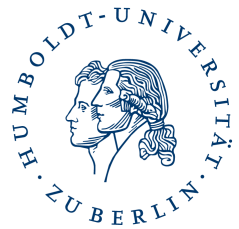


HUMBOLDT-UNIVERSITÄT ZU BERLIN



Vorlesungsskript

## **Analysis II**

**Vorlesung für Studierende der Physik,  
gelesen im Sommersemester 2019**

PD Angela Ortega

Stand: 5. Juli 2019

Die Veranstaltung „Analysis II (für Physiker\*innen)“ ist eine Service-Veranstaltung des Instituts für Mathematik für das Institut für Physik. Sie wird in der Regel von Studierenden der Physik im zweiten Bachelor-Semester besucht. Im Sommersemester 2019 wird die Vorlesung von Frau Angela Ortega gelesen. Gemeinsam mit Dennis Groh betreut sie die Übungen. Das vorliegende Dokument dient als Skript für die Veranstaltung.

# Inhaltsverzeichnis

<b>1</b>	<b>Konvergenz und Stetigkeit im Mehrdimensionalen</b>	<b>5</b>
1.1	Strukturen auf Vektorräumen . . . . .	5
1.2	Konvergenz von Folgen und Reihen . . . . .	9
1.3	Offenheit und Abgeschlossenheit von Mengen . . . . .	14
1.4	Grenzwerte von Funktionen in einem Punkt . . . . .	17
1.5	Stetige Funktionen zwischen normierten Vektorräumen . . . . .	23
1.6	Stetige Funktionen auf kompakten Mengen . . . . .	27
1.7	Zusammenhängende Mengen und Gebiete . . . . .	29
<b>2</b>	<b>Differentialrechnung in <math>\mathbb{R}^n</math></b>	<b>31</b>
2.1	Totale Differenzierbarkeit und Ableitungen . . . . .	31
2.2	Partielle Differenzierbarkeit . . . . .	33
2.3	Das Hauptkriterium für Differenzierbarkeit . . . . .	37
2.4	Rechenregeln für differenzierbare Funktionen . . . . .	38
2.5	Stetige Differenzierbarkeit . . . . .	41
2.6	Reellwertige Funktionen . . . . .	43
2.7	Lokale Extrema . . . . .	49
2.8	Der Satz von der Umkehrabbildung und der Satz über implizite Funktionen . . . . .	52
2.9	Lokale Extrema unter Nebenbedingungen . . . . .	57
<b>3</b>	<b>Integralrechnung in <math>\mathbb{R}^n</math></b>	<b>62</b>
3.1	Das Riemann-Integral auf Quadern . . . . .	63
3.2	Das Riemann-Integral von Funktionen über Jordan-messbaren Mengen . . . . .	66
3.3	Der Satz von Fubini . . . . .	68
3.4	Das Riemann-Integral stetiger Funktionen über offenen Mengen . . . . .	75
<b>4</b>	<b>Grundlagen der Vektoranalysis</b>	<b>85</b>
4.1	Kurvenintegrale . . . . .	85
4.2	Oberflächenintegrale . . . . .	95
4.3	Klassische Integralsätze . . . . .	103
	<b>Literatur</b>	<b>120</b>

# Einleitung

Diese Vorlesung beschäftigt sich mit der Differential- und Integralrechnung für mehrdimensionale Funktionen, das heißt, ganz allgemein mit Funktionen

$$f : X \rightarrow Y,$$

wobei  $X$  und  $Y$  endlichdimensionale Vektorräume sind. Diese Funktionen bilden also Vektoren auf Vektoren ab. Ein Beispiel ist durch

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

gegeben. Obwohl der Fall  $X = Y = \mathbb{R}$  ein Spezialfall ist, handelt es sich hierbei in keiner Weise um eine Wiederholung von Analysis I. Vielmehr werden in den folgenden Kapiteln Verallgemeinerungen von bisher behandelten Konzepten betrachtet. Beispiele hierfür sind

**Ableitung** Für eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  ist die Ableitung als

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

definiert.

**Hauptsatz der Differential- und Integralrechnung** Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  stetig differenzierbar. Dann gilt

$$\int_a^b f'(x) \, dx = f(b) - f(a).$$

**Kettenregel** Für differenzierbare Funktionen  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  gilt

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

# Kapitel 1

## Konvergenz und Stetigkeit im Mehrdimensionalen

### 1.1 Strukturen auf Vektorräumen

Um einen Konvergenzbegriff sinnvoll definieren zu können, benötigen wir das Verständnis, wie genau der Abstand zweier Vektoren  $x, y \in X$  in einem beliebigen (endlich-dimensionalen) Vektorraum gemessen werden kann. Hierfür führen wir den Begriff der *Norm* ein.

**Definition 1.1 (Norm):** Sei  $X$  ein endlich-dimensionaler reeller Vektorraum.<sup>1</sup> Eine Abbildung  $\|\cdot\| : X \rightarrow [0, \infty)$  heißt *Norm in  $X$* , falls folgende drei Eigenschaften gelten:

1. (Positive Definitheit) Für alle  $x \in X$  gilt: Aus  $\|x\| = 0$  folgt  $x = 0$ .
2. (Homogenität) Für alle  $x \in X$  und alle  $\lambda \in \mathbb{R}$  gilt:

$$\|\lambda x\| = |\lambda| \|x\|.$$

3. (Dreiecksungleichung) Für alle  $x, y \in X$  gilt

$$\|x + y\| \leq \|x\| + \|y\|.$$

Das Tupel  $(X, \|\cdot\|)$  heißt *normierter Vektorraum*.

**Bemerkung 1.2:** In einem normierten Vektorraum  $(X, \|\cdot\|)$  gelten:

- a) (Umgekehrte Dreiecksungleichung) Für alle  $x, y \in X$  gilt

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|.$$

- b) (Vierecksungleichung) Für alle  $v, w, x, y \in X$  gilt

$$\left| \|v - w\| - \|x - y\| \right| \leq \|v - x\| + \|w - y\|.$$

---

<sup>1</sup>Ein Vektorraum heißt *reeller Vektorraum*, wenn es ein Vektorraum über dem Körper der reellen Zahlen  $\mathbb{R}$  ist.

*Beweis.* Zunächst gilt

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|,$$

da die Norm homogen ist und die Dreiecksungleichung erfüllt. Analog gilt

$$\|y\| \leq \|y - x\| + \|x\|.$$

Aus beiden Abschätzungen folgt die umgekehrte Dreiecksungleichung.

Der Beweis der Vierecksungleichung lässt sich ebenso leicht zeigen und bleibt den Studierenden zur Übung. □

**Beispiele** Folgend angegeben sind wichtige Beispiele für Normen im Vektorraum  $\mathbb{R}^n$ . Sei hierfür  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  ein beliebiger Spaltenvektor.

**$p$ -Normen** Es sei eine natürliche Zahl  $p \in \mathbb{N} \setminus \{0\}$  gegeben. Dann ist die dazugehörige  $p$ -Norm definiert als

$$\|x\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}. \quad (1.1)$$

Die 2-Norm, also die  $p$ -Norm für  $p = 2$ , wird auch *Euklidische Norm* genannt. Sie ist von besonderer Bedeutung, da sie eng mit dem üblichen Skalarprodukt in  $\mathbb{R}^n$  zusammenhängt.

**Maximumsnorm** Die Maximumsnorm  $\|\cdot\|_\infty$  ist definiert als

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|. \quad (1.2)$$

Es lassen sich auch Normen in anderen endlich-dimensionalen Vektorräumen, beispielsweise dem Raum der Matrizen  $M(m, n)$  definieren. Der Raum ist gegeben durch

$$M(m, n) := \left\{ A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \text{ wobei } a_{ij} \in \mathbb{R}, \text{ für alle } i = 1, \dots, m, j = 1, \dots, n \right\}.$$

Über diesen Raum lässt sich beispielsweise die *Frobenius-Norm* durch

$$\|A\|_F := \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} \quad (1.3)$$

definieren.

Die Frobenius-Norm ist als Matrixnorm *verträglich* mit der Euklidischen Vektornorm. Das bedeutet, für alle Matrizen  $A \in M(m, n)$  und alle Vektoren  $x \in \mathbb{R}^n$  gilt

$$\|Ax\|_2 \leq \|A\|_F \|x\|_2.$$

Außerdem ist die Frobeniusnorm submultiplikativ, das heißt, für alle Matrizen  $A \in M(l, m)$  und  $B \in M(m, n)$  gilt

$$\|AB\|_F \leq \|A\|_F \|B\|_F.$$

Insbesondere gilt für alle regulären Matrizen  $A \in M(m, n)$ , also Matrizen mit  $\det A \neq 0$ , die Abschätzung

$$\|A^{-1}\|_F \geq \frac{1}{\|A\|_F}.$$

**Definition 1.3 (Normäquivalenz):** In einem normierten Vektorraum heißen zwei Normen  $\|\cdot\|$  und  $\|\cdot\|_X$  *äquivalent*, falls es positive Konstanten  $m, M > 0$  gibt, sodass für alle  $x \in X$

$$m\|x\| \leq \|x\|_X \leq M\|x\| \tag{1.4}$$

gilt.

**Satz 1.4:** In einem endlich-dimensionalen normierten Vektorraum sind stets alle Normen äquivalent.

*Beweis.* Der Beweis kann in Standardlehrbüchern der Analysis nachgelesen werden. Siehe dazu auch das Literaturverzeichnis am Ende des Skripts. □

**Definition 1.5 (Skalarprodukt):** Sei  $X$  ein (reeller) Vektorraum. Ein Skalarprodukt über  $X$  ist eine Abbildung  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$ , die folgende Eigenschaften erfüllt:

1. (Positive Definitheit) Für alle  $x \in X$  gilt  $\langle x, x \rangle \geq 0$ . Außerdem gilt die Äquivalenz

$$\langle x, x \rangle = 0 \iff x = 0.$$

2. (Symmetrie) Für alle  $x, y \in X$  gilt

$$\langle x, y \rangle = \langle y, x \rangle.$$

3. (Linearität) Für alle Vektoren  $x, y, z \in X$  und alle Skalare  $\lambda, \mu \in \mathbb{R}$  gilt

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle.$$

Man bemerke, dass die Linearität in der zweiten Komponente nicht gefordert werden muss, da sie automatisch aus der Symmetrie folgt.

Ein Tupel  $(X, \langle \cdot, \cdot \rangle)$  heißt *Skalarproduktraum* (auch: *Prähilbertraum*).

**Bemerkung 1.6:** In einem Skalarproduktraum  $(X, \langle \cdot, \cdot \rangle)$  gelten:

- a) (Cauchy-Schwarzsche Ungleichung) Für alle  $x, y \in X$  gilt

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle. \tag{1.5}$$

Hierbei gilt Gleichheit genau dann, wenn die Vektoren  $x$  und  $y$  linear abhängig sind.

b) (Binomische Formel) Für alle  $x, y \in X$  gilt

$$\langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle. \quad (1.6)$$

*Beweis.* Der Beweis kann in Standardlehrbüchern der Analysis nachgelesen werden. Siehe dazu auch das Literaturverzeichnis am Ende des Skripts. □

**Proposition 1.7:** Sei  $(X, \langle \cdot, \cdot \rangle)$  Skalarprodukt in  $X$ . Dann wird durch

$$\|x\| := \sqrt{\langle x, x \rangle}$$

eine Norm in  $X$  induziert. Insbesondere ist jeder Skalarproduktraum ein normierter Vektorraum.

Für diese Norm lässt sich die Cauchy-Schwarzsche Ungleichung (1.5) auch in folgender Form schreiben:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|. \quad (1.7)$$

Ebenso ist dann die binomische Formel (1.6) äquivalent zu

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2. \quad (1.8)$$

*Beweis.* Man weise die Normeigenschaften der angegebenen Abbildung als Übung nach. □

**Beispiele** Sei  $X = \mathcal{P}_2(\mathbb{R})$  der Raum der quadratischen Polynome über  $\mathbb{R}$ , also

$$\mathcal{P}_2(\mathbb{R}) := \left\{ p(x) = a_0 + a_1x + a_2x^2, \text{ wobei } a_0, a_1, a_2 \in \mathbb{R} \right\}.$$

Dann wird durch

$$\langle p, q \rangle := \int_{-1}^1 p(x)q(x) \, dx$$

ein Skalarprodukt auf  $X$  definiert.

Sei nun  $X = \mathbb{R}^n$  und  $x, y \in \mathbb{R}^n$ . Dann ist die Euklidische Norm  $\|\cdot\|_2$  die Norm, die vom Euklidischen Skalarprodukt

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i = x^T y \quad (1.9)$$

induziert wird. Dieses ist das kanonische (Standard-)Skalarprodukt in  $\mathbb{R}^n$ .

**Bemerkung 1.8:** Da in  $\mathbb{R}^n$  alle Normen äquivalent sind, ist es im Folgenden für die Definition der Konvergenz unerheblich, welche Norm wir verwenden. Wir schreiben deshalb einfach nur  $\|\cdot\|$ . Bemerke jedoch: Die Wahl der Norm ist nur bei der Frage, *ob* etwas konvergiert irrelevant, jedoch nicht bei der Frage, *wie schnell* etwas konvergiert.



## 1.2 Konvergenz von Folgen und Reihen

**Definition 1.9 (Konvergenz von Folgen und Reihen):** Sei  $(X, \|\cdot\|)$  ein beliebiger endlich-dimensionaler reeller normierter Vektorraum.

1. Eine Folge  $(x_j)_{j \in \mathbb{N}} \subset X$  heißt *konvergent*, falls ein  $x \in X$  existiert, sodass

$$\forall \varepsilon > 0 \exists j_0 \in \mathbb{N} \forall j \geq j_0 : \|x_j - x\| \leq \varepsilon. \quad (1.10)$$

Der Vektor  $x \in X$  heißt *Grenzwert* der Folge. Man sagt: Die Folge  $(x_j)_{j \in \mathbb{N}}$  konvergiert gegen  $x$  und schreibt

$$x = \lim_{j \rightarrow \infty} x_j \quad \text{oder} \quad x_j \rightarrow x \quad \text{für } j \rightarrow \infty.$$

2. Sei  $(x_j)_{j \in \mathbb{N}} \subset X$  eine Folge. Dann bezeichnet man als *Reihe* die Folge ihrer Partialsummen. Für die Reihe  $(s_n)_{n \in \mathbb{N}}$  gilt also

$$s_1 = x_1, \quad s_2 = x_1 + x_2, \quad \dots \quad s_n = \sum_{j=1}^n x_j.$$

Die Reihe  $(s_n)_{n \in \mathbb{N}}$  heißt *konvergent*, falls ein  $s$  existiert, sodass

$$s = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{j=1}^n x_j.$$

Anstelle  $s$  schreiben wir dann auch  $\sum_{j=1}^{\infty} x_j$ .

**Bemerkung 1.10:** Jede konvergente Folge  $(x_j)_{j \in \mathbb{N}}$  ist beschränkt. Es gibt also ein  $M > 0$ , sodass für alle  $j \in \mathbb{N}$

$$\|x_j\| \leq M. \quad (1.11)$$

Andernfalls ließe sich für jedes  $n \in \mathbb{N}$  ein Folgenglied  $x_{j_n}$  finden, sodass  $\|x_{j_n}\| > n$ . Das widerspricht der Definition von Konvergenz.

**Definition 1.11 (Konvergenz von Vektoren):** Sei  $X$  ein  $n$ -dimensionaler reeller Vektorraum. Wir bezeichnen die Basis von  $X$  mit  $\{b_1, \dots, b_n\}$ . Dann lässt sich jedes Element in  $X$  in der Basisdarstellung schreiben. Insbesondere gilt für jedes Element  $x_j \in X$  einer Folge  $(x_j)_{j \in \mathbb{N}}$ , dass Koordinaten  $x_{j_k} \in \mathbb{R}$  für  $k = 1, \dots, n$  existieren, sodass

$$x_j = \sum_{k=1}^n x_{j_k} b_k.$$

1. Eine Folge von Vektoren  $(x_j)_{j \in \mathbb{N}} \subset X$  konvergiert genau dann gegen einen Vektor

$$y = \sum_{k=1}^n y_k b_k, \quad y_k \in \mathbb{R},$$

wenn für jedes  $k = 1, \dots, n$  die Folge der Koordinaten  $(x_{j_k})_{j \in \mathbb{N}}$  gegen  $y_k$  konvergiert. Es gilt also

$$\lim_{j \rightarrow \infty} \sum_{k=1}^n x_{j_k} b_k = \sum_{k=1}^n \lim_{j \rightarrow \infty} x_{j_k} b_k = \sum_{k=1}^n y_k b_k = y.$$

Die Konvergenz der Vektoren ist somit äquivalent zu koordinatenweiser Konvergenz.

2. Die Vektorreihe  $(s_n)$  mit

$$s_n := \sum_{j=1}^n x_j$$

konvergiert genau dann gegen  $y \in X$ , falls für jedes  $k = 1, \dots, n$  die Reihe

$$\sum_{j=1}^n x_{j_k}$$

gegen die  $k$ -te Koordinate  $y_k$  von  $y$  konvergiert. Es gilt also

$$\sum_{j=1}^{\infty} x_j = \sum_{j=1}^{\infty} \left( \sum_{k=1}^n x_{j_k} b_k \right) = \sum_{k=1}^n \left( \sum_{j=1}^{\infty} x_{j_k} \right) b_k = \sum_{k=1}^n y_k b_k = y.$$

**Beispiele** Die Folge der Vektoren  $((\sqrt[j]{j}, 3, \frac{1}{j})^T)_{j \in \mathbb{N}}$  mit  $i \in \mathbb{N}$  konvergiert. Es gilt nämlich

$$\lim_{j \rightarrow \infty} \begin{pmatrix} \sqrt[j]{j} \\ 3 \\ \frac{1}{j} \end{pmatrix} = \begin{pmatrix} \lim_{j \rightarrow \infty} \sqrt[j]{j} \\ \lim_{j \rightarrow \infty} 3 \\ \lim_{j \rightarrow \infty} \frac{1}{j} \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.$$

Die Folge der Partialsummen der Folge  $((\frac{1}{j!}, \frac{1}{j^2})^T)_{j \in \mathbb{N}}$  konvergiert. Es gilt nämlich

$$\sum_{j=0}^{\infty} \begin{pmatrix} \frac{1}{j!} \\ \frac{1}{j^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{e} \\ \frac{\pi^2}{6} \end{pmatrix}.$$

Die Folge der Partialsummen der Folge  $((\frac{1}{2^j}, \frac{1}{j})^T)_{j \in \mathbb{N}}$  konvergiert nicht; sie divergiert. Zwar konvergiert die Folge der ersten Komponenten (Stichwort: geometrische Reihe), aber die Folge der zweiten Komponenten divergiert. In der Tat strebt die Reihe

$$\sum_{j=1}^n \frac{1}{j}$$

für  $n \rightarrow \infty$  gegen unendlich. Somit divergiert die Reihe der Vektoren.

