_2 \sqrt{\frac{c}{m}} \cos\left(0 \cdot \sqrt{\frac{c}{m}}\right) = c_2 \sqrt{\frac{c}{m}},$$

also $y(t) = \cos\left(t\sqrt{\frac{c}{m}}\right)$.

Wir kehren zurück zur **allgemeinen Gleichung** $y''(t) + ay'(t) + by(t) = 0$. Aus physikalischen Gründen sind $b > 0$ und $a \geq 0$, denn $b = \frac{c}{m}$ ist die Federkonstante pro Masse, und $a = \frac{d}{m}$ ist die Dämpfung pro Masse.

Fall 1: sei $D > 0$. Das heißt $\frac{a^2}{4} > b$. Dann folgt

$$\lambda_1 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b} \in \mathbb{R}, \quad \lambda_2 = -\frac{a}{2} - \sqrt{\frac{a^2}{4} - b} \in \mathbb{R},$$

und die allgemeine Lösung ist $y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$. Wegen $a \geq 0$ und $b > 0$ gilt $\lambda_2 < 0$ und

$$\lambda_1 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b} < \frac{a}{2} + \sqrt{\frac{a^2}{4}} = 0,$$

also klingt $y(t)$ exponentiell ab für $t \rightarrow \infty$. Dieser Fall wird *Kriechfall* genannt.

Fall 2: Sei $D = 0$. Das heißt $\frac{a^2}{4} = b$, also $\lambda_1 = \lambda_2 = -\frac{a}{2}$. Die allgemeine Lösung ist dann

$$y(t) = c_1 e^{-\frac{a}{2}t} + c_2 t e^{-\frac{a}{2}t},$$

mit wählbaren Konstanten $c_1, c_2 \in \mathbb{R}$. Dieser Fall heißt *aperiodischer Grenzfall*.

Fall 3: Sei $D < 0$. Das heißt $\frac{a^2}{4} < b$. Dann ist

$$\lambda_1 = -\frac{a}{2} + i\sqrt{b - \frac{a^2}{4}}, \quad \lambda_2 = -\frac{a}{2} - i\sqrt{b - \frac{a^2}{4}},$$

und die allgemeine Lösung lautet

$$y(t) = e^{-\frac{a}{2}t} \left(c_1 \cos(\omega t) + c_2 \sin(\omega t) \right),$$

wobei $c_1, c_2 \in \mathbb{R}$ frei wählbar sind, und wir haben definiert:

$$\omega := \sqrt{-D} = \sqrt{b - \frac{a^2}{4}}.$$

Dies ist der Schwingungsfall, und wir haben eine gedämpfte Schwingung vorliegen.

Kapitel 5

Integration von Funktionen mehrerer Variabler

15 Zweidimensionale Integrale

Aus der HM I kennen wir Integrale der Form $\int_a^b f(x) dx$, die wir interpretieren als „Fläche unter dem Graphen“.

Jetzt wollen wir Funktionen $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ betrachten und eine Menge $G \subseteq \mathbb{R}^2$, und wir wollen definieren, was

$$\int_G f(x, y) d(x, y)$$

sein soll. Unser Ziel dabei wird sein, dieses Integral auf eindimensionale Integrale zurückzuführen.

Aber zuerst müssen wir beschreiben, welche Eigenschaften die Menge G haben soll.

Definition 15.1. Sei $G \subseteq \mathbb{R}^2$.

(a) Die Menge G heißt y -projizierbar, wenn es Funktionen $\underline{y}, \bar{y}: [a, b] \rightarrow \mathbb{R}$ gibt mit

$$G = \{(x, y) \in \mathbb{R}^2 : x \in [a, b] \text{ und } \underline{y}(x) \leq y \leq \bar{y}(x)\}.$$

(b) Die Menge G heißt x -projizierbar, wenn es Funktionen $\underline{x}, \bar{x}: [c, d] \rightarrow \mathbb{R}$ gibt mit

$$G = \{(x, y) \in \mathbb{R}^2 : y \in [c, d] \text{ und } \underline{x}(y) \leq x \leq \bar{x}(y)\}.$$

(c) Die Menge G heißt Standardmenge, wenn (a) und (b) gelten.