**Lemma 1.12 (Eigenschaften des Grenzwerts):** Sei  $X$  ein normierter reeller Vektorraum und  $(x_j)_{j \in \mathbb{N}}, (y_j)_{j \in \mathbb{N}} \subset X$  zwei konvergente Folgen von Vektoren aus  $X$ . Außerdem seien zwei konvergente Folgen reeller Zahlen  $(\lambda_j)_{j \in \mathbb{N}}$  und  $(\mu_j)_{j \in \mathbb{N}}$  gegeben. Dann konvergiert die Folge

$$(\lambda_j x_j + \mu_j y_j)_{j \in \mathbb{N}}$$

ebenfalls, und es gilt

$$\lim_{j \rightarrow \infty} (\lambda_j x_j + \mu_j y_j) = \lim_{j \rightarrow \infty} \lambda_j \lim_{j \rightarrow \infty} x_j + \lim_{j \rightarrow \infty} \mu_j \lim_{j \rightarrow \infty} y_j.$$

Diese Eigenschaft heißt *Linearität des Grenzwerts*.

Ist  $X$  ein Skalarproduktraum, gilt für eine konvergente Folge  $(x_j)_{j \in \mathbb{N}} \subset X$  und  $y \in X$  außerdem

$$\lim_{j \rightarrow \infty} \langle x_j, y \rangle = \langle \lim_{j \rightarrow \infty} x_j, y \rangle.$$

Seien  $(A_j)_{j \in \mathbb{N}}, (B_j)_{j \in \mathbb{N}} \subset M(m, n)$  zwei konvergente Folgen von Matrizen und  $(x_j)_{j \in \mathbb{N}} \subset \mathbb{R}^n$  eine konvergente Folge von Vektoren. Dann gilt sowohl

$$\lim_{j \rightarrow \infty} (A_j x_j) = \lim_{j \rightarrow \infty} A_j \lim_{j \rightarrow \infty} x_j$$

als auch

$$\lim_{j \rightarrow \infty} (A_j B_j) = \lim_{j \rightarrow \infty} A_j \lim_{j \rightarrow \infty} B_j.$$

*Beweis.* Wir beweisen hier beispielhaft die vorletzte Aussage. Sei also  $(A_j)_{j \in \mathbb{N}} \subset M(m, n)$  eine konvergente Folge von Matrizen mit Grenzwert  $A \in M(m, n)$ , und  $(x_j)_{j \in \mathbb{N}} \subset \mathbb{R}^n$  eine konvergente Folge von Vektoren mit Grenzwert  $x \in \mathbb{R}^n$ . Dann existieren nach Definition der Konvergenz reelle Zahlen  $\varepsilon_1 > 0$  und  $\varepsilon_2 > 0$  und Indizes  $j_1, j_2 \in \mathbb{N}$ , sodass

$$\|A_j - A\|_* \leq \varepsilon_1 \quad \text{für alle } j \geq j_1$$

und

$$\|x_j - x\| \leq \varepsilon_2 \quad \text{für alle } j \geq j_2,$$

wobei  $\|\cdot\|_*$  eine beliebige mit der Vektornorm  $\|\cdot\|$  verträgliche Matrixnorm ist. Insbesondere gelten beide Aussagen für alle  $j \geq \max\{j_1, j_2\}$ , sodass  $\|A_j - A\|_* \leq \varepsilon$  und  $\|x_j - x\| \leq \varepsilon$  für  $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$ . Außerdem gibt es nach Bemerkung 1.10 eine Schranke  $M$ , sodass  $\|A_j\|_* \leq M$ . Es folgt, dass

$$\begin{aligned} \|A_j x_j - Ax\| &= \|A_j x_j - A_j x + A_j x - Ax\| \\ &\leq \|A_j(x - x_j)\| + \|(A_j - A)x\| && \text{(CSU)} \\ &\leq \|A_j\|_* \|x - x_j\| + \|A_j - A\|_* \|x\| && \text{(Verträglichkeit)} \\ &\leq M\varepsilon + \varepsilon \|x\| = (M + \|x\|)\varepsilon. && \text{(Bem. 1.10)} \end{aligned}$$

Da  $\varepsilon_1$  und  $\varepsilon_2$  beliebig gewählt waren, somit also auch  $\varepsilon > 0$  beliebig ist, und daher insbesondere beliebig klein gewählt werden kann, gibt es für alle  $\tilde{\varepsilon} := (M + \|x\|)\varepsilon > 0$  ein  $j_0 \in \mathbb{N}$ , sodass für alle  $j \geq j_0$

$$\|A_j x_j - Ax\| \leq \tilde{\varepsilon} := (M + \|x\|)\varepsilon$$

gilt. Die Folge  $(A_j x_j)_{j \in \mathbb{N}}$  von Vektoren konvergiert also gegen  $Ax$ . Insbesondere existiert der Grenzwert.  $\square$

**Definition 1.13 (Cauchy-Folge):** Eine Folge  $(x_j) \subset X$  von Vektoren eines normierten Vektorraums  $X$  heißt *Cauchy-Folge*, wenn für

$$\forall \varepsilon > 0 : \exists j_0 \in \mathbb{N} \forall k, l \geq j_0 : \|x_k - x_l\| \leq \varepsilon \quad (1.12)$$

gilt.

**Satz 1.14 (Cauchy-Kriterium):** 1. Eine konvergente Folge ist stets eine Cauchy-Folge. Es gilt also: Erfüllt eine Folge das Cauchy-Kriterium (1.12) nicht, so kann sie nicht konvergieren.

2. Eine konvergente Vektorreihe erfüllt das Cauchy-Kriterium. Es gilt also:

$$\forall \varepsilon > 0 \exists j_0 \in \mathbb{N} \forall k, l \geq j_0 : \left\| \sum_{j=l+1}^k x_j \right\| \leq \varepsilon.$$

Insbesondere muss die Folge, über die die Partialsummen gebildet werden, stets einen Nullfolge sein, damit überhaupt die Chance besteht, dass die Reihe konvergiert.

Bemerkung: Die Umkehrung, dass eine Cauchy-Folge eine konvergente Folge ist, gilt im Allgemeinen nicht! Diese Äquivalenz gilt nur in sogenannten *vollständigen Räumen*.

**Beispiel: Die Neumann-Reihe** Die Neumann-Reihe ist ein wichtiges und sehr allgemein gültiges Beispiel, mit deren Hilfe die Existenz der Inversen einer Matrix (oder allgemeiner: eines linearen Operators) bewiesen werden kann.

Sei der Vektorraum  $M(m, m)$  der quadratischen Matrizen mit einer submultiplikativen Matrixnorm  $\|\cdot\|_*$  (beispielsweise der Frobeniusnorm) versehen. Sei  $A \in M(m, m)$  und gelte  $\|A\|_* < 1$ . Dann ist die *Neumann-Reihe* zu  $A$  gegeben durch

$$\left( \sum_{k=0}^{\infty} A^k \right)_{n \in \mathbb{N}}.$$

Es gilt folgender Satz.

**Satz 1.15 (Neumann-Reihe):** Unter genannten Voraussetzungen konvergiert die Neumann-Reihe. Ihr Grenzwert erfüllt die Identität

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Hierbei bezeichnet  $I \in M(m, m)$  die Einheitsmatrix. Bemerke, dass die obige Gleichheit eine Verallgemeinerung der Aussage für geometrische Reihen ist: Ist nämlich  $|x| < 1$ , so gilt

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

*Beweis.* Definiere mit  $s_n$  die  $n$ -te Partialsumme, also

$$s_n := \sum_{k=0}^n A^k \in M(m, m).$$

Dann gilt

$$s_n - A \cdot s_n = \sum_{k=0}^n A^k - A \sum_{k=0}^n A^k = \sum_{k=0}^n A^k - \sum_{k=1}^{n+1} A^k = I - A^{n+1},$$

da sich bis auf den ersten und den letzten Summanden alle anderen gegenseitig auslöschen. Es ist also

$$(I - A)s_n = I - A^{n+1}, \quad \text{weswegen} \quad (I - A) \lim_{n \rightarrow \infty} s_n = I - \lim_{n \rightarrow \infty} A^{n+1} \quad (*)$$

folgt. Da  $\|A\|_* < 1$ , ist  $(\|A\|_*^n)_{n \in \mathbb{N}}$  eine Nullfolge. Daraus folgt nun wegen der Submultiplikativität der gewählten Matrixnorm, dass

$$\|A^n - 0\|_* = \|A^n\|_* \leq \|A\|_*^n \xrightarrow{n \rightarrow \infty} 0.$$

Damit ist also auch  $(A^n)_{n \rightarrow \infty}$  eine Nullfolge. Die rechte Seite in obiger Abschätzung (\*) konvergiert somit gegen  $I$ , woraus

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{j=0}^n A^j = (I - A)^{-1}$$

folgt. Damit ist alles gezeigt. □

## Die Exponentialfunktion für Matrizen

**Definition 1.16:** Sei  $A \in M(m, m)$  und  $\|\cdot\|_*$  eine beliebige Matrixnorm. Dann existiert der Wert

$$\exp(A) := \sum_{j=0}^{\infty} \frac{1}{j!} A^j,$$

und die Abbildung  $\exp : M(m, m) \rightarrow \mathbb{R}$  mit  $A \mapsto \exp(A)$  wird auch *Exponentialfunktion für Matrizen* genannt.

*Hinweis:* Die Existenz der Werte muss streng genommen gezeigt werden. Der Beweis läuft analog zum Beweis der Existenz des Grenzwerts für die Neumann-Reihe und verbleibt daher zur Übung.

**Lemma 1.17 (Eigenschaften der Exponentialfunktion für Matrizen):** Für beliebige Matrizen  $A, B \in M(m, m)$  gelten folgende Aussagen.

a) Falls  $AB = BA$ , so gilt

$$\exp(A + B) = \exp(A) + \exp(B).$$

b) Falls  $B$  regulär ist, also  $\det B \neq 0$ , so gilt

$$\exp(BAB^{-1}) = B \exp(A) B^{-1}.$$

c) Für alle  $A \in M(m, m)$  gilt die Gleichheit

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

d) Für alle  $A \in M(m, m)$  ist  $\exp(A)$  regulär. Es gilt also  $\det(\exp(A)) \neq 0$ . Außerdem gilt

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

### Beispiele

1. Sei  $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$  mit  $a, b \in \mathbb{R}$ . Dann gilt

$$\exp(A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} + \frac{1}{2} \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} + \dots = \begin{pmatrix} \exp(a) & 0 \\ 0 & \exp(b) \end{pmatrix}.$$

2. Sei  $B = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$  mit  $a \in \mathbb{R}$ . Dann gilt

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}^n = 0$$

für  $n \geq 2$ , und deshalb

$$\exp(B) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} + 0 = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

3. Sei  $C = \begin{pmatrix} 0 & -\varphi \\ \varphi & 0 \end{pmatrix}$  mit  $\varphi \in [0, 2\pi)$ . Dann ist

$$\exp(C) = \begin{pmatrix} \sum_{k=0}^{\infty} (-1)^k \frac{\varphi^{2k}}{(2k)!} & - \sum_{k=0}^{\infty} (-1)^k \frac{\varphi^{2k+1}}{(2k+1)!} \\ \sum_{k=0}^{\infty} (-1)^k \frac{\varphi^{2k+1}}{(2k+1)!} & \sum_{k=0}^{\infty} (-1)^k \frac{\varphi^{2k}}{(2k)!} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

## 1.3 Offenheit und Abgeschlossenheit von Mengen

Sei  $(X, \|\cdot\|)$  ein normierter Vektorraum und  $M \subset X$  eine Teilmenge von  $X$ .

**Definition 1.18:** Sei  $x \in X$  und  $r > 0$  eine positive reelle Zahl.

1. Die Mengen

$$K_r(x) := \{y \in X, \|x - y\| < r\} \tag{1.13}$$

beziehungsweise

$$\overline{K_r}(x) := \{y \in X, \|x - y\| \leq r\} \tag{1.14}$$

heißen *offene* beziehungsweise *abgeschlossene* Kugel mit Radius  $r$  um  $x$ .

2. Ein Punkt  $x \in M$  heißt *innerer Punkt von  $M$* , falls ein  $r > 0$  mit  $K_r(x) \subset M$  existiert.
3. Die Menge aller inneren Punkte einer Menge  $M$  heißt *innerer Kern von  $M$*  und wird mit  $\text{int } M$  oder  $\overset{\circ}{M}$  bezeichnet. Es gilt stets  $\text{int } M \subset M$ .
4. Eine Menge  $M$  heißt *offen*, wenn  $\overset{\circ}{M} = M$ .
5. Ein Punkt  $x \in M$  heißt *Randpunkt von  $M$* , wenn für alle  $r > 0$  gilt:

$$\left( K_r(x) \cap M \neq \emptyset \right) \wedge \left( K_r(x) \cap (X \setminus M) \neq \emptyset \right).$$

Hierbei bezeichnet  $X \setminus M := M^C := \{x \in X, x \notin M\}$  das *Komplement von  $M$  in  $X$* .

6. Die Menge aller Randpunkte von  $M$  heißt *Rand von  $M$*  und wird mit  $\partial M$  bezeichnet. Beispielsweise gilt

$$\partial K_r(x) = \{x \in X, \|x - y\| = r\}.$$

7. Die Menge  $\overline{M} := M \cup \partial M$  heißt *abgeschlossene Hülle von  $M$*  oder *Abschluss von  $M$* . Allgemein gilt  $M \subset \overline{M}$ .
8.  $M$  heißt *abgeschlossen*, wenn  $M = \overline{M}$ .

**Beispiele**

- a) Endliche Mengen  $M = \{x_1, \dots, x_n\}$  sind abgeschlossen. Es gibt nämlich für alle  $x_i$  ein  $r_i > 0$ , sodass  $K_{r_i}(x_i) \cap M = \{x_i\}$ . Insbesondere ist jeder Punkt  $x_i$  Randpunkt von  $M$ , sodass  $M = \partial M$ .
- b) Die Mengen  $X$  und  $\emptyset$  sind die einzigen Teilmengen von  $X$ , die gleichzeitig offen und abgeschlossen sind.
- c) Sei  $X = \mathbb{R}$ . Dann gilt für alle  $a, b \in \mathbb{R}$  mit  $a < b$ , dass das Intervall  $(a, b)$  offen, das Intervall  $[a, b]$  abgeschlossen, und die Intervalle  $(a, b]$  und  $[a, b)$  weder offen noch abgeschlossen sind. Außerdem gilt für den Rand  $\partial((a, b)) = \partial((a, b]) = \partial([a, b)) = \partial([a, b]) = \{a, b\}$  und für das Innere  $\text{int}(a, b) = \text{int}(a, b]) = \text{int}([a, b)) = \text{int}([a, b]) = (a, b)$ .
- d) Sei  $X = \mathbb{R}$  und  $M = \mathbb{Q} \subset \mathbb{R}$ . Dann gilt:  $\text{int } \mathbb{Q} = \emptyset$  und  $\partial \mathbb{Q} = \overline{\mathbb{Q}} = \mathbb{R}$ .
- e) Sei  $X = \mathbb{R}^2$ . Dann sind die Achsen  $\{(x, y), x = 0\}$  und  $\{(x, y), y = 0\}$  abgeschlossen. Ebenso sind die beliebige abgeschlossene Intervalle  $[a, b] \subset \mathbb{R}^2$  abgeschlossen. Keine der genannten Teilmengen besitzt innere Punkte.

**Lemma 1.19 (Dualität von Offenheit und Abgeschlossenheit):** *Es gelten folgende Aussagen:*

1. Eine Menge  $M \subset X$  ist genau dann offen, wenn  $X \setminus M$  abgeschlossen ist.
2. Der Durchschnitt beliebig vieler und die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.
3. Der Durchschnitt endlich vieler und die Vereinigung beliebig vieler offener Mengen ist offen.
4. Eine Menge ist genau dann abgeschlossen, wenn sie alle ihre Randpunkte enthält.
5. Eine Menge ist genau dann offen, wenn sie keinen ihrer Randpunkte enthält.
6. Sei  $I \subset \mathbb{N}$  eine abzählbare Indexmenge und  $M_i \subset X$  für alle  $i \in I$ . Dann gelten die De Morgan'schen Regeln:

$$X \setminus \left( \bigcup_{i \in I} M_i \right) = \bigcap_{i \in I} X \setminus M_i; \quad (1.15)$$

$$X \setminus \left( \bigcap_{i \in I} M_i \right) = \bigcup_{i \in I} X \setminus M_i; \quad (1.16)$$

und

$$X \setminus X = \emptyset. \quad (1.17)$$

*Beweis.* Wir beweisen hier beispielhaft nur, dass die Vereinigung beliebig vieler offener Mengen offen ist. Sei dazu  $x \in \bigcup_{i \in I} M_i$  beliebig gewählt. Dann gibt es ein  $i_0$ , sodass  $x \in M_{i_0}$ . Da  $M_{i_0}$  offen ist, existiert also ein Radius  $r > 0$  und eine Kugel  $K_r(x)$ , sodass  $K_r(x) \subset M_{i_0}$ . Diese Kugel ist also auch Teilmenge der Vereinigung aller  $M_i$ . Insbesondere ist  $x \in \text{int} \bigcup_{i \in I} M_i$ . Da dies beliebiges  $x$  galt, ist die Vereinigung beliebig vieler offener Mengen offen.  $\square$

**Bemerkung 1.20:** Sei  $j \in \mathbb{N} \setminus \{0\}$ . Für jedes  $j$  ist das Intervall  $(-\frac{1}{j}, \frac{1}{j}) \subset \mathbb{R}$  offen. Für den Durchschnitt all dieser (abzählbar) unendlich vielen Intervalle gilt jedoch

$$\bigcap_{j=1}^{\infty} \left( -\frac{1}{j}, \frac{1}{j} \right) = \{0\}.$$

Er ist also abgeschlossen. Das illustriert, dass der Durchschnitt unendlich vieler offener Mengen im Allgemeinen nicht offen ist.

Sei nun  $[-1 + \frac{1}{j}, 1 - \frac{1}{j}] \subset \mathbb{R}$ . Für jedes  $j$  ist dies ein abgeschlossenes Intervall. Für die Vereinigung all dieser (abzählbar) unendlich vielen Intervalle gilt jedoch

$$\bigcup_{j=1}^{\infty} \left[ 1 + \frac{1}{j}, 1 - \frac{1}{j} \right] = (-1, 1).$$

Sie ist also offen. Das illustriert, dass die Vereinigung unendlich vieler abgeschlossener Mengen im Allgemeinen nicht abgeschlossen ist.



**Satz 1.21:** *Es gelten folgende Aussagen:*

1. Die Menge  $\text{int } M$  ist die größte offene Teilmenge von  $M$ . Mathematisch ausgedrückt: Für alle  $A \subset M$  mit  $A$  offen gilt  $A \subset \text{int } M$ .
2. Die Menge  $\overline{M}$  ist die kleinste abgeschlossene Menge, die  $M$  enthält. Mathematisch ausgedrückt: Für alle  $A \supset M$  mit  $A$  abgeschlossen gilt  $\overline{M} \subset A$ .
3. Ein Element  $x \in X$  ist genau dann ein Randpunkt von  $M$ , d. h.  $x \in \partial M$ , wenn Folgen  $(x_j) \subset M$  und  $(y_j) \subset X \setminus M$  existieren, sodass  $x_j \rightarrow x$  und  $y_j \rightarrow x$  für  $j \rightarrow \infty$ .
4. Eine Menge  $M$  ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten Folge  $(x_j) \subset M$  ebenfalls in  $M$  liegt. Dieses Kriterium heißt auch *Folgenabgeschlossenheit* von  $M$ .

## 1.4 Grenzwerte von Funktionen in einem Punkt

Abbildungen (Funktionen) zwischen höher-dimensionalen Vektorräumen haben, je nach Dimension des Definitions- und Wertebereichs, verschiedene Namen beziehungsweise physikalische Interpretationen. Zum Beispiel sind

- *skalare Felder* Abbildungen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Sie beschreiben beispielsweise die zeitliche Änderung einer Temperaturverteilung:  $T : \mathbb{R}^4 \rightarrow \mathbb{R}$  mit  $T(x, y, z, t)$  ist die Temperatur an einem Ort  $(x, y, z)$  zu einem Zeitpunkt  $t$ .
- *Vektorfelder* Abbildungen  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Sie beschreiben beispielsweise die Geschwindigkeit einer fließenden Flüssigkeit an einem Ort:  $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit  $v(x, y, z) = (v_x, v_y, v_z)$ , wobei  $v_x, v_y, v_z$  die Geschwindigkeitskomponenten an  $(x, y, z)$  parallel zu den entsprechenden Achsen sind.
- *Tensorfelder* Abbildungen  $\mathbb{R}^n \rightarrow M(l, m)$ . Jedem Punkt  $x \in \mathbb{R}^n$  wird also eine Matrix zugeordnet. Beispielsweise lassen sich in der Elektrodynamik einem Punkt der Feldstärketensor zuordnen.

Im Folgenden seien  $X$  und  $Y$  normierte Vektorräume,  $M \subset X$  und  $f : X \rightarrow Y$  eine Abbildung, zunächst ohne weitere spezielle Eigenschaften.

**Definition 1.22 (Häufungspunkt):** Ein Element  $x_0 \in X$  heißt *Häufungspunkt* von  $M$ , wenn

$$\forall \delta > 0 \exists x \in M \setminus \{x_0\} : \|x - x_0\| \leq \delta. \quad (1.18)$$

Es gibt also eine Folge  $(x_j) \subset M \setminus \{x_0\}$  mit  $x_j \rightarrow x_0$ .

**Beispiel** Sei  $X = \mathbb{R}$ ,  $M \subset X$  und bezeichne  $\text{HP}(M)$  die Menge der Häufungspunkte von  $M$ . Dann ist

$$\text{HP}(\mathbb{N}) = \emptyset,$$

die Menge der natürlichen Zahlen besitzt also keine Häufungspunkte;

$$\text{HP}(\mathbb{Q}) = \mathbb{R},$$

jede reelle Zahl ist also Häufungspunkt der Menge der rationalen Zahlen. Man sagt auch: Die rationalen Zahlen liegen *dicht in*  $\mathbb{R}$ ; sowie

$$\text{HP}(\{1 + \frac{1}{n}, n \in \mathbb{N}\}) = \{1\}.$$

**Definition 1.23 (Konvergenz von Funktionswerten):** Sei nun  $x_0$  ein Häufungspunkt von  $M$ ,  $y_0 \in Y$  beliebig, und es gelte

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in M : \|x - x_0\|_X \leq \delta \implies \|f(x) - y_0\|_Y \leq \varepsilon. \quad (1.19)$$

Dann heißt  $f$  *konvergent für*  $x \rightarrow x_0$  und  $y_0$  heißt *Grenzwert von*  $f$  *für*  $x \rightarrow x_0$ . Falls so ein  $y_0$  existiert, schreibt man auch

$$y_0 = \lim_{x \rightarrow x_0} f(x) \quad \text{oder} \quad f(x) \rightarrow y_0 \quad \text{für } x \rightarrow x_0.$$

**Bemerkung 1.24:** Man bemerke:

- a) Für die Definition der Konvergenz von Funktionswerten (1.19) muss  $f$  nicht in  $x_0$  definiert sein.
- b) Durch die Bedingung (1.19) ist  $y_0$  eindeutig festgelegt.
- c) Es gilt die Äquivalenz folgender Aussagen:
  1. Der Punkt  $y_0$  ist Grenzwert von  $f$  für  $x \rightarrow x_0$ , also  $y_0 = \lim_{x \rightarrow x_0} f(x)$ .
  2. Für jede Folge  $(x_j) \subset M \setminus \{x_0\}$  mit  $x_j \rightarrow x_0$  gilt:  $f(x_j) \rightarrow y_0$ .

Ebenso wie bei der Konvergenz von Vektoren im vorigen Abschnitt ist Konvergenz von vektorwertigen Funktionen äquivalent zur Konvergenz aller Komponenten. Konvergenz bedeutet also koordinatenweise Konvergenz. Um das einzusehen, betrachte eine Funktion  $f : M \subset X \rightarrow \mathbb{R}^n$ , wobei  $X$  ein beliebiger Vektorraum und  $M$  eine Teilmenge dieses Vektorraums ist. Wir wählen eine Basis  $\{b_1, \dots, b_n\}$  des Zielraums  $\mathbb{R}^n$ . Dann lässt sich der Funktionswert  $f(x)$  für alle  $x \in M$  darstellen als

$$f(x) = \sum_{k=1}^n f_k(x) b_k.$$

Hierbei bezeichnen die Funktionen  $f_k : M \rightarrow \mathbb{R}$  die Koordinatenfunktionen zu  $f$  bezüglich der Basis  $\{b_1, \dots, b_n\}$ . Für einen Häufungspunkt  $x_0$  von  $M$  gilt dann

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} \left( \sum_{k=1}^n f_k(x) b_k \right) = \sum_{k=1}^n \left( \lim_{x \rightarrow x_0} f_k(x) \right) b_k.$$

Der Grenzwert auf der linken Seite existiert also genau dann, wenn alle Grenzwerte der Koordinatenfunktionen  $f_k$  für  $x \rightarrow x_0$  auf der rechten Seite existieren.

Betrachten wir zu guter Letzt noch den Spezialfall  $X = \mathbb{R}^n$ , wobei wir als Basis die kanonische Basis der Einheitsvektoren im  $\mathbb{R}^n$  wählen. Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine beliebige Funktion mit

$$f(x_1, \dots, x_n) := (f_1(x), \dots, f_m(x)).$$

Dann ist

$$\lim_{x \rightarrow x_0} (f_1(x), \dots, f_m(x)) = \left( \lim_{x \rightarrow x_0} f_1(x), \dots, \lim_{x \rightarrow x_0} f_m(x) \right).$$

### Beispiele

1. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  gegeben durch

$$f(x_1, x_2) := (\sin x_1, \cos x_2, \exp(x_1 + x_2)).$$

Dann ist  $\lim_{x \rightarrow 0} f(x_1, x_2) = (0, 1, 1)$ .

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

$$f(x, y) := (x^2 \cdot y^3 + 2, \sin(xy)).$$

Dann liegt der Grenzwert von  $f$  für  $(x, y) \rightarrow (0, 0)$  bei  $(2, 0, 0)$ .

3. Sei  $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$  definiert durch

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

Dann existiert der Grenzwert von  $f$  für  $(x, y) \rightarrow (0, 0)$  nicht. Dies lässt sich leicht einsehen, denn wie zuvor bedeutet Konvergenz koordinatenweise Konvergenz. Sei jedoch  $(x_k)$  eine Nullfolge und betrachten wir die Folge  $(x_k, 0)$ , die dann gegen  $(0, 0)$  konvergiert, so gilt für die Funktionswerte, dass

$$f(x_k, 0) = \frac{x_k}{x_k^2} = \frac{1}{x_k}.$$

Die rechte Seite strebt jedoch für  $x_k \rightarrow 0$  gegen unendlich. Deswegen ist  $f$  nicht konvergent für  $(x, y) \rightarrow (0, 0)$ .

4. Sei  $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$  gegeben durch

$$f(x, y) := \frac{x^2 y^2}{x^2 + y^2}.$$

Wir fragen uns, ob der Grenzwert von  $f$  für  $(x, y) \rightarrow (0, 0)$  existiert. Dazu schätzen wir die Funktionswerte folgendermaßen nach unten und nach oben ab:

$$0 \leq \frac{x^2 y^2}{x^2 + y^2} = x^2 \left( \frac{y^2}{x^2 + y^2} \right) \leq x^2,$$

da  $y^2 \geq 0$ . Wählen wir nun eine beliebige Nullfolge  $(x_k, y_k)$ , so liegen alle Funktionswerte dieser Folge zwischen im Intervall  $[0, x_k^2]$ . Es gilt also

$$0 \leq f(x_k, y_k) \leq x_k^2.$$

Da die obere Schranke für  $x_k \rightarrow 0$  nun gegen Null konvergiert, müssen auch die Funktionswerte gegen Null konvergieren, da die obere und die untere Schranke gleich Null sind (Sandwich-Argument). Deswegen existiert der Grenzwert und wir dürfen schreiben

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = (0, 0).$$

Alternativ lässt sich der Nachweis auch folgendermaßen führen: Wir vermuten 0 als Grenzwert. Wir müssen also die Bedingung (1.19) nachprüfen. Dazu halten  $\varepsilon > 0$  beliebig fixiert und wählen  $\delta \leq \sqrt{\varepsilon}$ . Dann folgt aus der Voraussetzung

$$|x| \leq \|(x, y) - (0, 0)\|_2 = \sqrt{x^2 + y^2} \leq \delta,$$

dass

$$\left| \frac{x^2 y^2}{x^2 + y^2} - 0 \right| = |x^2| \left| \frac{y^2}{x^2 + y^2} \right| \leq |x^2| \leq |x|^2 \leq \delta^2 \leq \varepsilon,$$

weswegen  $|f(x, y) - 0| \leq \varepsilon$ . Das war zu zeigen.

5. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definiert durch

$$f(x, y) := \frac{y}{x - 1}.$$

Hier ist die interessante Frage, ob der Grenzwert für  $(x, y) \rightarrow (1, 0)$  existiert. Sei dafür  $(x_k)$  eine Folge mit  $x_k \rightarrow 1$  und betrachte die Funktionswerte zu  $(x_k, 0)$ . Es gilt zwar, dass

$$f(x_k, 0) = 0,$$

weswegen  $f$  entlang dieser Folge gegen 0 konvergiert. Für die Funktionswerte der Folge  $(x_k, x_k - 1)$ , die ebenfalls gegen  $(1, 0)$  konvergiert, gilt jedoch

$$f(x_k, x_k) = \frac{x_k - 1}{x_k - 1} = 1.$$

Der vermeintliche Grenzwert ist also nicht eindeutig bestimmt; die Funktion divergiert für  $(x, y) \rightarrow (1, 0)$ .

**Lemma 1.25 (Eigenschaften des Grenzwerts):** Sei  $M \subset X$  und seien skalare Funktionen  $\lambda, \mu : M \rightarrow \mathbb{R}$  sowie vektorwertige Funktionen  $f, g : M \rightarrow \mathbb{R}^n$  gegeben. Außerdem seien Matrixabbildungen (Tensorfelder)  $A : X \rightarrow M(m, n)$  und  $B : X \rightarrow M(n, l)$  gegeben. Sei  $x_0$  ein Häufungspunkt von  $M$ . Dann gilt

$$1. \quad \lim_{x \rightarrow x_0} (\lambda(x)f(x) + \mu(x)g(x)) = \lim_{x \rightarrow x_0} \lambda(x) \cdot \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} \mu(x) \cdot \lim_{x \rightarrow x_0} g(x), \quad (1.20)$$

$$2. \quad \lim_{x \rightarrow x_0} (A(x)f(x)) = \lim_{x \rightarrow x_0} A(x) \cdot \lim_{x \rightarrow x_0} f(x), \quad (1.21)$$

$$3. \quad \lim_{x \rightarrow x_0} (A(x)B(x)) = \lim_{x \rightarrow x_0} A(x) \lim_{x \rightarrow x_0} B(x), \quad (1.22)$$

sowie

$$4. \quad \lim_{x \rightarrow x_0} \langle f(x), g(x) \rangle = \left\langle \lim_{x \rightarrow x_0} f(x), \lim_{x \rightarrow x_0} g(x) \right\rangle. \quad (1.23)$$

Hierbei ist  $\langle \cdot, \cdot \rangle$  das Standardskalarprodukt in  $\mathbb{R}^n$ .

Die Grenzwerte auf den linken Seiten existieren genau dann, wenn die Grenzwerte auf den entsprechenden rechten Seiten konvergieren.

*Beweis.* Der Beweis verwendet die Eigenschaften des Grenzwerts nach Lemma 1.12 und verbleibt zur Übung. □

Wir haben nun bereits einige Eigenschaften der Grenzwertbildung wie Linearität kennengelernt. Eine Frage ist dabei bisher unbeantwortet geblieben, nämlich die Frage nach der Vertauschbarkeit von Grenzwerten. Gilt also beispielsweise für eine Funktion  $f : M \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ , dass

$$\lim_{y \rightarrow y_0} \left( \lim_{x \rightarrow x_0} f(x, y) \right) = \lim_{x \rightarrow x_0} \left( \lim_{y \rightarrow y_0} f(x, y) \right), \quad (1.24)$$

wobei  $(x_0, y_0)$  ein Häufungspunkt von  $M$  ist? Existiert also der rechte Grenzwert genau dann, wenn der linke Grenzwert existiert, und müssen, falls dies der Fall ist, auch beide übereinstimmen? Wie die nächsten Beispiele zeigen, ist leider weder das eine noch das andere zutreffend.

**Beispiele** Zunächst zur Existenz lediglich eines der beiden Grenzwerte. Sei hierfür eine Funktion  $f : M \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben, wobei  $M = [0, \infty) \times [0, \infty) \setminus \{(0, 0)\}$  und

$$f(x, y) := \frac{x^2}{x^2 + y^2}.$$

Dann gilt

$$\lim_{y \rightarrow y_0} \left( \lim_{x \rightarrow x_0} f(x, y) \right) = \lim_{y \rightarrow y_0} 0 = 0$$

und

$$\lim_{x \rightarrow x_0} \left( \lim_{y \rightarrow y_0} f(x, y) \right) = \lim_{x \rightarrow x_0} 1 = 1.$$

Außerdem existiert der Grenzwert von  $f$  für  $(x, y) \rightarrow (0, 0)$  nicht, da

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

aber für  $y = x$ , also entlang der Gerade  $(x, x)$  gilt

$$\lim_{(x,x) \rightarrow (0,0)} f(x, y) = \lim_{x \rightarrow 0} \frac{x^2}{x^2 + x^2} = \frac{1}{2}.$$

Es ist sogar möglich, dass die koordinatenweise Grenzwerte übereinstimmen, aber der punktweise Grenzwert dennoch nicht existiert. Das zeigt das folgende Beispiel: Sei  $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$  mit

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

Dann existiert wegen

$$\lim_{(x,0) \rightarrow (0,0)} f(x, y) = 0 \neq \frac{1}{2} = \lim_{(x,x) \rightarrow (0,0)} f(x, y)$$

der Grenzwert von  $f$  bezüglich  $(x, y) \rightarrow (0, 0)$  nicht, obwohl

$$\lim_{x \rightarrow 0} \left( \lim_{y \rightarrow 0} f(x, y) \right) = 0 = \lim_{y \rightarrow 0} \left( \lim_{x \rightarrow 0} f(x, y) \right).$$

Wir wollen die Beobachtungen nun in einem kleinen Lemma zusammenfassen.