Die **Idee** hinter der Definition von $\int_G f(x, y) d(x, y)$ ist es, zunächst für jedes $x \in [a, b]$ ein Integral

$$\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) dy$$

auszurechnen, und anschließend

$$\int_a^b \left(\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) dy \right) dx$$

auszuwerten. Natürlich kann man stattdessen auch die Rollen von x und y tauschen:

$$\int_c^d \left(\int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) dx \right) dy.$$

Definition 15.2. Sei $G \subseteq \mathbb{R}^2$ x -projizierbar oder y -projizierbar, und sei f stetig auf G . Dann setzen wir

$$\int_G f(x, y) \, d(x, y) = \int_a^b \left(\int_{\underline{y}(x)}^{\overline{y}(x)} f(x, y) \, dy \right) dx, \quad \text{falls } G \text{ } y\text{-projizierbar ist,}$$

$$\int_G f(x, y) \, d(x, y) = \int_c^d \left(\int_{\underline{x}(y)}^{\overline{x}(y)} f(x, y) \, dx \right) dy, \quad \text{falls } G \text{ } x\text{-projizierbar ist.}$$

Falls G eine Standardmenge ist, dann führen beide Gleichungen zum selben Integralwert.

Beispiel 15.3. (a) Sei $G = [0, 1] \times [0, 1]$ das Einheitsquadrat und $f(x, y) = x^3 - y^3 - x^2y + 2$. Dann folgt

$$\begin{aligned} \int_G f(x, y) \, d(x, y) &= \int_0^1 \left(\int_0^1 (x^3 - y^3 - x^2y + 2) \, dy \right) dx \\ &= \int_0^1 \left(x^3y - \frac{1}{4}y^4 - \frac{1}{2}x^2y^2 + 2y \right) \Big|_{y=0}^{y=1} dx = \int_0^1 \left(x^3 - \frac{1}{4} - \frac{1}{2}x^2 + 2 \right) dx \\ &= \int_0^1 \left(x^3 - \frac{1}{2}x^2 + \frac{7}{4} \right) dx = \left(\frac{1}{4}x^4 - \frac{1}{6}x^3 + \frac{7}{4}x \right) \Big|_{x=0}^{x=1} = \frac{1}{4} - \frac{1}{6} + \frac{7}{4} = 2 - \frac{1}{6} = \frac{11}{6}. \end{aligned}$$

(b) Sei $f(x, y) = x^2 + y^2$ und G sei ein Flächenstück oberhalb einer Parabel, d.h.

$$G = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 2, \quad x^2 \leq y \leq 4\}.$$

Dann rechnen wir:

$$\begin{aligned} \int_G f(x, y) \, d(x, y) &= \int_0^2 \left(\int_{x^2}^4 (x^2 + y^2) \, dy \right) dx = \int_0^2 \left(x^2y + \frac{1}{3}y^3 \right) \Big|_{y=x^2}^{y=4} dx \\ &= \int_0^2 \left(4x^2 + \frac{1}{3}4^3 - x^4 - \frac{1}{3}x^6 \right) dx = \left(\frac{4}{3}x^3 + \frac{64}{3}x - \frac{1}{5}x^5 - \frac{1}{21}x^7 \right) \Big|_{x=0}^2 \\ &= \frac{4 \cdot 8}{3} + \frac{128}{3} - \frac{32}{5} - \frac{128}{21} = \dots = \frac{4288}{105}. \end{aligned}$$

Verfahren 15.4 (Substitutionsregel). Wir betrachten ein Integral $\int_G f(x, y) \, d(x, y)$ und setzen voraus, dass für eine geeignete Funktion Φ gilt, dass $\binom{x}{y} = \Phi(u, v)$, wobei $\Phi: H \rightarrow G$ bijektiv ist. Wir haben also $x = \Phi_1(u, v)$ und $y = \Phi_2(u, v)$. Dann gilt

$$\int_G f(x, y) \, d(x, y) = \int_H f(\Phi(u, v)) |\det J_\Phi(u, v)| \, d(u, v),$$

wobei J_Φ die Jacobi-Matrix von Φ ist.

Beispiel 15.5. Sei G der Kreisring im \mathbb{R}^2 um den Ursprung mit innerem Radius 1 und äußerem Radius 4. Wir fragen nach dem Wert des Integrals $\int_G e^{x^2+y^2} \, d(x, y)$.