**Lemma 1.26:** Sei  $f : M \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  eine beliebige Funktion und  $(x_0, y_0)$  ein Häufungspunkt von  $M$ . Folgende Bedingungen sind hinreichend dafür, dass der Grenzwert

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$$

nicht existiert.

1. Die Grenzwerte

$$\lim_{y \rightarrow y_0} \left( \lim_{x \rightarrow x_0} f(x, y) \right) \quad \text{und} \quad \lim_{x \rightarrow x_0} \left( \lim_{y \rightarrow y_0} f(x, y) \right)$$

existieren, sind aber verschieden.

2. Es existieren konvergente Folgen  $x_k \rightarrow x_0$  und  $y_k \rightarrow y_0$ , sodass  $f(x_k, y_k)$  nicht konvergiert.

3. Es existieren konvergente Folgen  $x_k \rightarrow x_0$  und  $y_k \rightarrow y_0$  sowie  $\bar{x}_k \rightarrow x_0$  und  $\bar{y}_k \rightarrow y_0$ , sodass  $f(x_k, y_k)$  und  $f(\bar{x}_k, \bar{y}_k)$  existieren, aber beide Grenzwerte stimmen nicht überein.

**Definition 1.27:** Sei eine Funktion  $f : M \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben. Dann *konvergiert*  $f(x, \cdot)$  für  $x \rightarrow x_0$  *gleichmäßig gegen eine Funktion  $g$  auf  $(y_0 - r, y_0 + r)$* , wenn

$$\forall \varepsilon > 0 \forall y \in (y_0 - r, y_0 + r) \exists \delta > 0 : |x - x_0| \leq \delta \implies |f(x, y) - g(y)| \leq \varepsilon. \quad (1.25)$$

Bemerkung: Der Ausdruck  $f(x, \cdot)$  kann als Schar von Funktionen verstanden werden, die mit  $x$  parametrisiert wird: Zu jedem  $x$  gehört eine Funktion  $f(x, \cdot) : M \cap \mathbb{R} \rightarrow \mathbb{R}$ . Die gleichmäßige Konvergenz im zweiten Argument für  $x \rightarrow x_0$  bedeutet nun, dass die Konvergenz unabhängig vom gewählten  $x$  ist: Das  $\delta > 0$  in obiger Definition hängt nur von  $x_0$ , nicht jedoch von  $x$ .

**Satz 1.28:** Wenn  $f(x, \cdot)$  für  $x \rightarrow x_0$  *gleichmäßig gegen  $g$  auf  $(y_0 - r, y_0 + r)$  konvergiert* und  $g(y)$  für  $y \rightarrow y_0$  *gegen ein  $z_0$  konvergiert*, dann existiert der *gleichmäßige Limes* von  $f$  für  $(x, y) \rightarrow (x_0, y_0)$  existiert. Außerdem ist

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = z_0.$$

## 1.5 Stetige Funktionen zwischen normierten Vektorräumen

Kommen wir nun zu einem der zentralen Begriffe der Analysis.

**Definition 1.29 (Stetigkeit):** Seien  $X$  und  $Y$  zwei normierte Vektorräume. Eine Abbildung  $f : M \subset X \rightarrow Y$  heißt *stetig in einem Punkt  $x_0 \in M$* , wenn gilt

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in M : \|x - x_0\|_X \leq \delta \implies \|f(x) - f(x_0)\|_Y \leq \varepsilon. \quad (1.26)$$

Die Abbildung heißt *stetig*, falls sie in jedem Punkt in  $M$  stetig ist.

**Lemma 1.30:** Sei  $M \subset X$ ,  $x_0$  ein Häufungspunkt von  $M$  und  $f : M \rightarrow Y$ . Dann sind folgende Aussagen äquivalent.

1. Die Funktion  $f$  ist stetig in  $x_0$ .
2. Für jede Folge  $(x_k) \subset M$  mit  $x_k \rightarrow x_0$  gilt:  $\lim_{x_k \rightarrow x_0} f(x_k) = f(x_0)$ . Diese Eigenschaft heißt auch *Folgenstetigkeit*.

*Beweis.* □

**Bemerkung 1.31:** Wenn  $x_0 \in M$  kein Häufungspunkt von  $M$  sondern ein isolierter Punkt ist, dann ist jede Abbildung  $f : M \rightarrow Y$  stetig in  $x_0$ .

### Beispiele

1. Jede konstante Funktion  $f : X \rightarrow Y$  ist stetig.
2. Sei  $Y = X$ . Dann ist die Abbildung  $\text{id} : X \rightarrow X$  mit  $x \mapsto x$  stetig. Diese Abbildung heißt auch *Identitätsabbildung*.
3. Jede lineare Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist stetig. Bemerke, dass die Menge der linearen Abbildung isomorph zur Menge der Matrizen  $M(m, n)$  ist. Für jede lineare Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  existiert also genau eine Matrix  $A \in M(m, n)$  mit

$$f(x) = Ax.$$

Zum Beweis der Stetigkeit betrachte ein beliebiges  $x \in \mathbb{R}^n$  und irgendeine submultiplikative Matrixnorm  $\|\cdot\|_*$  in  $M(m, n)$ , die mit der gewählten Vektornorm  $\|\cdot\|_n$  in  $\mathbb{R}^n$  verträglich ist. Sei  $\varepsilon > 0$  beliebig gewählt und setze  $\delta := \frac{\varepsilon}{\|A\|_*}$ . Dann gilt

$$\|x - y\|_n \leq \delta \implies \|f(x) - f(y)\|_m = \|Ax - Ay\|_m \leq \|A\|_* \|x - y\|_n \leq \varepsilon$$

nach Wahl von  $\delta$ . Bemerke, dass wir weiterhin zum Nachweis von Konvergenz, insbesondere zum Nachweis von Stetigkeit der Abbildungen, beliebige Normen wählen können, weil sie in unserem endlich-dimensionalen Setting äquivalent sind.

4. Alle Polynomfunktionen sind stetig.
5. Die Abbildung  $\det : M(n, n) \rightarrow \mathbb{R}$ , die einer Matrix ihre Determinante zuweist, ist stetig.
6. Die Exponentialfunktion für Matrizen  $\exp : M(n, n) \rightarrow M(n, n)$  mit  $A \mapsto \exp(A)$  ist stetig.

**Satz 1.32:** Seien  $X, Y$  und  $Z$  normierte Vektorräume und seien  $M \subset X$  und  $N \subset Y$  beliebige Teilmengen. Weiterhin seien die Funktionen  $f : M \rightarrow Y$  und  $g : N \rightarrow Z$  mit  $f(M) \subset N$  gegeben. Sind die Funktionen  $f$  in  $x_0 \in M$  und  $g$  in  $f(x_0) =: y_0 \in N$  stetig, so ist auch die Komposition (Verknüpfung) beider Funktionen

$$g \circ f : M \rightarrow Z \quad \text{mit } (g \circ f)(x) := g(f(x))$$

stetig in  $x_0$ .

*Beweis.* Sei  $\varepsilon > 0$  beliebig fixiert. Da  $g$  stetig in  $y_0$  ist, existiert ein  $\delta_1 > 0$ , sodass für alle  $y \in N$  gilt

$$\|y - y_0\|_Y \leq \delta_1 \implies \|g(y) - g(y_0)\|_Z \leq \varepsilon.$$

Das gilt insbesondere für die  $y \in N$ , die im Bild von  $f$  liegen. Zudem ist  $f$  nach Voraussetzung stetig in  $x_0 \in M$ , sodass  $\delta_2 > 0$  existiert, sodass für alle  $x \in M$  gilt

$$\|x - x_0\|_X \leq \delta_2 \implies \|f(x) - f(x_0)\|_Y \leq \delta_1.$$

Wir können für beliebiges  $\varepsilon > 0$  insbesondere  $\delta_1 := \varepsilon$  wählen. Dann folgt

$$\|x - x_0\|_X \leq \delta_2 \implies \|f(x) - f(x_0)\|_Y \leq \varepsilon = \delta_1 \implies \|g(f(x)) - g(f(x_0))\|_Z \leq \varepsilon$$

und damit die Stetigkeit der Komposition in  $x_0$ . □



**Beispiele** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stückweise definiert durch

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

Dann ist  $f$  nicht stetig in  $(0, 0)$ , da die Folge  $(\frac{1}{n}, \frac{1}{\sqrt{n}})$  gegen  $(0, 0)$  konvergiert, aber

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{1}{\sqrt{n}}\right) = 1$$

ist. Die Funktion konvergiert also für die angegebene Nullfolge nicht gegen den Funktionswert  $f(0, 0) = 0$  und ist somit in diesem Punkt unstetig.

Sei nun  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  erneut stückweise definiert durch

$$f(x, y) := \begin{cases} \frac{\sin(xy)}{xy}, & \text{für } (x, y) \in (\mathbb{R} \setminus \{0\})^2 \\ 1, & \text{sonst} \end{cases}.$$

Zum Nachweis der Stetigkeit von  $f$  definieren wir die Funktionen

$$g : \mathbb{R}^2 \rightarrow \mathbb{R} \quad \text{mit } g(x, y) := xy$$

und

$$h : \mathbb{R} \rightarrow \mathbb{R} \quad \text{mit } h(x) := \begin{cases} \frac{\sin x}{x}, & \text{für } x \neq 0 \\ 1, & \text{sonst} \end{cases}.$$

Dann ist  $g$  stetig in ganz  $\mathbb{R}^2$ , da es ein Polynom ist. Zudem ist  $h$  stetig auf  $\mathbb{R}$ , insbesondere stetig  $x = 0$ , da

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

wie man leicht aus der Reihendarstellung der Sinusfunktion ablesen kann. Insgesamt ist also auch die Komposition  $h \circ g = f$  stetig.

**Lemma 1.33:** Seien zwei Funktionen  $f, g : M \subset X \rightarrow Y$  gegeben, die stetig auf  $M$  sind. Dann ist auch die Funktion  $\lambda f + \mu g : M \rightarrow Y$  für alle  $\lambda, \mu \in \mathbb{R}$  stetig.

*Beweis.* Ohne Beschränkung der Allgemeinheit nehmen wir  $\lambda \neq 0 \neq \mu$  an. Da  $f$  und  $g$  stetig in allen Punkten  $x \in M$  sind, gibt es für beliebiges  $\varepsilon > 0$  Zahlen  $\delta_1 > 0$  und  $\delta_2 > 0$ , sodass für alle  $y \in M$  gilt

$$\|x - y\|_X \leq \delta_1 \implies \|f(x) - f(y)\|_Y \leq \frac{\varepsilon}{2|\lambda|}$$

sowie

$$\|x - y\|_X \leq \delta_2 \implies \|g(x) - g(y)\|_Y \leq \frac{\varepsilon}{2|\mu|}.$$

Wähle  $\delta := \min\{\delta_1, \delta_2\}$ . Dann gelten beide Implikationen für alle  $y$  mit  $\|x - y\|_X \leq \delta$ . Es folgt

$$\|(\lambda f + \mu g)(x) - (\lambda f + \mu g)(y)\|_Y = \|\lambda f(x) + \mu g(x) - \lambda f(y) - \mu g(y)\|_Y$$

$$\begin{aligned} &= \|\lambda(f(x) - f(y)) + \mu(g(x) - g(y))\|_Y \\ &\leq |\lambda| \|f(x) - f(y)\|_Y + |\mu| \|g(x) - g(y)\|_Y \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Damit ist alles gezeigt.  $\square$

**Lemma 1.34:** Seien  $f, g : M \subset X \rightarrow \mathbb{R}$  zwei auf  $M$  stetige Funktionen, so ist auch das Produkt  $f \cdot g : M \rightarrow \mathbb{R}$  mit  $(f \cdot g)(x) := f(x) \cdot g(x)$  stetig auf  $M$ .

*Beweis.* Die Aussage folgt aus Lemma 1.25. Es gilt nämlich für alle  $x_0 \in M$ , dass

$$\lim_{x \rightarrow x_0} (f \cdot g)(x) = \lim_{x \rightarrow x_0} f(x) \cdot g(x) = \lim_{x \rightarrow x_0} f(x) \cdot \lim_{x \rightarrow x_0} g(x) = f(x_0) \cdot g(x_0) = (f \cdot g)(x_0),$$

weil beide Grenzwerte auf der rechten Seite aufgrund der Stetigkeit beider Funktionen existieren.  $\square$

**Bemerkung 1.35:** Bemerke, dass aus den Lemmata 1.33 und 1.34 direkt die Stetigkeit von Polynomfunktionen folgt.

Wir untersuchen nun, wie sich offene und abgeschlossene Mengen unter stetigen Funktionen verhalten.

**Satz 1.36:** Sei  $M \subset X$  eine Teilmenge und  $f : M \rightarrow Y$  stetig.

1. Sei  $M$  offen. Dann gilt: Ist  $B \subset Y$  offen, so ist das Urbild  $f^{-1}(B) \subset M$  offen.
2. Sei  $M$  abgeschlossen. Dann gilt: Ist  $B \subset Y$  abgeschlossen, so ist auch  $f^{-1}(B) \subset M$  abgeschlossen.

*Beweis.* Sei  $B \subset Y$  offen. Ist  $f^{-1}(B) = \emptyset$ , das Urbild also leer, so ist die Behauptung erfüllt. Andernfalls wähle einen beliebigen Punkt  $x_0 \in f^{-1}(B)$ . Dann gilt  $f(x_0) \in B$  und da  $B$  offen ist, gibt es einen Radius  $r > 0$ , sodass

$$K_r(f(x_0)) \subset B.$$

Da  $f$  stetig ist, gibt es zu diesem  $r > 0$  ein  $\delta > 0$ , sodass für alle  $x \in M$  aus

$$\|x - x_0\|_X \leq \delta \implies \|f(x) - f(x_0)\|_Y \leq r.$$

Da  $M$  offen ist, kann  $\delta > 0$  insbesondere so klein gewählt werden, dass die Ungleichung scharf ist. Das Bild all dieser  $x \in M$  mit  $\|x - x_0\|_X < \delta$  liegt ja nun in  $B$ . Also liegen diese  $x$  insbesondere im Urbild  $f^{-1}(B)$ . Damit ist aber die gesamte Kugel

$$K_\delta(x_0) \subset f^{-1}(B).$$

Da  $x_0 \in f^{-1}(B)$  beliebig gewählt war, ist damit für alle  $x_0 \in f^{-1}(B)$  gezeigt, dass ein  $\delta > 0$  existiert, sodass die offene Kugel um  $x_0$  ebenfalls Teilmenge des Urbilds  $f^{-1}(B)$  ist. Damit ist die Offenheit dieser Menge gezeigt.

Für den zweiten Teil des Beweises nehmen wir an,  $B \subset Y$  sei abgeschlossen. Dann liegt der Grenzwert  $x$  jeder konvergenten Folge  $(x_k) \subset f^{-1}(B)$  in  $M$ . Es gilt also:  $x_k \rightarrow x \in M$ . Um zu zeigen, dass  $f^{-1}(B)$  abgeschlossen ist, müssen wir nun begründen, warum sogar  $x \in f^{-1}(B)$  gelten muss. Da  $x \in M$ , ist der Funktionswert  $f(x)$  definiert, und wegen der Stetigkeit von  $f$  folgt, dass

$$\lim_{k \rightarrow \infty} f(x_k) = f(x).$$

Da  $f(x_k) \in B$  für alle  $k \in \mathbb{N}$ , und da  $B$  abgeschlossen ist, so muss auch  $f(x) \in B$  liegen. Damit ist jedoch  $x \in f^{-1}(B)$  nach Definition des Urbilds. Somit ist das Urbild abgeschlossen.  $\square$

Bemerkung: Es gilt sogar die Rückrichtung: Ist das Urbild jeder offenen Menge von einer Funktion  $f$  offen, so ist  $f$  stetig. Ebenso gilt: Ist das Urbild jeder abgeschlossenen Menge von einer Funktion  $f$  abgeschlossen, so ist  $f$  stetig. Beide Eigenschaften in Satz 1.36 sind also äquivalente Charakterisierungen der Stetigkeit.

**Beispiele** Seien  $f, g : M \subset X \rightarrow \mathbb{R}$  stetige Funktionen.

- Sei  $M$  abgeschlossen. Dann ist auch die Menge

$$\{x \in M, f(x) \geq 0\} = f^{-1}([0, \infty))$$

abgeschlossen.

- Sei  $M$  offen. Dann ist auch die Menge

$$\{x \in M, f(x) < g(x)\} = (g - f)^{-1}(-\infty, 0)$$

offen.

- Sei  $M$  abgeschlossen. Dann ist auch die Menge

$$\{x \in M, f(x) \cdot g(x) = c\} = (f \cdot g)^{-1}(\{c\})$$

abgeschlossen.

## 1.6 Stetige Funktionen auf kompakten Mengen

**Definition 1.37:** Eine Teilmenge  $M \subset X$  heißt *kompakt*, wenn jede Folge aus  $M$  eine konvergente Teilfolge besitzt, deren in  $M$  liegt.

**Satz 1.38 (Heine-Borel):** In einem endlich-dimensionalen Vektorraum  $X$  ist eine Teilmenge  $M \subset X$  genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

*Beweis.* ( $\implies$ ) Wir setzen Kompaktheit voraus und müssen Abgeschlossenheit und Beschränktheit nachweisen. Sei also  $(x_n) \subset M$  eine gegen  $x \in X$  konvergente Teilfolge. Da  $M$  kompakt ist, besitzt  $x_n$  eine konvergente Teilfolge  $(x_{n_k})$  mit Grenzwert  $\tilde{x} \in M$ . Da alle Teilfolgen von konvergenten Folgen gegen den gleichen Grenzwert konvergieren – man überlege sich, warum –, muss  $\tilde{x} = x$  und somit  $x \in M$  sein.

Nehmen wir nun an,  $M$  sei nicht beschränkt. Dann existiert zu jedem  $N \in \mathbb{N}$  ein Folgenglied  $x_N$  mit  $\|x_N\| > N$ . Diese Teilfolge  $(x_N) \subset (x_n)$  kann jedoch selbst keine beschränkte, insbesondere keine konvergente Teilfolge haben. Damit existiert, im Widerspruch zur vorausgesetzten Kompaktheit, eine Folge ohne konvergente Teilfolgen.

( $\impliedby$ ) Sei  $M$  zunächst beschränkt. Dann besitzt jede Folge nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge. Ist  $M$  nun auch noch abgeschlossen, so liegt der Grenzwert dieser Teilfolge ebenfalls in  $M$ . Damit ist die Kompaktheit von  $M$  bewiesen.  $\square$

**Bemerkung 1.39:** Die Hinrichtung gilt auch in unendlich-dimensionalen Vektorräumen. Die Rückrichtung gilt jedoch im Allgemeinen nicht, da der Satz von Bolzano-Weierstraß nur in endlich-dimensionalen Vektorräumen gilt. Es gibt aber in unendlich-dimensionalen vollständigen normierten Vektorräumen ein vergleichbares Resultat: den Satz von Eberlein-Šmulian.

### Beispiele

- Für beliebige Werte  $a, b \in \mathbb{R}$  ist das abgeschlossene Intervall  $[a, b]$  kompakt.
- In endlich-dimensionalen Vektorräumen  $X$  ist die abgeschlossene Kugel  $\overline{K_r(x)} \subset X$  (insbesondere also die abgeschlossene Einheitskugel) kompakt.

**Lemma 1.40:** *Ist  $M \subset X$  eine beliebige Teilmenge und  $f : M \rightarrow X$  stetig. Ist  $K \subset M$  kompakt, so ist  $f(K)$  ebenfalls kompakt. Stetige Funktionen bilden also kompakte Mengen auf kompakte Mengen ab.*

*Beweis.* Sei  $(y_n) \subset f(K)$  eine beliebige Folge. Dann existiert nach Definition des Bildes für jedes Folgenglied  $y_n$  ein  $x_n \in K$  mit  $f(x_n) = y_n$ . Da  $K$  kompakt ist, besitzt also die Folge  $(x_n)$  eine konvergente Teilfolge  $(x_{n_k})$  mit Grenzwert  $x \in K$ . Da  $f$  stetig ist, muss dann die Folge  $(f(x_{n_k})) = (y_{n_k})$  konvergieren, und zwar gegen  $f(x) \in K$ . Somit besitzt  $(y_n) \subset f(K)$  eine in  $f(K)$  konvergente Teilfolge.  $\square$

**Satz 1.41 (Weierstraß):** *Sei  $K \subset X$  eine nicht-leere kompakte Teilmenge und  $f : K \rightarrow \mathbb{R}$  stetig. Dann besitzt die Funktion  $f$  über  $K$  ein Maximum und ein Minimum.*

*Beweis.* Sei  $f : K \rightarrow \mathbb{R}$  stetig. Nach Lemma 1.40 ist das Bild  $f(K)$  ebenfalls kompakt, insbesondere also beschränkt. Wir setzen

$$s := \sup\{f(K)\} \in \mathbb{R}.$$

Dann existiert eine Folge  $(s_n) \subset f(K)$  mit  $s_n \rightarrow s$ . Da  $f(K)$  auch abgeschlossen ist, muss  $s \in f(K)$  sein. Es gibt also ein Element  $x_0 \in K$ , sodass  $f(x_0) = s$ . Damit gilt jedoch für alle  $x \in K$ , dass

$$f(x) \leq f(x_0)$$

nach Definition des Supremums. Außerdem ist das Supremum sogar ein Maximum, da es angenommen wird.

Zum Nachweis des Minimums argumentiere analog. □

## 1.7 Zusammenhängende Mengen und Gebiete

**Definition 1.42:** Eine Menge  $M \subset X$  heißt *wegzusammenhängend*, falls zu je zwei Punkten  $a, b \in M$  eine stetige Abbildung  $\varphi : [0, 1] \subset \mathbb{R} \rightarrow M$  mit  $\varphi(0) = a$  und  $\varphi(1) = b$  existiert.

Die Abbildung  $\varphi$  heißt auch *Weg von a nach b*.

**Bemerkung 1.43:** Sei  $\varphi_1 : [0, 1] \rightarrow X$  ein Weg von  $a$  nach  $b$  und  $\varphi_2 : [0, 1] \rightarrow X$  ein Weg von  $c$  nach  $d$ , wobei  $b = c$ . Dann definiert  $\varphi_3 : [0, 1] \rightarrow X$  mit

$$\varphi_3(t) := \begin{cases} \varphi_1(2 \cdot t), & \text{für } t \in [0, 1/2] \\ \varphi_2(2 \cdot t), & \text{für } t \in (1/2, 1] \end{cases}$$

einen Weg von  $a$  nach  $d$ .

### Beispiele

1. Sei  $X$  ein endlich-dimensionaler Vektorraum und  $x, y \in X$ . Dann liefert die Abbildung  $\varphi : [0, 1] \rightarrow X$  mit  $\varphi(t) := x + t(y - x)$  die stetige Strecke zwischen  $x$  und  $y$ . Insbesondere ist jeder endlich-dimensionale Vektorraum wegzusammenhängend.
2. Die Kugel  $K_r(x) \subset X$  ist wegzusammenhängend.
3. In  $\mathbb{R}$  ist jedes Intervall wegzusammenhängend und jede wegzusammenhängende Teilmenge von  $\mathbb{R}$  ein Intervall.
4. Die Menge  $M = \{(x, y, z) \in \mathbb{R}^3, \text{ sodass } x^2 + y^2 = 1\}$  ist wegzusammenhängend.
5. Die Menge

$$\{(x, y) \in \mathbb{R}^2 \text{ mit } (x - 1)^2 + y^2 < 1\} \cup \{(x, y) \in \mathbb{R}^2 \text{ mit } (x + 1)^2 + y^2 < 1\}$$

ist nicht wegzusammenhängend.

**Satz 1.44:** Sei  $M \subset X$  und  $f : M \rightarrow Y$  stetig. Ist  $M$  wegzusammenhängend, so ist auch  $f(M)$  wegzusammenhängend.

*Beweis.* Seien  $y_1, y_2 \in f(M)$  beliebig. Dann existieren  $x_1, x_2 \in M$  mit  $f(x_1) = y_1$  und  $f(x_2) = y_2$ . Da  $M$  wegzusammenhängend ist, existiert ein stetiger Weg  $\varphi : [0, 1] \rightarrow X$  mit  $\varphi(0) = x_1$  und  $\varphi(1) = x_2$ . Für die Verknüpfung  $f \circ \varphi : [0, 1] \rightarrow Y$  gilt dann, dass  $(f \circ \varphi)(0) = y_1$  und  $(f \circ \varphi)(1) = y_2$ . Außerdem ist  $f \circ \varphi$  als Verknüpfung stetiger Funktionen stetig. Die Komposition ist also ein stetiger Weg zwischen  $y_1$  und  $y_2$ . Da die Punkte im Bild von  $f$  beliebig gewählt waren, ist  $f(M)$  wegzusammenhängend.  $\square$

**Definition 1.45:** Eine nicht-leere offene wegzusammenhängende Teilmenge  $\Omega \subset \mathbb{R}^n$  heißt *Gebiet*.

**Bemerkung 1.46:** In einem Gebiet lassen sich je zwei Punkte durch einen Polygonzug verbinden. Das Aneinandersetzen funktioniert wie in der Bemerkung 1.43: Die Punkte  $x_1, x_2, \dots, x_n$  lassen sich durch

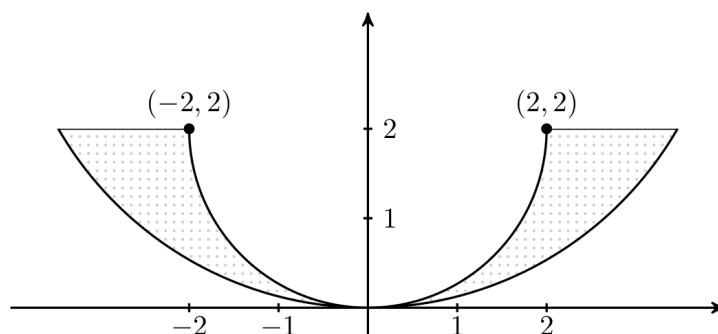
$$\varphi(t) := \begin{cases} x_1 + nt(x_2 - x_1) & \text{für } t \in [0, \frac{1}{n}] \\ x_2 + (nt - 1)(x_3 - x_2) & \text{für } t \in (\frac{1}{n}, \frac{2}{n}] \\ \vdots & \\ x_{n-1} + (nt - (n-1))(x_n - x_{n-1}) & \text{für } t \in (\frac{n-1}{n}, 1] \end{cases}$$

verbinden.

**Beispiel** In der Menge

$$M = \{(x, y) \in \mathbb{R}^2, \|(x, y) - (0, 4)\|_2 \leq 4 \wedge \|(x, y) - (0, 2)\|_2 \geq 2 \wedge 0 \leq y \leq 2\}$$

gibt es einen Weg zwischen den Punkten  $(-2, 2)$  und  $(2, 2)$ , der ganz in  $M$  liegt. Es gibt jedoch keinen endlichen Polygonzug, also keinen Polygonzug durch endlich viele Punkte, der die beiden Punkte verbindet. Wir verzichten ihr auf einen Beweis und verweisen auf die Anschaulichkeit der folgenden Illustration.



Für einen beliebigen Punkt  $x \in M$  ist die Strecke von  $x$  zum Ursprung  $(0, 0)$  zwangsläufig eine Sekante des inneren Kreises, da dessen Krümmung für  $x$  nahe bei Null größer ist als die des äußeren Kreises. Deshalb kann der Ursprung durch einen Polygonzug durch Punkte innerhalb der Menge nicht erreicht werden.

# Kapitel 2

## Differentialrechnung in $\mathbb{R}^n$

### 2.1 Totale Differenzierbarkeit und Ableitungen

In diesem Kapitel sei  $\Omega \subset \mathbb{R}^n$  stets ein Gebiet. Der Vektorraum  $\mathbb{R}^n$  sei standardmäßig mit der Euklidischen Norm und dem Euklidischen Skalarprodukt versehen.

**Definition 2.1 (Umgebung):** Unter einer *Umgebung* eines Punktes  $x \in \Omega$  verstehen wir eine offene Menge  $U$ , die  $x$  enthält.

**Definition 2.2 (Differenzierbarkeit):** Sei  $f : \Omega \rightarrow \mathbb{R}^m$  eine Abbildung und  $U$  eine Umgebung von  $a \in \Omega$ . Dann heißt  $f$  (*total*) *differenzierbar in*  $a \in \Omega$ , wenn eine lineare Abbildung  $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$  existiert, sodass

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0 \quad (2.1)$$

gilt. Alternativ muss für jedes  $x \in U$  die Gleichheit

$$f(x) = f(a) + L(x-a) + R(x-a) \quad (2.2a)$$

gelten, wobei

$$\lim_{x \rightarrow a} \frac{R(x-a)}{\|x-a\|} = 0. \quad (2.2b)$$

Die Funktion  $f$  ist also in  $a \in \Omega$  differenzierbar, wenn sie in einer Umgebung von  $a$  durch die affine Abbildung

$$x \mapsto f(a) + L(x-a)$$

approximiert werden kann.

Die Funktion  $f$  heißt *differenzierbar in*  $\Omega$ , wenn  $f$  für alle  $a \in \Omega$  differenzierbar ist.

Die Abbildung  $L$  ist durch die Bedingung (2.1) eindeutig festgelegt und heißt *Ableitung von*  $f$  *an der Stelle*  $a$ . Statt  $L$  schreiben wir auch  $df(a)$ ,  $f'(a)$  oder  $Df(a)$ .

Bemerkung: Man erkenne, dass  $h \in \mathbb{R}^n$  hier ein Vektor ist. Außerdem soll noch bemerkt werden, dass eine Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  genau dann differenzierbar ist, wenn all ihre Komponentenfunktionen  $f_1, \dots, f_m$  differenzierbar sind. Diese Tatsache bedarf streng genommen eines Beweises; es soll uns aber genügen, sie an dieser Stelle zu akzeptieren.

### Beispiele

1. Eine affine Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $f(x) := Ax + c$  ist überall differenzierbar. Die Ableitung ist für alle  $a \in \Omega$  gegeben durch  $f'(a) = A$ , da

$$f(x) = f(a) + A(x - a),$$

sodass  $R = 0$ .

2. Sei  $A$  eine symmetrische Matrix. Dann ist eine quadratische Form  $Q : \mathbb{R}^n \rightarrow \mathbb{R}$  mit

$$Q(x) := \langle x, Ax \rangle$$

überall differenzierbar. Aufgrund der Symmetrie von  $A$  und des Skalarprodukts ist

$$\begin{aligned} Q(a+h) &= \langle a+h, A(a+h) \rangle = \langle a, A(a) \rangle + \langle h, A(a) \rangle + \langle a, A(h) \rangle + \langle h, A(h) \rangle \\ &= Q(a) + 2\langle h, A(a) \rangle + \langle h, A(h) \rangle. \end{aligned}$$

Somit ist die Ableitung von  $Q$  durch

$$Q'(a) = 2\langle A(a), \cdot \rangle$$

gegeben. Da das Skalarprodukt bilinear ist, ist das tatsächlich eine lineare Abbildung (von  $\mathbb{R}^n$  nach  $\mathbb{R}$ ). Außerdem gilt für  $R(h) := \langle h, A(h) \rangle$  unter Verwendung der Cauchy-Schwarzschen Ungleichung

$$\frac{R(h)}{\|h\|} = \frac{\langle h, Ah \rangle}{\|h\|} \leq \frac{\|h\| \|Ah\|}{\|h\|} \leq \|A\| \|h\| \implies \lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0.$$

3. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  die Abbildung, die mit  $x = (x_1, x_2) \in \mathbb{R}^2$  durch

$$f(x) := \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 \\ x_2 \end{pmatrix}$$

definiert ist. Dann ist  $f$  differenzierbar auf  $\mathbb{R}^2$ . Die Ableitung an  $x$  ist durch

$$df(x) := \begin{pmatrix} 2x_1 & 2x_2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

gegeben. Es gilt nämlich

$$\begin{aligned} \frac{\|R(h)\|}{\|h\|} &= \|f(x+h) - f(x) - df(x)h\| \\ &= \left\| \begin{pmatrix} (x_1+h_1)^2 + (x_2+h_2)^2 \\ x_1+h_1 \\ x_2+h_2 \end{pmatrix} - \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 2x_1 & 2x_2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\| \end{aligned}$$



$$\begin{aligned}
 &= \left\| \begin{pmatrix} (x_1 + h_1)^2 + (x_2 + h_2)^2 \\ x_1 + h_1 \\ x_2 + h_2 \end{pmatrix} - \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 2x_1h_1 + 2x_2h_2 \\ h_1 \\ h_2 \end{pmatrix} \right\| \\
 &= \left\| \begin{pmatrix} h_1^2 + h_2^2 \\ 0 \\ 0 \end{pmatrix} \right\|.
 \end{aligned}$$

Die rechte Seite strebt nun für  $h \rightarrow 0$  gegen Null, also strebt auch die linke Seite gegen Null. Damit ist Bedingung (2.2) erfüllt.

**Bemerkung:** Es sei nochmals betont, dass die Ableitung einer Funktion  $f$  an einem Punkt  $a$  eine lineare Abbildung ist. Die übliche Schreibweise  $df(a)$  kann etwas irreführend sein, denn nicht etwa  $df$  ist linear in  $a$ , sondern „ $df(a)$ “ ist eine lineare Abbildung über den Richtungen, auf die sie angewendet wird. Es gilt also

$$df(a)(\lambda h_1 + \mu h_2) = \lambda df(a)(h_1) + \mu df(a)(h_2)$$

für alle  $\lambda, \mu \in \mathbb{R}$  und  $h_1, h_2$  aus dem gleichen Raum wie  $a$ . Wie bei linearen Abbildungen und Matrizen üblich, schreibt man häufig  $df(a)h$  statt  $df(a)(h)$ . Diese Schreibweise heißt aber hier nicht, dass  $df(a)$  und  $h$  multipliziert werden, sondern bedeutet vielmehr „ $df(a)$  angewendet auf  $h$ “. Für reellwertige Funktionen ist dieser Unterschied vernachlässigbar, ähnlich wie der Unterschied zwischen linearen Abbildungen und Matrizen. In allgemeinen normierten Vektorräumen, sogar schon bei vektorwertigen Funktionen, muss aber mehr Sorgfalt walten.

## 2.2 Partielle Differenzierbarkeit

Viele Funktionen sind nicht *total*, sondern nur in bestimmte Richtungen differenzierbar. Der Differentialquotient (der Grenzwert) existiert also nicht für beliebiges  $h \in \mathbb{R}^n$  mit  $\|h\| \rightarrow 0$  sondern lediglich für bestimmte Komponenten. Das führt auf die Definition der partiellen Differenzierbarkeit.

**Definition 2.3 (Partielle Differenzierbarkeit):** Sei  $f : \Omega \rightarrow \mathbb{R}$  eine Abbildung,  $a \in \Omega$  ein Punkt und  $v = (v_1, \dots, v_n) \in \mathbb{R}^n$  eine Richtung. Wir betrachten die Funktion  $\varphi_v : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$\varphi_v(t) := f(a + tv).$$

Falls die Abbildung  $\varphi_v$  an der Stelle  $t = 0$  differenzierbar ist, so nennen wir

$$D_v f(a) := \frac{d}{dt} \varphi_v(0) := \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} \quad (2.3)$$

die *Richtungsableitung von  $f$  an  $a$  in Richtung  $v$* .