Unser **erster Versuch** besteht in einer direkten Berechnung des Integrals, so wie wir es bisher mehrfach getan hatten. Die Menge G ist leider weder x -projizierbar noch y -projizierbar, weshalb wir sie entlang der x -Achse zerschneiden:

$$G = G_1 \cup G_2,$$

$$G_1 = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4, \quad x \geq 0\}, \quad G_2 = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4, \quad x \leq 0\},$$

und jetzt sind G_1 und G_2 y -projizierbar. Wir haben dann

$$\begin{aligned} \int_G e^{x^2+y^2} \, d(x, y) &= \int_{G_1} e^{x^2+y^2} \, d(x, y) + \int_{G_2} e^{x^2+y^2} \, d(x, y), \\ \int_{G_1} e^{x^2+y^2} \, d(x, y) &= \int_4^{-1} \left(\int_0^{\sqrt{16-x^2}} e^{x^2+y^2} \, dy \right) dx + \int_{-1}^1 \left(\int_{\sqrt{1-x^2}}^{\sqrt{16-x^2}} e^{x^2+y^2} \, dy \right) dx \\ &\quad + \int_1^4 \left(\int_0^{\sqrt{16-x^2}} e^{x^2+y^2} \, dy \right) dx, \end{aligned}$$

und jetzt stecken wir fest, weil wir die inneren Integrale nicht ausrechnen können, da wir für $e^{x^2+y^2}$ keine Stammfunktion finden.

Der **zweite Versuch** verwendet die Substitution. Wir erinnern an die Polarkoordinaten (r, φ) der Ebene mit $x = r \cos \varphi$ und $y = r \sin \varphi$. Es ist also

$$\begin{pmatrix} x \\ y \end{pmatrix} = \Phi(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.$$

Dann ist $x^2 + y^2 = r^2$ sowie $(x, y) = \Phi(r, \varphi) \in G$ genau dann, wenn $1 \leq \sqrt{x^2 + y^2} \leq 4$, also $1 \leq r \leq 4$. Wir bekommen somit

$$H = \{(r, \varphi) \in (0, \infty) \times (-\pi, \pi] : 1 \leq r \leq 4\}.$$

Somit ist nach Substitutionsregel

$$\int_G e^{x^2+y^2} d(x, y) = \int_H e^{r^2} |\det J_\Phi(r, \varphi)| d(r, \varphi).$$

Wir benötigen jetzt die Jacobi-Matrix von Φ :

$$J_\Phi(r, \varphi) = \begin{pmatrix} \partial_r x(r, \varphi) & \partial_\varphi x(r, \varphi) \\ \partial_r y(r, \varphi) & \partial_\varphi y(r, \varphi) \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Also ist $\det J_\Phi(r, \varphi) = r \cos^2 \varphi + r \sin^2 \varphi = r$, und wir finden dann insgesamt

$$\begin{aligned} \int_G e^{x^2+y^2} d(x, y) &= \int_H e^{r^2} r d(r, \varphi) = \int_{-\pi}^{\pi} \left(\int_1^4 e^{r^2} r dr \right) d\varphi = \int_1^4 \left(\int_{-\pi}^{\pi} e^{r^2} r d\varphi \right) dr \\ &= \int_1^4 2\pi e^{r^2} r dr = 2\pi \cdot \frac{1}{2} e^{r^2} \Big|_{r=1}^{r=4} = \pi (e^4 - e). \end{aligned}$$

Beispiel 15.6. Wir betrachten einen Viertelkreis mit Radius 1 im ersten Quadranten,

$$G = \{(x, y) \in \mathbb{R}^2 : x \geq 0, \quad y \geq 0, \quad x^2 + y^2 \leq 1\}$$

und fragen nach $\int_G xy d(x, y)$.