Die partiellen Ableitungen von  $f$  sind nun die Richtungsableitungen in die Einheitsrichtungen  $e_k$ . Die Vektoren  $e_k$  sind also die kanonischen Einheitsvektoren, die an der  $k$ -ten Stelle eine 1 und ansonsten nur Nulleinträge haben. Die  $k$ -te partielle Ableitung von  $f$  ist also durch

$$D_{e_k} f(a) := \frac{\partial f}{\partial x_k}(a) := \lim_{t \rightarrow 0} \frac{f(a + te_k) - f(a)}{t} \quad (2.4)$$

gegeben. Statt  $\frac{\partial f}{\partial x_k}$  schreiben wir auch  $\partial_k f$ . Bezeichnen wir einen Punkt in  $\mathbb{R}^3$  mit  $(x, y, z)$ , so schreiben wir auch  $\partial_x f$ ,  $\partial_y f$  und  $\partial_z f$ .

**Satz 2.4:** Ist  $f : \Omega \rightarrow \mathbb{R}^m$  an der Stelle  $a \in \Omega$  differenzierbar, so ist  $f$  dort stetig.

*Beweis.* Da  $f$  an  $a$  differenzierbar ist, gilt

$$f(a+h) = f(a) + df(a)h + R(h),$$

wobei  $R$  die Eigenschaft (2.2b) erfüllt. Da  $\Omega$  ein Gebiet ist, also offen, gibt es eine offene Kugel  $K_\delta(a) \subset \Omega$  für ein gewissen  $\delta > 0$ . Außerdem lässt sich o. B. d. A. annehmen, dass

$$\frac{R(h)}{\|h\|} \leq 1$$

für alle  $h \in \mathbb{R}^n$  mit  $0 < \|h\| < \delta$ . (Ansonsten wähle ein kleineres  $\delta > 0$ . Damit gilt

$$\begin{aligned} \|f(a+h) - f(a)\| &= \|df(a)h + R(h)\| \leq \|df(a)\| \|h\| + \|R(h)\| \\ &\leq \|df(a)\| \|h\| + \|h\| = (\|df(a)\| + 1) \|h\| \xrightarrow{h \rightarrow 0} 0. \end{aligned}$$

Damit ist alles gezeigt. □

**Satz 2.5:** Ist  $f : \Omega \rightarrow \mathbb{R}^m$  in einem Punkt  $a \in \Omega$  differenzierbar, so existiert die Richtungsableitung von  $f$  in  $a$  in jede Richtung  $v$ . Insbesondere sind die Komponentenfunktionen  $f_1, \dots, f_m$  partiell differenzierbar. Da die Ableitung  $df(a)$  eine lineare Abbildung ist, gibt es eine dazugehörige Matrix  $J_f(a)$ , sodass

$$df(a)(v) = J_f(a)v \tag{2.5}$$

für alle  $v \in \mathbb{R}^n$  gilt. Diese Matrix heißt *Jacobi-Matrix* von  $f$  an der Stelle  $a$ . Es gilt

$$J_f(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} = \left( \frac{\partial f_i}{\partial x_k}(a) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}} = \left( \partial_k f_i(a) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}}. \tag{2.6}$$

Insbesondere ist die Ableitung von  $f$  in  $a$  eindeutig bestimmt.

*Beweis.* Sei  $f$  in  $a$  differenzierbar. Dann gilt Bedingung (2.1) für alle  $h \in \mathbb{R}^n$  mit  $h \rightarrow 0$ . Wir wählen  $h := tv$ , sodass  $tv \rightarrow 0 \iff t \rightarrow 0$ . Damit gilt

$$0 = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a) - df(a)(tv)}{\|tv\|} = \frac{1}{\|v\|} \lim_{t \rightarrow 0} \left\| \frac{f(a+tv) - f(a)}{t} - df(a)(v) \right\|,$$

wobei wir hier die Linearität der Ableitung ausgenutzt haben. Insbesondere existiert der Limes auf der rechten Seite, und wir haben

$$D_v f(a) = df(a)(v). \tag{2.7}$$

Dies gilt für beliebige Richtungen, insbesondere also für die Einheitsvektoren, also die partiellen Ableitungen. Deswegen ist

$$df(a)(e_k) = \frac{\partial f}{\partial x_k}(a) = \left( \frac{\partial f_i}{\partial x_k}(a) \right)_{1 \leq i \leq m} = \sum_{i=1}^m \frac{\partial f_i}{\partial x_k}(a) e'_i, \quad (2.8)$$

wobei  $e_1, \dots, e_n$  die  $n$  kanonischen Einheitsvektoren in  $\mathbb{R}^n$  und  $e'_1, \dots, e'_m$  die  $m$  kanonischen Einheitsvektoren in  $\mathbb{R}^m$  sind.

Dass zu jeder linearen Abbildung eine Matrix mit der Eigenschaft (2.5) existiert, ist aus der linearen Algebra bekannt. Dass die Matrix durch die Formel (2.6) gegeben ist, ist durch Multiplikation der Matrix mit den Einheitsvektoren ersichtlich.  $\square$

### Beispiele

1. Sei  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  gegeben durch  $f(x, y, z) := x^2 \exp(y) + \sin(z)$ . Dann lauten die partiellen Ableitungen

$$\frac{\partial f}{\partial x}(x, y, z) = 2x \exp(y), \quad \frac{\partial f}{\partial y}(x, y, z) = x^2 \exp(y) \quad \text{und} \quad \frac{\partial f}{\partial z}(x, y, z) = \cos(z).$$

2. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  gegeben durch

$$f(x, y) := \begin{pmatrix} x^2 + y^2 \\ x \\ y \end{pmatrix}.$$

Dann lauten die partiellen Ableitungen

$$\begin{aligned} \frac{\partial f_1}{\partial x}(x, y) &= 2x, & \frac{\partial f_1}{\partial y}(x, y) &= 2y, \\ \frac{\partial f_2}{\partial x}(x, y) &= 1, & \frac{\partial f_2}{\partial y}(x, y) &= 0, \\ \frac{\partial f_3}{\partial x}(x, y) &= 0, & \frac{\partial f_3}{\partial y}(x, y) &= 1. \end{aligned}$$

Konsequenterweise ist daher die Jacobi-Matrix gegeben durch

$$J_f(a) = \begin{pmatrix} 2x & 2y \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

**Bemerkung 2.6:** Für eine differenzierbare Funktion  $f : \Omega \rightarrow \mathbb{R}$  erhalten wir für alle  $h \in \mathbb{R}^n$

$$df(a)h = \frac{\partial f}{\partial x_1}(a)h_1 + \dots + \frac{\partial f}{\partial x_n}(a)h_n.$$

Bemerke jedoch, dass die Existenz der partiellen Ableitungen allein nicht die Existenz der totalen Ableitung garantiert!

**Beispiel** Sei die Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch

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

Diese Funktion ist in  $(0, 0)$  partiell, jedoch nicht total differenzierbar. Bilden wir die partiellen Ableitungen, ergeben sich

$$\frac{\partial g}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{g(t, 0) - g(0, 0)}{t} = 0$$

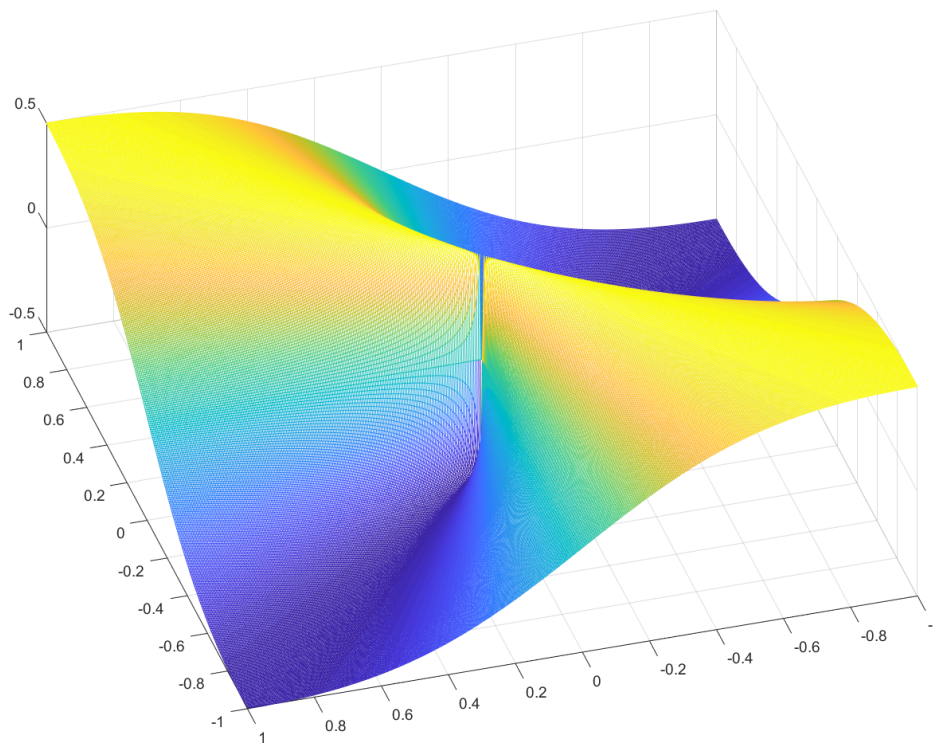
und

$$\frac{\partial g}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{g(0, t) - g(0, 0)}{t} = 0.$$

Wäre die Funktion  $g$  in  $(0, 0)$  differenzierbar, wäre sie dort nach Satz 2.4 insbesondere stetig. Allerdings ist

$$\frac{1}{2} = \lim_{n \rightarrow 0} g\left(\frac{1}{n}, \frac{1}{n}\right) \neq \lim_{n \rightarrow 0} g\left(\frac{1}{n}, 0\right) = 0,$$

was der Stetigkeit von  $g$  widerspricht.



Bemerkung: Die totale Differenzierbarkeit liefert Richtungs-differenzierbarkeit in *alle* Richtungen, nicht nur in Einheitsrichtungen.

## 2.3 Das Hauptkriterium für Differenzierbarkeit

**Satz 2.7:** *Existieren alle partiellen Ableitungen einer Abbildung  $f : \Omega \rightarrow \mathbb{R}^m$  an allen  $x \in \Omega$ , und sind diese partiellen Ableitungen stetige Funktionen in  $\Omega$ , so ist  $f$  in ganz  $\Omega$  (total) differenzierbar und es gilt*

$$df(x) = \left( \frac{\partial f_i}{\partial x_k}(x) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}}. \quad (2.9)$$

*Beweisskizze.* Zunächst genügt es, diese Aussage für reellwertige Funktionen zu zeigen, da die Aussage dann für alle Komponentenfunktionen von  $f$  und damit für  $f$  selbst gilt.

Sei also  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine Abbildung, deren partielle Ableitungen auf ganz  $\Omega$  existieren. Da  $\Omega$  offen ist, gibt es für  $a \in \Omega$  eine Kugel  $K_\delta(a) \subset \Omega$ . Sei nun  $h = (h_1, \dots, h_n) \in K_\delta(a)$ . Dann gilt

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2, \dots, a_n) - f(a_1, \dots, a_n) \\ &\quad + f(a_1 + h_1, a_2 + h_2, a_3, \dots, a_n) - f(a_1 + h_1, a_2, \dots, a_n) \\ &\quad \vdots \\ &\quad + f(a_1 + h_1, \dots, a_n + h_n) - f(a_1 + h_1, \dots, a_{n-1} + h_{n-1}, a_n). \end{aligned}$$

Wir setzen  $\varphi_1(x) := f(x, a_2, \dots, a_n)$ . Nach Voraussetzung existiert die Ableitung

$$\varphi_1'(x) = \frac{\partial f}{\partial x_1}(x, a_2, \dots, a_n)$$

von  $\varphi_1$  für alle Werte zwischen  $a_1$  und  $a_1 + h$ . Daher können wir den Mittelwertsatz der Differentialrechnung in  $\mathbb{R}$  auf  $\varphi_1$  anwenden (vgl. Analysis I) und erhalten

$$f(a_1 + h, a_2, \dots, a_n) - f(a_1, \dots, a_n) = h_1 \cdot \frac{\partial f}{\partial x_1}(\xi_1, a_2, \dots, a_n).$$

Dieser Schritt lässt sich nun auch analog für alle anderen Komponenten von  $a$  durchführen. Es lässt sich daher

$$f(a+h) - f(a) = \sum_{k=1}^n h_k \cdot \frac{\partial f}{\partial x_k}(a) + \sum_{k=1}^n h_k \cdot \left( \frac{\partial f}{\partial x_k}(c^{(k)}) - \frac{\partial f}{\partial x_k}(a) \right)$$

schreiben, wobei

$$c^{(k)} := (a_1 + h_1, \dots, a_{k-1} + h_{k-1}, \xi_k, a_{k+1}, \dots, a_n).$$

ist. Die erste Summe wird nun definiert als der lineare Operator  $L(h)$ , der zweite als  $R(h)$ . Für diesen Term müssen wir die Bedingung (2.2b) nachweisen. Mit der Cauchy-Schwarzschen Ungleichung gilt nun

$$|R(h)| := \sum_{k=1}^n h_k \cdot \left( \frac{\partial f}{\partial x_k}(c^{(k)}) - \frac{\partial f}{\partial x_k}(a) \right) \leq \|h\| \cdot \|b\|, \quad (*)$$

wobei  $b \in \mathbb{R}^n$  ist und die Komponenten durch

$$b_k := \frac{\partial f}{\partial x_k}(c^{(k)}) - \frac{\partial f}{\partial x_k}(a)$$

gegeben sind. Da die partiellen Ableitungen nach Voraussetzung stetig sind, geht mit  $h \rightarrow 0$  auch jedes  $c^{(k)} \rightarrow a$ . Bemerke, dass die  $\xi_k \rightarrow a_k$  für alle  $k = 1, \dots, n$ , da jedes  $\xi_k$  auf der Strecke zwischen  $a_k$  und  $a_k + h_k$  liegt. Deswegen gilt

$$\frac{\partial f}{\partial x_k}(c^{(k)}) \xrightarrow{h \rightarrow 0} \frac{\partial f}{\partial x_k}(a),$$

weswegen  $b_k \rightarrow 0$  für alle  $k$ , mithin  $\|b\| \rightarrow 0$ . Es folgt aus (\*), dass

$$\lim_{h \rightarrow 0} \frac{|R(h)|}{\|h\|} = 0.$$

Damit ist alles gezeigt. □

**Beispiel** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $f(x, y) := (\exp(x) \cos y, \exp(x) \sin(y))^T$  gegeben. Beide Komponentenfunktionen sind partiell differenzierbar, wobei

$$\begin{aligned} \frac{\partial f_1}{\partial x}(x, y) &= \exp(x) \cos y & \frac{\partial f_1}{\partial y}(x, y) &= -\exp(x) \sin y \\ \frac{\partial f_2}{\partial x}(x, y) &= \exp(x) \sin y & \frac{\partial f_2}{\partial y}(x, y) &= \exp(x) \cos y. \end{aligned}$$

Alle partiellen Ableitungen sind überall stetig. Somit ist  $f$  total differenzierbar und es gilt

$$df(x, y) = \begin{pmatrix} \exp(x) \cos y & -\exp(x) \sin y \\ \exp(x) \sin y & \exp(x) \cos y \end{pmatrix}.$$

## 2.4 Rechenregeln für differenzierbare Funktionen

### 2.4.1 Linearität der Ableitung

Wie wir bereits gesehen haben, ist die Ableitung einer Funktion eine lineare Abbildung, die einer Richtung, also einem Vektor  $h \in \mathbb{R}^n$ , einen Wert oder einen Vektor zuordnet. Wir wissen also schon, dass die Ableitung (oder auch: das Differential) einer Funktion an einem Punkt linear ist. Auf der anderen Seite lässt sich auch untersuchen, ob der Differentialoperator, also der Operator, der einer Funktion  $f$  ihre Ableitung  $f'(a)$  an einem Punkt zuordnet, linear ist. Und tatsächlich ist das ja auch schon aus der Schule bekannt. Sind nämlich  $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  zwei an der Stelle  $a \in \Omega$  differenzierbare Funktionen, so ist auch  $\alpha f + \beta g : \Omega \rightarrow \mathbb{R}^m$  für alle  $\alpha, \beta \in \mathbb{R}$  eine in  $a$  differenzierbare Funktion. Es gilt:

$$d(\alpha f + \beta g)(a) = \alpha df(a) + \beta dg(a). \tag{2.10}$$

### 2.4.2 Kettenregel

Seien zwei Abbildungen  $g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $f : \tilde{\Omega} \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$  gegeben, wobei  $g(\Omega) \subset \tilde{\Omega}$  sein soll. Ist  $g$  an  $x \in \Omega$  und  $f$  an  $y := g(x) \in \tilde{\Omega}$  differenzierbar, so ist die Komposition  $f \circ g : \Omega \rightarrow \mathbb{R}^p$  an  $x$  differenzierbar und es gilt

$$d(f \circ g)(x) = df(y) \circ dg(x). \quad (2.11)$$

Für die partiellen Ableitungen für  $k = 1, \dots, n$  gilt

$$\frac{\partial(f \circ g)}{\partial x_k}(x) = \sum_{j=1}^m \frac{\partial f}{\partial y_j}(g_1(x), \dots, g_m(x)) \cdot \frac{\partial g_j}{\partial x_k}(x). \quad (2.12)$$

#### Beispiele

1. Seien  $g : [0, T] \rightarrow \mathbb{R}^3$  und  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  zwei Abbildungen mit

$$g(t) := \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix} \quad \text{und} \quad f(x, y, z) := x + yz.$$

Definiere zunächst  $h : [0, T] \rightarrow \mathbb{R}$  als  $h := f \circ g$ . Dann ist  $h$  gegeben durch

$$h(t) := \cos t + t \sin t.$$

Die Ableitung von  $h$  lässt sich dann mithilfe der Produktregel der Differentialrechnung in  $\mathbb{R}$  bestimmen. Sie lautet.

$$h'(t) = -\sin t + \sin t + t \cos t = t \cos t.$$

Bestimmen wir die Ableitung der Verknüpfung  $f \circ g : [0, T] \rightarrow \mathbb{R}$  an  $t \in [0, T]$  direkt mithilfe der Kettenregel, so erhalten wir

$$d(f \circ g)(t) = df(g(t)) \cdot dg(t) = \begin{pmatrix} 1 \\ t \\ \sin t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix} = t \cos t.$$

Beide Ableitungen stimmen also tatsächlich überein. Bemerke, dass die Multiplikation hier Skalarmultiplikation bedeutet.

2. Seien  $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  und  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  zwei Abbildungen mit

$$g(x, y, z) := \begin{pmatrix} x^2 y \\ y^2 \\ \exp -xz \end{pmatrix} \quad \text{und} \quad f(u, v, w) := u^2 + v^2 - w.$$

Wir bezeichnen die Komponentenfunktionen von  $g$  wie üblich. Die Ableitungen von  $f$  und  $g$  sind gegeben durch

$$df(u, v, w) = \left( \frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w} \right)^T = (2u, 2v, -1)^T$$

sowie

$$dg(x, y, z) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} & \frac{\partial g_1}{\partial z} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} & \frac{\partial g_2}{\partial z} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial y} & \frac{\partial g_3}{\partial z} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2y & 0 \\ -z \exp(-xz) & 0 & -x \exp(-xz) \end{pmatrix}.$$

Die Ableitung der Verknüpfung  $f \circ g$  ergibt sich damit als

$$\begin{aligned} df(g(x), g(y), g(z)) \cdot dg(x, y, z) &= \begin{pmatrix} 2x^2y \\ 2y^2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2y & 0 \\ -z \exp(-xz) & 0 & -x \exp(-xz) \end{pmatrix} \\ &= \begin{pmatrix} 4x^3y^2 + z \exp(-xz) \\ 2x^4y + 4y^3 \\ x \exp(-xz) \end{pmatrix}. \end{aligned}$$

Hierbei ist die Multiplikation von wie eine spaltenweise Skalarmultiplikation zu verstehen.

3. Seien die Funktionen  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$g(u, v) := \begin{pmatrix} \exp(u^2) \\ u - \sin v \end{pmatrix} \quad \text{und} \quad f(x, y) := \begin{pmatrix} \cos y + x^2 \\ \exp(x + y) \end{pmatrix}$$

gegeben. Die Ableitung der Verknüpfung  $f \circ g$  an  $(0, 0)$  ergibt sich dann aus

$$df(x, y) = \begin{pmatrix} 2x & -\sin y \\ \exp(x + y) & \exp(x + y) \end{pmatrix} \quad \text{und} \quad dg(x, y) = \begin{pmatrix} 2u \exp(u^2) & 0 \\ 1 & -\cos v \end{pmatrix}$$

als

$$d(f \circ g)(0, 0) = df(g(0, 0)) \cdot dg(0, 0) = \begin{pmatrix} 2 & 0 \\ e & e \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ e & -e \end{pmatrix}$$

da  $g(0, 0) = (1, 0)^T$ .

### 2.4.3 Produkt- und Quotientenregel

Seien zwei Abbildungen  $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  gegeben, die an  $a \in \Omega$  differenzierbar sind. Dann ist auch das punktweise Produkt  $f \cdot g : \Omega \rightarrow \mathbb{R}$  in  $a$  differenzierbar und es gilt

$$d(f \cdot g)(a) = f(a)dg(a) + g(a)df(a). \quad (2.13)$$

Bemerke, dass  $f(a), g(a) \in \mathbb{R}$  und  $df(a), dg(a) \in \mathbb{R}^n$ .



Falls  $g(a) \neq 0$ , so ist auch der punktweise Quotient  $\frac{f}{g}$  an  $a$  differenzierbar, wobei

$$d\left(\frac{f}{g}\right)(a) = \frac{g(a)df(a) - f(a)dg(a)}{(g(a))^2} \quad (2.14)$$

gilt.

Die Produktregel lässt sich auf die Kettenregel zurückführen. Ist nämlich  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  und  $G : \mathbb{R}^n \rightarrow \mathbb{R}^2$  mit

$$F(u, v) := uv \quad \text{und} \quad G(x) := \begin{pmatrix} f(x) \\ g(x) \end{pmatrix},$$

sodass

$$dF(u, v) = (v, u)^T \quad \text{und} \quad dG(x) = \begin{pmatrix} df(x) \\ dg(x) \end{pmatrix} \in M(2, n),$$

so gilt  $f \cdot g = F \circ G$  nach Definition des punktweisen Produkts von Funktionen. Insbesondere ist

$$\begin{aligned} d(fg)(x) &= d(F \circ G)(x) \\ &= dF(G(x)) \cdot dG(x) = \begin{pmatrix} g(x) \\ f(x) \end{pmatrix} \cdot \begin{pmatrix} df(x) \\ dg(x) \end{pmatrix} = g(x)f'(x) + f(x)g'(x). \end{aligned}$$

Ganz analog lässt sich auch die Quotientenregel auf die Kettenregel zurückführen. Definiere hierfür  $H : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $H(u, v) := \frac{u}{v}$ , sodass

$$dH(u, v) = \left(\frac{1}{v}, -\frac{u}{v^2}\right)^T,$$

und leite anschließend die Verknüpfung  $H \circ G$  ab. Es ergibt sich,

$$\begin{aligned} d\left(\frac{f}{g}\right)(x) &= d(H \circ G)(x) = dH(G(x)) \cdot dG(x) \\ &= \begin{pmatrix} \frac{1}{g(x)} \\ -\frac{f(x)}{g(x)^2} \end{pmatrix} \cdot \begin{pmatrix} df(x) \\ dg(x) \end{pmatrix} \\ &= \frac{df(x)}{g(x)} - \frac{f(x)dg(x)}{g(x)^2} = \frac{g(x)df(x) - f(x)dg(x)}{g(x)^2}. \end{aligned}$$

## 2.5 Stetige Differenzierbarkeit

**Definition 2.8:** Eine Abbildung  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  heißt *stetig differenzierbar*, wenn alle partiellen Ableitungen  $\frac{\partial f_k}{\partial x_i}(x)$  für alle  $x \in \Omega$  existieren und stetig sind. Wir definieren

$$C^1(\Omega, \mathbb{R}^n) := \left\{ f : \Omega \rightarrow \mathbb{R}^n, \text{ sodass } f \text{ stetig differenzierbar auf } \Omega \right\}. \quad (2.15)$$

Bemerke, dass Abbildungen  $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$  nach Satz 2.7 differenzierbar sind. Ist  $n = 1$ , schreiben wir auch  $\mathcal{C}^1(\Omega)$ .

Analog lassen sich (rekursiv) stetig differenzierbare Funktionen höherer Ordnung definieren: Eine Abbildung heißt  $k$ -mal stetig differenzierbar ( $f \in \mathcal{C}^k(\Omega, \mathbb{R}^n)$ ), wenn  $f \in \mathcal{C}^{k-1}(\Omega, \mathbb{R}^n)$  und alle partiellen Ableitungen  $k$ -ter Ordnung von  $f$  stetig sind. Wir schreiben

$$\mathcal{C}^\infty(\Omega, \mathbb{R}^n) := \bigcap_{k=1}^{\infty} \mathcal{C}^k(\Omega, \mathbb{R}^n) \quad (2.16)$$

für beliebig oft stetig differenzierbare Funktionen. Wir erweitern die Schreibweise auch für stetige Funktionen und bezeichnen den Raum der stetigen Funktionen von  $\Omega$  nach  $\mathbb{R}^n$  mit  $\mathcal{C}(\Omega, \mathbb{R}^n) := \mathcal{C}^0(\Omega, \mathbb{R}^n)$ .

Bemerke, dass vektorwertige Funktionen  $f : \Omega \rightarrow \mathbb{R}^m$  genau dann  $k$ -mal stetig differenzierbar sind, wenn alle Komponentenfunktionen  $f_i : \Omega \rightarrow \mathbb{R}$   $k$ -mal stetig differenzierbar sind.

**Bemerkung 2.9:** Für alle  $k \in \mathbb{N} \cup \{0\}$  ist  $\mathcal{C}^k(\Omega, \mathbb{R}^n)$  ein reeller Vektorraum.

Offensichtlich gilt  $\mathcal{C}^\infty(\Omega, \mathbb{R}^n) \subset \dots \subset \mathcal{C}^1(\Omega, \mathbb{R}^n) \subset \mathcal{C}^0(\Omega, \mathbb{R}^n)$ .

**Satz 2.10 (Schwarz):** Sei  $\Omega \subset \mathbb{R}^2$  und  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ . Dann gilt für alle  $i, k = 1, \dots, n$

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_k} \right) = \frac{\partial}{\partial x_k} \left( \frac{\partial f}{\partial x_i} \right). \quad (2.17)$$

Die partiellen Ableitungen sind also unabhängig von der Reihenfolge des Differenzierens.

Bemerkung: Auch wenn der Satz hier nicht bewiesen wird, sei darauf hingewiesen, dass zur Anwendung des Satzes von Schwarz die totale Differenzierbarkeit nicht ausreicht! Man benötigt die Stetigkeit der Ableitungen für den Beweis.

Die Aussage lässt sich leicht induktiv für partielle Ableitungen höherer Ordnung zeigen. Auch hierfür wird die Stetigkeit der partiellen Ableitung höchster Ordnung benötigt.

**Beispiel** Sei  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  durch  $f(x, y, z) := \exp(xy) + z \cos x$  definiert. Die partiellen Ableitungen

$$\frac{\partial f}{\partial x}(x, y, z) = y \exp(xy) - z \sin x, \quad \frac{\partial f}{\partial y}(x, y, z) = x \exp(xy), \quad \text{und} \quad \frac{\partial f}{\partial z}(x, y, z) = \cos x$$

von  $f$  sind allesamt stetig. Somit ist  $f \in \mathcal{C}^1(\mathbb{R}^3, \mathbb{R})$ . Betrachten wir nun auch die zweiten partiellen Ableitungen, so gilt

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x, y, z) &= y^2 \exp(xy) - z \cos x, & \frac{\partial^2 f}{\partial y^2}(x, y, z) &= x^2 \exp(xy), & \frac{\partial^2 f}{\partial z^2}(x, y, z) &= 0 \\ \frac{\partial^2 f}{\partial x \partial y}(x, y, z) &= \exp(xy)(1 + x), & \frac{\partial^2 f}{\partial x \partial z}(x, y, z) &= -\sin x, & \text{und} & \frac{\partial^2 f}{\partial y \partial z}(x, y, z) = 0. \end{aligned}$$

Da auch diese Ableitungen allesamt global stetig sind, ist sogar  $f \in \mathcal{C}^2(\mathbb{R}^3, \mathbb{R})$ .

## 2.6 Reellwertige Funktionen

Wie am Anfang des Kapitels gesehen, kann die Ableitung einer Funktion durch die dazugehörige Jacobi-Matrix beschrieben werden. Für reellwertige Funktionen gibt es noch eine andere Darstellungsmöglichkeit, die auf einer allgemeinen und mathematisch sehr tiefgreifenden Aussage beruht, die als *Darstellungssatz von Riesz* bekannt ist. Sie besagt, anschaulich gesprochen, dass zu jeder linearen Abbildung  $f$  über einem Hilbertraum (vollständiger Skalarproduktraum)  $X$  ein eindeutiges Element  $y \in X$  existiert, sodass

$$f(x) = \langle x, y \rangle \quad \text{für alle } x \in X$$

gilt. Betrachten wir beispielsweise lineare Funktionen in  $\mathbb{R}$ , also Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = mx$ , so lässt sich  $f$  eindeutig mit  $m$  identifizieren, denn es gilt

$$f(x) = m \cdot x = \langle m, x \rangle_{\mathbb{R}}.$$

Diese Beobachtung führt, im Kontext der Differentialrechnung von Funktionen in mehreren Variable, auf den Begriff des *Gradienten*.

**Definition 2.11 (Gradient):** Sei  $\Omega \subset \mathbb{R}^n$  und  $f \in \mathcal{C}^1(\Omega, \mathbb{R})$  eine stetig differenzierbare Funktion. Die Ableitung von  $f$  an  $a \in \Omega$  wird durch die lineare Abbildung  $df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$  beschrieben, die auch Differential genannt wird. Im vektorwertigen Fall gehört zu einer linearen Abbildung die Jacobi-Matrix. Im skalarwertigen Fall gibt es zu  $df(a)$  einen eindeutig bestimmten Vektor  $\nabla f(a) \in \mathbb{R}^n$ , sodass für alle  $h = (h_1, \dots, h_n)^T$

$$df(a)(h) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) h_k = \langle \nabla f(a), h \rangle_{\mathbb{R}^n} \quad (2.18)$$

gilt. Dieser Vektor  $\nabla f(a)$  heißt (*Vektor-*)*Gradient von  $f$  in  $a$* . Er ist durch

$$\nabla f(a) := \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \quad (2.19)$$

gegeben.

**Lemma 2.12:** Seien  $f, g \in \mathcal{C}^1(\Omega)$  und  $a \in \Omega$ . Für den Gradienten gelten folgende Rechenregeln:

1. (*Linearität*) Für alle  $\alpha, \beta \in \mathbb{R}$  gilt

$$\nabla(\alpha f + \beta g)(a) = \alpha \nabla f(a) + \beta \nabla g(a). \quad (2.20)$$

2. (*Produktregel*) Es gilt

$$\nabla(f \cdot g)(a) = f(a) \nabla g(a) + g(a) \nabla f(a). \quad (2.21)$$

3. (Quotientenregel) Für alle  $a \neq 0$  gilt

$$\nabla \left( \frac{f}{g} \right) (a) = \frac{g(a)\nabla f(a) - f(a)\nabla g(a)}{g(a)^2}. \quad (2.22)$$

Das folgt direkt aus den bereits bekannten Ableitungsregeln.

**Definition 2.13:** Sei  $I \subset \mathbb{R}$  ein Intervall. Dann heißt eine einmal stetig differenzierbare Abbildung  $\varphi : I \rightarrow \mathbb{R}^n$   $\mathcal{C}^1$ -Weg. Die Ableitung einer solchen Funktion an  $t_0 \in I$  kann durch den Tangentenvektor

$$\varphi'(t_0) := \lim_{t \rightarrow t_0} \frac{1}{t - t_0} (\varphi(t) - \varphi(t_0)) = \begin{pmatrix} \varphi_1(t_0) \\ \vdots \\ \varphi'_n(t_0) \end{pmatrix} \in \mathbb{R}^n \quad (2.23)$$

beschrieben werden.

**Bemerkung 2.14:** Im physikalischen Kontext, wenn  $I$  ein Zeitintervall beschreibt, wird statt  $\varphi'$  auch  $\dot{\varphi}$  geschrieben, um die Zeitableitung von den Ableitungen bezüglich der Raumvariablen  $x_k$  zu unterscheiden.

1. Beschreibt  $\varphi(t)$  nun beispielsweise den Ort eines Massepunktes zu einem Zeitpunkt  $t$ , so lässt sich  $\dot{\varphi}(t)$  als Geschwindigkeitsvektor des Teilchens zum Zeitpunkt  $t$  interpretieren.
2. Ist  $f \in \mathcal{C}^1(\Omega)$  und  $\varphi : I \rightarrow \Omega$  ein  $\mathcal{C}^1$ -Weg, so gilt

$$\begin{aligned} \frac{d}{dt} f(\varphi(t)) &= df(\varphi(t))d\varphi(t) = (\partial_1 f(\varphi(t)), \dots, \partial_n f(\varphi(t))) \dot{\varphi}(t) \\ &= \sum_{k=1}^n \partial_k f(\varphi(t)) \dot{\varphi}_k(t) = \langle \nabla f(\varphi(t)), \dot{\varphi}(t) \rangle. \end{aligned}$$

3. Sei  $a \in \Omega$  und fixiere einen Richtungsvektor  $v \in \mathbb{R}^n$  mit  $\|v\| = 1$ . Sei  $\varphi$  der  $\mathcal{C}^1$ -Weg mit  $\varphi(t) := a + tv$ . Betrachten wir die Ableitung der Verknüpfung  $f \circ \varphi$  an der Stelle  $t = 0$ , so ergibt sich

$$\left. \frac{d}{dt} f(\varphi(t)) \right|_{t=0} = \langle \nabla f(\varphi(0)), \dot{\varphi}(0) \rangle = \langle \nabla f(a), v \rangle = df(a)(v) = D_v f(a).$$

So lässt sich die Richtungsableitung von  $f$  an  $a$  in Richtung  $v$  auch einfach als Ableitung der Verknüpfung  $f$  mit dem  $\mathcal{C}^1$ -Weg  $\varphi$  verstehen. Insbesondere folgt

$$|D_v f(a)| = |\langle \nabla f(a), v \rangle| \leq \|\nabla f(a)\| \|v\| = \|\nabla f(a)\|.$$

Wie in Bemerkung 1.6 bemerkt, gilt hierbei Gleichheit genau dann, wenn  $v$  und  $\nabla f(a)$  linear abhängig sind, also wenn

$$v = \pm \frac{\nabla f(a)}{\|\nabla f(a)\|}.$$

Daher wird  $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$  auch *Richtung des steilsten Anstiegs* und  $v = -\frac{\nabla f(a)}{\|\nabla f(a)\|}$  *Richtung des steilsten Abstiegs* genannt.