In Polarkoordinaten ist G gegeben durch die Bedingungen $0 \leq r \leq 1$ und $0 \leq \varphi \leq \pi/2$. Mit der Substitutionsregel bekommen wir dann

$$\begin{aligned} \int_G xy d(x, y) &= \int_0^{\pi/2} \int_0^1 r \cos \varphi r \sin \varphi \cdot r d(r, \varphi) = \int_0^{\pi/2} \cos \varphi \sin \varphi d\varphi \cdot \int_0^1 r^3 dr \\ &= \frac{1}{2} \sin^2 \varphi \Big|_{\varphi=0}^{\varphi=\pi/2} \cdot \frac{1}{4} r^4 \Big|_{r=0}^{r=\pi/2} = \left(\frac{1}{2} \sin^2(\pi/2) - \frac{1}{2} \sin^2(0) \right) \cdot \frac{1}{4} = \frac{1}{8}. \end{aligned}$$

16 Dreidimensionale Integrale

Wir betrachten ein **Beispielproblem**: Sei K ein Quader, Zylinder, Kegel o.ä. mit nicht-konstanter Maschendichte $\varrho(x, y, z): K \rightarrow \mathbb{R}$. Dann ist die Gesamtmasse von K gleich

$$\int_K \varrho(x, y, z) d(x, y, z).$$

Beispiel 16.1. Sei $K = [0, 2] \times [0, 1] \times [0, 1]$ ein Quader. Dann ist die Gesamtmasse gleich

$$\int_K \varrho(x, y, z) d(x, y, z) = \int_0^2 \int_0^1 \int_0^1 \varrho(x, y, z) dz dy dx.$$

Wir verfahren genauso, wenn K eine projizierbare Menge ist, d.h. zum Beispiel

$$K = \{(x, y, z) \in \mathbb{R}^3 : a \leq z \leq b, \quad \underline{y}(z) \leq y \leq \bar{y}(z), \quad \underline{x}(y, z) \leq x \leq \bar{x}(y, z)\}.$$

Dann ist

$$\int_K \varrho(x, y, z) d(x, y, z) = \int_a^b \left(\int_{\underline{y}(z)}^{\bar{y}(z)} \left(\int_{\underline{x}(y,z)}^{\bar{x}(y,z)} \varrho(x, y, z) dx \right) dy \right) dz.$$

Beispiel 16.2. Wir betrachten die Pyramide K mit der quadratischen Grundfläche, die gebildet wird von den vier Eckpunkten $(2, 2, 0)$, $(2, -2, 0)$, $(-2, 2, 0)$, $(-2, -2, 0)$; und die Höhe der Pyramide sei 2. Die Dichtefunktion der Pyramide sei $\rho(x, y, z) = z$. Es ist also

$$K = \{(x, y, z) \in \mathbb{R}^3: 0 \leq z \leq 2, \quad -2 + z \leq y \leq 2 - z, \quad -2 + z \leq x \leq 2 - z\},$$

und wir suchen $\int_K z \, d(x, y, z)$. Die Rechnung ist dann:

$$\begin{aligned} \int_K z \, d(x, y, z) &= \int_0^2 \left(\int_{-2+z}^{2-z} \left(\int_{-2+z}^{2-z} z \, dx \right) dy \right) dz = \int_0^2 \left(\int_{-2+z}^{2-z} \left(zx \Big|_{x=-2+z}^{x=2-z} \right) dy \right) dz \\ &= \int_0^2 \left(\int_{-2+z}^{2-z} (z(2-z+2-z)) \, dy \right) dz = \int_0^2 \left(\int_{-2+z}^{2-z} (z \cdot (4-2z)) \, dy \right) dz \\ &= \int_0^2 \left(z \cdot (4-2z) y \Big|_{y=-2+z}^{2-z} \right) dz = \int_0^2 (z \cdot (4-2z) \cdot (4-2z)) \, dz \\ &= \int_0^2 z \cdot (16 - 16z + 4z^2) \, dz = \int_0^2 (4z^3 - 16z^2 + 16z) \, dz \\ &= \left(z^4 - \frac{16}{3}z^3 + 8z^2 \right) \Big|_{z=0}^2 = 16 - \frac{16}{3} \cdot 8 + 8 \cdot 4 = 48 - \frac{128}{3} = \frac{144}{3} - \frac{128}{3} = \frac{16}{3}. \end{aligned}$$

Beispiel 16.3. Wir betrachten einen vertikal stehenden Zylinder mit Durchmesser 2 und Höhe 8:

$$K = \{(x, y, z) \in \mathbb{R}^3: -4 \leq z \leq 4, \quad x^2 + y^2 \leq 1\}$$

und setzen $\rho(x, y, z) = x^2 y^2$. Gesucht ist das Integral $\int_K \rho(x, y, z) \, d(x, y, z)$. Die Symmetrie des Zylinders suggeriert die Einführung von **Zylinderkoordinaten**, analog zu den Polarkoordinaten der Ebene:

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z.$$

Weiterhin geben wir der Querschnittsfläche des Zylinders einen Namen,

$$S = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\},$$

und wir rechnen wie folgt:

$$\begin{aligned} \int_K x^2 y^2 \, d(x, y, z) &= \int_{-4}^4 \int_S x^2 z^2 \, d(x, y) \, dz = \int_{-4}^4 \int_{-\pi}^{\pi} \int_0^1 (r \cos \varphi)^2 z^2 \cdot r \, dr \, d\varphi \, dz \\ &= \int_{-4}^4 z^2 \, dz \cdot \int_{-\pi}^{\pi} \cos^2(\varphi) \, d\varphi \cdot \int_0^1 r^3 \, dr = \frac{1}{3} z^3 \Big|_{z=-4}^{z=4} \cdot \left(\frac{\varphi}{2} + \frac{1}{4} \sin(2\varphi) \right) \Big|_{\varphi=-\pi}^{\varphi=\pi} \cdot \frac{1}{4} r^4 \Big|_{r=0}^1 \\ &= \frac{2}{3} \cdot 64 \cdot \pi \cdot \frac{1}{4} = \frac{32}{3} \pi. \end{aligned}$$

Bemerkung 16.4 (Kugelkoordinaten). Bei kugelsymmetrischen Körpern empfiehlt sich die Verwendung von Kugelkoordinaten, siehe Abbildung XXXXXXXXXXXX.

Die Koordinaten sind

$$\begin{array}{ll} r \in (0, \infty) & (\text{Radius}), \\ \varphi \in (-\pi, \pi) & (\text{geographische Länge}), \\ \theta \in (0, \pi) & (\text{geographische Breite}). \end{array}$$

Hierbei entspricht $\theta = 0$ dem Nordpol und $\theta = \pi/2$ dem Äquator. Die Umrechnung läuft über die Formeln

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta,$$

und wir haben $r^2 = x^2 + y^2 + z^2$. Das Volumendifferential bestimmt sich zu

$$d(x, y, z) = r^2 \sin \theta \, d(r, \varphi, \theta).$$

Beispiel 16.5. Sei K die Einheitskugel,

$$K = \{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 + z^2 \leq 1\},$$

und sei

$$\varrho(x, y, z) = e^{(x^2+y^2+z^2)^{3/2}}.$$

Dann folgt

$$\begin{aligned} \int_K \varrho(x, y, z) \, d(x, y, z) &= \int_{\{0 \leq r \leq 1\}} e^{(r^2)^{3/2}} r^2 \sin \theta \, d(r, \varphi, \theta) = \int_0^1 \int_{-\pi}^{\pi} \int_0^{\pi} e^{r^3} r^2 \sin \theta \, d\theta \, d\varphi \, dr \\ &= \int_0^1 e^{r^3} r^2 \, dr \cdot \int_0^{\pi} \sin \theta \, d\theta \cdot \int_{-\pi}^{\pi} d\varphi = \frac{1}{3} e^{r^3} \Big|_0^1 \cdot (-\cos \theta) \Big|_0^{\pi} \cdot 2\pi \\ &= \left(\frac{1}{3} e - \frac{1}{3} \right) \cdot (-(-1) - (-1)) \cdot 2\pi = \frac{4\pi}{3} (e - 1). \end{aligned}$$

Bemerkung 16.6 (Einige Formeln aus der Physik). Sei $K \subset \mathbb{R}^3$ eine (gutartige) Menge, $\varrho: K \rightarrow \mathbb{R}$ eine Dichteverteilung. Dann sind

$$V = \int_K d(x, y, z) \quad \text{das Volumen von } K,$$

$$M = \int_K \varrho(x, y, z) \, d(x, y, z) \quad \text{die Masse von } K.$$

Der Punkt $(x_S, y_S, z_S) \in \mathbb{R}^3$ mit

$$x_S := \frac{1}{M} \int_K x \varrho(x, y, z) \, d(x, y, z),$$

$$y_S := \frac{1}{M} \int_K y \varrho(x, y, z) \, d(x, y, z),$$

$$z_S := \frac{1}{M} \int_K z \varrho(x, y, z) \, d(x, y, z)$$

ist der Schwerpunkt von K .

Beispiel 16.7. Wir bestimmen den Schwerpunkt des Kugeloktanten mit Radius R

$$K = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, x^2 + y^2 + z^2 \leq R^2\},$$

wobei wir eine konstante Dichte ϱ voraussetzen.

Die Masse M ist

$$\begin{aligned} M &= \int_K \varrho \, d(x, y, z) = \varrho \int_0^R \int_0^{\pi/2} \int_0^{\pi/2} r^2 \sin \theta \, d\theta \, d\varphi \, dr \\ &= \varrho \int_0^R r^2 \, dr \cdot \int_0^{\pi/2} \sin \theta \, d\theta \cdot \int_0^{\pi/2} d\varphi = \frac{1}{3} R^3 \varrho \cdot \frac{\pi}{2} \cdot (-\cos \theta) \Big|_0^{\pi/2} \\ &= \frac{1}{6} \pi \varrho R^3 \cdot (-0 + 1) = \frac{1}{6} \pi \varrho R^3. \end{aligned}$$

Jetzt ermitteln wir die Koordinaten des Schwerpunkts, beginnend mit x_S :

$$\begin{aligned} x_S &= \frac{1}{M} \int_K x \varrho \, d(x, y, z) = \frac{\varrho}{M} \int_0^R \int_0^{\pi/2} \int_0^{\pi/2} r \sin \theta \cos \varphi \cdot r^2 \sin \theta \, d\theta \, d\varphi \, dr \\ &= \frac{\varrho}{M} \int_0^R r^3 \, dr \cdot \int_0^{\pi/2} \cos \varphi \, d\varphi \cdot \int_0^{\pi/2} \sin^2 \theta \, d\theta \\ &= \frac{\varrho}{M} \cdot \frac{1}{4} R^4 \cdot \sin \varphi \Big|_0^{\pi/2} \cdot \left(\frac{\theta}{2} - \frac{1}{2} \sin \theta \cos \theta \right) \Big|_0^{\pi/2} = \frac{\varrho}{M} \cdot \frac{R^4}{4} \cdot 1 \cdot \frac{\pi}{4} = \frac{\varrho R^4 \pi \cdot 6}{16 \pi \varrho R^3} = \frac{3R}{8}. \end{aligned}$$

Genauso ist $y_S = z_S = \frac{3R}{8}$.

17 Kurvenintegrale

Gegeben sei eine „geschwungene Linie im \mathbb{R}^2 (oder \mathbb{R}^3)“, wofür wir ab jetzt das Wort „Kurve“ verwenden wollen und eine stetige Funktion $\varrho: \mathbb{R}^2 \rightarrow \mathbb{R}$. Wir stellen uns die Linie als einen Faden vor, der mit Masse belegt ist, und diese Masse hat die Dichte $\varrho(x, y)$. Wir fragen uns nach der Gesamtmasse des Fadens.

Wir brauchen dafür ein Verfahren, das uns die „Kurve“ rechnerisch beschreibt.

Definition 17.1. Sei $[a, b] \subset \mathbb{R}$ ein Intervall. Eine stetig differenzierbare Funktion $\gamma: [a, b] \rightarrow \mathbb{R}^2$ heißt Weg im \mathbb{R}^2 .

Beispiel 17.2. (a) Ein kleines Parabelstück wird beschrieben durch $\gamma: [0, 1] \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (t, t^2)^\top$.

(b) Es ist möglich, daß verschiedene Wege γ dieselbe Figur im \mathbb{R}^2 erzeugen. Zum Beispiel wird der rechte Halbkreis (im Uhrzeigersinn durchlaufen von 12 Uhr bis 6 Uhr) beschrieben durch den Weg

$$\gamma: [0, \pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} \sin t \\ \cos t \end{pmatrix},$$

aber er wird auch beschrieben durch den Weg

$$\tilde{\gamma}: [-1, 1] \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(t) = \begin{pmatrix} \sqrt{1-t^2} \\ -t \end{pmatrix}.$$

All diejenigen Wege, welche dieselbe Figur erzeugen, nennen wir *Kurve*. Für die Durchführung der Rechnung (also beim Berechnen der Kurvenintegrale) ist es dann egal, welchen Weg γ man nimmt zur Beschreibung der Linie, denn es kommt immer dasselbe Ergebnis heraus.

Definition 17.3 (Kurvenintegral erster Art). Sei $\gamma: [a, b] \rightarrow \mathbb{R}^2$ ein Weg und $\varrho: \mathbb{R}^2 \rightarrow \mathbb{R}$ eine stetige Funktion. Dann definieren wir

$$\int_{\gamma} \varrho(x) \, ds := \int_{t=a}^b \varrho(\gamma(t)) |\gamma'(t)| \, dt$$

und nennen diesen Ausdruck Kurvenintegral (erster Art) von ϱ entlang γ . Hierbei ist $|\gamma'(t)|$ die pythagoräische Länge des Vektors $\gamma'(t)$.

Beispiel 17.4. Sei $\gamma: [0, 1] \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (t, t^2)^\top$ und $\varrho(x, y) = x^2 + y$. Dann rechnen wir wie folgt:

$$\begin{aligned} \int_{\gamma} \varrho(x, y) \, ds &= \int_{t=0}^1 \varrho(t, t^2) |\gamma'(t)| \, dt = \int_{t=0}^1 (t^2 + t^2) \left| \begin{pmatrix} 1 \\ 2t \end{pmatrix} \right| \, dt = \int_{t=0}^1 2t^2 \sqrt{1 + (2t)^2} \, dt \\ &= \left(\frac{t}{32} \sqrt{1 + 4t^2} (1 + 8t^2) - \frac{1}{64} \ln(t + \frac{1}{2} \sqrt{1 + 4t^2}) \right) \Big|_{t=0}^{t=1} \\ &= \dots = \frac{9}{32} \sqrt{5} - \frac{1}{64} \ln(2 + \sqrt{5}). \end{aligned}$$

Es gibt noch eine zweite Art von Kurvenintegralen mit einer anderen physikalischen Interpretation: ist f ein Kraftfeld im Raum (zum Beispiel das Gravitationsfeld der Erde, oder ein elektrostatisches Feld) und γ ein Weg im \mathbb{R}^3 , so können wir die Frage aufwerfen, wieviel Energie es kostet, einen Körper entlang des Weges γ zu ziehen und dabei gegen die Kraft f zu arbeiten.

Definition 17.5. Sei $\gamma: [a, b] \rightarrow \mathbb{R}^3$ ein Weg und $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ sei stetig. Dann definieren wir

$$\int_{\gamma} f(x) \cdot dx := \int_{t=a}^b f(\gamma(t)) \cdot \gamma'(t) \, dt$$

und nennen diesen Ausdruck Kurvenintegral (zweiter Art) von f entlang γ . Hierbei ist $f(\gamma(t)) \cdot \gamma'(t)$ das Skalarprodukt der beiden Vektoren $f(\gamma(t))$ und $\gamma'(t)$.

Man kann solche Kurvenintegrale natürlich auch im \mathbb{R}^2 anstatt des \mathbb{R}^3 betrachten.

Beispiel 17.6. Sei $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) = (\sin(t), \cos(t))$ und $f(x, y) = (2xy, x^2)^\top$. Dann ist

$$\begin{aligned} \int_{\gamma} f(x) \cdot dx &= \int_{t=0}^{2\pi} f(\gamma(t)) \cdot \gamma'(t) \, dt = \int_{t=0}^{2\pi} \left\langle \begin{pmatrix} 2 \sin(t) \cos(t) \\ \cos^2(t) \end{pmatrix}, \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \right\rangle \, dt \\ &= \int_{t=0}^{2\pi} (2 \sin(t) \cos^2(t) - \sin(t) \cos^2(t)) \, dt = \int_{t=0}^{2\pi} \sin(t) \cos^2(t) \, dt \\ &= -\frac{1}{3} \cos^3(t) \Big|_{t=0}^{2\pi} = -\frac{1}{3} \cos^3(2\pi) + \frac{1}{3} \cos^3(0) = -\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 = 0. \end{aligned}$$