**Beispiel (Ableitung der Norm)** Sei  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  die Euklidische Norm, also die Abbildung

$$x \mapsto \left( \sum_{k=1}^n x_k^2 \right)^{\frac{1}{2}}.$$

Dann ist die partielle Ableitung von  $\|\cdot\|$  gegeben durch

$$\frac{\partial \|\cdot\|}{\partial x_k}(x) = 2x_k \cdot \frac{1}{2} \left( \sum_{k=1}^n x_k^2 \right)^{-\frac{1}{2}} = \frac{x_k}{\|x\|}.$$

Es folgt für die Ableitung der Normfunktion

$$d(\|\cdot\|)(x) = \frac{1}{\|x\|} (x_1, \dots, x_n) = \frac{x}{\|x\|}.$$

### 2.6.1 Hauptsätze für Funktionen in mehreren Veränderlichen

**Satz 2.15 (Hauptsatz der Differential- und Integralrechnung in  $\mathbb{R}^n$ ):** Seien  $\Omega \subset \mathbb{R}^n$  und  $a, b \in \Omega$ . Seien außerdem eine stetig differenzierbare Funktion  $f \in \mathcal{C}^1(\Omega, \mathbb{R})$  und ein  $\mathcal{C}^1$ -Weg  $\varphi : [\alpha, \beta] \rightarrow \Omega$  mit  $\varphi(\alpha) = a$  und  $\varphi(\beta) = b$  gegeben. Dann gilt

$$f(b) - f(a) = \int_a^\beta \langle \nabla f(\varphi(t)), \dot{\varphi}(t) \rangle dt. \quad (2.24)$$

*Beweis.* Der Satz lässt sich mittels  $\varphi$  leicht auf den eindimensionalen Fall zurückführen. Es gilt nämlich

$$\begin{aligned} f(b) - f(a) &= f(\varphi(\alpha)) - f(\varphi(\beta)) \\ &= (f \circ \varphi)(\alpha) - (f \circ \varphi)(\beta) = \int_\alpha^\beta \frac{d}{dt} f(\varphi(t)) dt = \int_\alpha^\beta \langle \nabla f(\varphi(t)), \dot{\varphi}(t) \rangle dt, \end{aligned}$$

wobei in der zweiten Zeile der aus Analysis I bekannte Hauptsatz der Differential- und Integralrechnung verwendet wurde.  $\square$

**Satz 2.16 (Schrankensatz):** Sei  $f \in \mathcal{C}^1(\Omega)$  und  $a, b \in \Omega$ , sodass die Verbindungsstrecke  $S$  zwischen  $a$  und  $b$  ganz in  $\Omega$  liegt.<sup>1</sup> Dann gilt

$$|f(b) - f(a)| \leq M \|b - a\|, \quad \text{wobei } M := \max \{ \|\nabla f(x)\|, x \in S \}. \quad (2.25)$$

Für vektorwertige Funktionen  $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$  gilt analog

$$\|f(b) - f(a)\| \leq L \|b - a\|, \quad \text{wobei } L := \max \{ \|J_f(x)\|_F, x \in S \}. \quad (2.26)$$

Hierbei ist  $\|df(x)\| := \sup_{\|h\|=1} \|df(x)(h)\|$  die Operatornorm von  $df(x)$ .

*Bemerke, dass das Maximum in beiden Fällen wohldefiniert ist, da die Ableitung nach Voraussetzung stetig ist, und deswegen das Maximum auf der kompakten Menge  $S$  annimmt. Insbesondere ist  $f$  Lipschitz-stetig, wenn die Ableitung beschränkt ist.*

<sup>1</sup>Mathematisch ausgedrückt:  $S := \{tb + (1-t)a \in \Omega \text{ für alle } t \in [0, 1]\} \subset \Omega$ .

*Beweis.* Die Strecke  $S$  lässt sich durch den Weg  $\varphi(t) := a + t(b - a)$  mit  $t \in [0, 1]$  beschreiben. Dann gilt nach dem Hauptsatz 2.15 für reellwertiges  $f$

$$\begin{aligned} |f(b) - f(a)| &= \left| \int_0^1 \langle \nabla f(\varphi(t)), \varphi'(t) \rangle dt \right| \\ &\leq \int_0^1 |\langle \nabla f(\varphi(t)), b - a \rangle| dt \\ &\leq \int_0^1 \|\nabla f(\varphi(t))\|_2 \|b - a\|_2 dt \leq M \|b - a\|. \end{aligned}$$

Falls  $f$  vektorwertig ist, so gilt die erste Aussage für alle Komponentenfunktionen  $f_i$ . Es folgt

$$\begin{aligned} \|f(b) - f(a)\|_2^2 &= \sum_{i=1}^m |f_i(b) - f_i(a)|_2^2 \\ &\leq \sum_{i=1}^m \left( \int_0^1 \|\nabla f_i(\varphi(t))\|_2^2 dt \right) \left( \int_0^1 \|b - a\|_2^2 dt \right) \quad (\text{CSU}) \\ &\leq \|b - a\|_2^2 \int_0^1 \|J_f(\varphi(t))\|_F^2 dt \leq M^2 \|b - a\|_2^2. \end{aligned}$$

Bemerke, dass die Gradienten  $\nabla f_i(x)$  die Spalten der Jacobi-Matrix  $J_f(x)$  sind. □

**Satz 2.17 (Mittelwertsatz):** Sei  $f \in \mathcal{C}^1(\Omega)$  und liege die Strecke von  $a$  nach  $b$  in  $\Omega$ . Dann gibt es ein  $\theta \in (0, 1)$ , sodass

$$f(b) - f(a) = \langle \nabla f(a + \theta(b - a)), b - a \rangle. \quad (2.27)$$

*Beweis.* Betrachte die Funktion  $g : [0, 1] \rightarrow \mathbb{R}$  mit  $g(t) := f(a + \theta(b - a))$ . Dann ist  $g(0) = f(a)$  und  $g(1) = f(b)$ . Wir nutzen die aus Analysis I bekannte Taylorreihenentwicklung für Funktionen in einer Variablen. In erster Ordnung gilt dann

$$g(1) = g(0) + g'(\theta)$$

für irgendein  $\theta \in (0, 1)$ . Parametrisiere nun  $\varphi$  die Strecke von  $a$  nach  $b$ , also  $\varphi : [0, 1] \rightarrow \Omega$  mit  $\varphi(t) := a + t(b - a)$ . Dann ist  $g = f \circ \varphi$  und es folgt

$$g'(t) = df(\varphi(t))\varphi'(t) = \langle \nabla f(a + t(b - a)), b - a \rangle.$$

Die zu beweisende Aussage folgt nun durch Einsetzen dieser Darstellung in obige Taylorreihenentwicklung von  $g$ . □

**Korollar 2.18:** Verschwindet der Gradient einer  $\mathcal{C}^1$ -Funktion auf einem Gebiet  $\Omega$ , so ist die Funktion dort konstant.

*Beweis.* Sei  $x_0 \in \Omega$  fest und wähle  $x \in \Omega$  beliebig. Da  $\Omega$  zusammenhängend ist, lassen sich  $x$  und  $x_0$  durch einen polygonalen Weg durch die Punkte  $x_0, x_1, \dots, x_N := x$  miteinander verbinden. Nach dem Mittelwertsatz 2.17 ist dann

$$f(x_0) = f(x_1) = \dots = f(x_{N-1}) = f(x_N) = f(x).$$

Das gilt für beliebig gewähltes  $x_0$  und beliebige Wege. □

**Bemerkung:** Ist  $\Omega$  nicht wegzusammenhängend (und damit insbesondere kein Gebiet), so gilt Korollar 2.18 nicht. Seien beispielsweise  $\Omega_1$  und  $\Omega_2$  zwei offene disjunkte Mengen ( $\Omega_1 \cap \Omega_2 = \emptyset$ ) und  $\Omega := \Omega_1 \cup \Omega_2$ , so kann  $f : \Omega \rightarrow \mathbb{R}$  als

$$f(x) := \begin{cases} 1, & \text{für } x \in \Omega_1 \\ 2, & \text{für } x \in \Omega_2 \end{cases}$$

gewählt werden. Dann ist  $f \in \mathcal{C}^1(\Omega)$ ,  $\nabla f(x) = 0$  auf ganz  $\Omega$ , aber  $f$  ist nicht konstant.

## 2.6.2 Taylorreihen und Taylorentwicklung

**Erinnerung** Für ein Gebiet  $\Omega \subset \mathbb{R}^n$  und eine Funktion  $f \in \mathcal{C}^k(\Omega)$  ist die  $k$ -Ableitung von  $f$  an  $x \in \Omega$  (auch: das Differential  $k$ -ter Ordnung von  $f$  an  $x$ ) gegeben durch die Abbildung  $D^{(k)}f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  mit

$$D^{(k)}f(x)h := D^{(k)}f(x)(h) := \underbrace{d(\cdots(df(x))\cdots)}_{k\text{-mal}} h. \quad (2.28)$$

**Satz 2.19 (Taylor-Formel):** Sei  $f \in \mathcal{C}^{k+1}(\Omega)$  und  $x \in \Omega$ . Dann gilt mit  $\delta > 0$  für alle  $h \in K_\delta(x) \subset \Omega$

$$f(x+h) = f(x) + \sum_{j=1}^k \frac{1}{j!} D^{(j)}f(x)h^j + R_k(x+h, x) \quad (2.29)$$

wobei der Restterm gegeben ist durch

$$R_k(x+h, x) := \frac{1}{(k+1)!} D^{(k+1)}f(x+\tau h)h^{k+1} \quad (2.30)$$

mit  $\tau \in (0, 1)$ .

*Beweis.* Wie so oft führen wir den mehrdimensionalen auf den eindimensionalen Fall zurück. Sei also  $\varphi : [0, 1] \rightarrow \Omega$  der Weg mit  $\varphi(t) := x + th$  und definiere  $g : [0, 1] \rightarrow \mathbb{R}$  als Verknüpfung  $f \circ \varphi$ . Dann gilt  $g(0) = f(x)$  und  $g(1) = f(x+h)$ , und insbesondere ist  $g \in \mathcal{C}^{k+1}([0, 1])$  als Verknüpfung zweier  $\mathcal{C}^{k+1}$ -Funktionen. Somit haben wir nach der Taylor-Formel für reellwertige Funktionen in einer Variablen

$$g(1) = g(0) + \sum_{j=1}^k \frac{1}{j!} g^{(j)}(0) \cdot (1-0)^j + \frac{1}{(k+1)!} g^{(k+1)}(\tau)(1-0)^{k+1} \quad (*)$$

für ein  $\tau \in (0, 1)$ .

Nun ist

$$\begin{aligned} g^{(1)}(t) &= D^{(1)}(f \circ \varphi)(t) = f'(\varphi(t))\varphi'(t) = f'(\varphi(t))h, \\ g^{(2)}(t) &= D^{(2)}(f \circ \varphi)(t) = D^{(1)}(f'(\varphi(t))h) = f''(\varphi(t))h^2 \end{aligned}$$

und deswegen induktiv für  $n \geq k + 1$

$$g^{(n)}(t) = D^{(n)}(f \circ \varphi)(t) = D^{(1)}(f^{(n-1)}(\varphi(t))h) = f^{(n)}(\varphi(t))h^n,$$

sodass (\*) äquivalent ist zu

$$f(x+h) = f(x) + \sum_{j=1}^k \frac{1}{j!} f^{(j)}(x)h^j + \frac{1}{(k+1)!} f^{(k+1)}(x+\tau h)h^{k+1}. \quad \square$$

**Bemerkung:** Die Taylor-Formel lässt sich auch auf vektorwertige Funktionen verallgemeinern. Der Beweis folgt ebenso aus dem eindimensionalen Fall. Allerdings muss dabei die Ordnung der partiellen Ableitungen mitgeführt, und die Kettenregel für partielle Ableitungen verwendet werden. Das ist an und für sich nicht schwieriger, erfordert aber eine etwas angepasste Notation und die Verwendung sogenannter Multiindizes.

**Beispiele** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  die Funktion mit  $f(x, y) := \exp(xy)$ . Wir suchen die Taylorreihenentwicklung der zweiten Ordnung für  $h = (h_1, h_2) \in \mathbb{R}^2$  im Punkt  $0 \in \mathbb{R}^2$ . Es gilt:

$$f(h_1, h_2) = f(0, 0) + df(0, 0) \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{1}{2} D^{(2)} f(0, 0) h^2 + R_2(h, 0).$$

Dabei ist die zweite Ableitungen gegeben durch die *Hesse-Matrix* (die Jacobi-Matrix der ersten Ableitung)

$$Hf(x, y) := D^{(2)} f(x) = \left( \frac{\partial^2 f}{\partial x_i \partial x_k} \right)_{1 \leq i, k \leq 2}.$$

Sie berechnet sich, wie sich zweite Ableitungen gewöhnlich berechnen, als Ableitung der ersten Ableitung von  $f$ . Da

$$df(x, y) = (y \exp(xy), x \exp(xy))$$

ist also

$$D^{(2)} f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x, y) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x, y) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x, y) & \frac{\partial^2 f}{\partial x_2^2}(x, y) \end{pmatrix} = \begin{pmatrix} y^2 \exp(xy) & (1+xy) \exp(xy) \\ (1+xy) \exp(xy) & x^2 \exp(xy) \end{pmatrix}.$$

Berechnen wir die Ableitungen bis zur zweiten Ordnung an  $0 \in \mathbb{R}^2$ , so gilt also an  $h \in \mathbb{R}^2$  die Taylordarstellung

$$\begin{aligned} f(h) &= 1 + 0 + \frac{1}{2} h \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot h + R_2(h, 0) \\ &= 1 + h_1 h_2 + R_2(h, 0). \end{aligned}$$

Man bemerke, dass es sich bei Ableitungen um vektorwertige Funktionen handelt. Deswegen bedeuten die Multiplikationen in der Taylorreihendarstellung wieder Skalarmultiplikationen in den entsprechenden Räumen.



Sei nun  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := \sin(x + 2y)$ . Dann sind die Ableitungen bis zur Ordnung 2 gegeben durch

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \cos(x + 2y), & \frac{\partial f}{\partial y}(x, y) &= 2 \cos(x + 2y) \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= -\sin(x + 2y), & \frac{\partial^2 f}{\partial y^2}(x, y) &= -4 \sin(x + 2y), \end{aligned}$$

sowie

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

Damit lautet die Taylorreihenentwicklung von  $f$  an der Stelle  $h \in \mathbb{R}^2$  um  $0 \in \mathbb{R}^2$

$$f(h) = 0 + h_1 + 2h_2 + R_2(h, 0).$$

## 2.7 Lokale Extrema

**Definition 2.20:** Eine Funktion  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  hat in  $a \in \Omega$  ein *lokales Minimum* (respektive *lokales Maximum*), falls ein  $\delta > 0$  existiert, sodass

$$f(a) \leq f(x) \quad (\text{beziehungsweise } f(a) \geq f(x)) \quad (2.31)$$

für alle  $x \in K_\delta(a)$  gilt. Ist  $a$  lokale Minimal- oder lokale Maximalstelle, so heißt  $a$  *lokale Extremalstelle*.

**Definition 2.21:** Sei  $\Omega \subset \mathbb{R}^n$  und  $f \in \mathcal{C}^1(\Omega)$ . Dann heißt  $a \in \Omega$  *stationärer (oder kritischer) Punkt von  $f$* , wenn

$$\nabla f(x) = 0 \quad (2.32)$$

gilt.

**Satz 2.22:** Ist  $\Omega \subset \mathbb{R}^n$  und  $f \in \mathcal{C}^2(\Omega)$  und ist  $a \in \Omega$  eine lokale Minimalstelle, so gilt

$$f'(a) = 0 \quad \text{und} \quad f''(a) \geq 0 \quad (2.33)$$

beziehungsweise

$$f'(a) = 0 \quad \text{und} \quad f''(a) \leq 0, \quad (2.34)$$

wenn  $a$  eine lokale Maximalstelle ist.

Erfüllt auf der anderen Seite eine Funktion die Bedingungen

$$f'(a) = 0 \quad \text{und} \quad f''(a) > 0 \quad (2.35)$$

beziehungsweise

$$f'(a) = 0 \quad \text{und} \quad f''(a) < 0, \quad (2.36)$$

so ist  $a$  eine lokale Maximal- beziehungsweise Minimalstelle. Die obere Aussage gibt somit ein notwendiges Kriterium, die untere ein hinreichendes Kriterium für Extremalstellen an. Hierbei schreiben wir  $f''(a) > 0$  ( $f''(a) \geq 0$ ), wenn die Hesse-Matrix positiv (semi-)definit und  $f''(a) < 0$  ( $f''(a) \leq 0$ ), wenn die Hesse-Matrix negativ (semi-)definit ist. Ist die Hesse-Matrix weder positiv noch negativ semi-definit (also indefinit), so ist  $a$  keine Extremalstelle von  $f$ .

*Beweis.* Wir zeigen zunächst die notwendigen Bedingungen. Da  $\min f(x) = -\max -f(x)$  ist, genügt es, den Beweis für das Minimum zu führen.

Sei  $h \in \mathbb{R}^n$  fixiert und definiere  $g : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$  durch

$$g(t) := f(a + th),$$

wobei  $a$  die lokale Minimalstelle von  $f$  ist. Dann gilt für  $\varepsilon$  hinreichend klein, dass

$$g(t) = f(a + th) \geq f(a) = g(0),$$

nach Definition der lokalen Minimalstelle. Da dies für alle  $t \in K_\varepsilon(0)$  gilt, ist 0 eine lokale Minimalstelle von  $g$ . Deswegen gilt nach Analysis I

$$g'(0) = 0 \quad \text{und} \quad g''(0) \geq 0.$$

Die Forderungen sind nach Definition von  $g$  und Kettenregel äquivalent zu

$$\langle \nabla f(a), h \rangle = 0 \quad \text{und} \quad \langle h, Hf(a)h \rangle \geq 0.$$

Beide Aussagen gelten für beliebige  $h$ , weswegen  $\nabla f(a) = 0$  und  $Hf(a) \geq 0$  gelten muss.

Der Beweis der hinreichenden Bedingungen sei nur skizziert: Nehmen wir an, die Hesse-Matrix von  $f$  an  $a$  sei positiv definit. Wir bezeichnen mit  $\lambda_{\min}$  ihren kleinsten Eigenwert. Dann gilt

$$\langle h, Hf(a)h \rangle \geq \lambda \|h\|^2$$

(siehe Lineare Algebra). Nun lässt sich zeigen, dass die Hesse-Matrix von  $f$  auch in einer hinreichend kleinen Kugel um  $a$  positiv definit sein muss und daher

$$\langle h, Hf(x)h \rangle \geq \frac{\lambda}{2} \|h\|^2 \quad \text{für } x \in K_\delta(a)$$

und alle  $h \in \mathbb{R}^n$  und  $\delta > 0$  hinreichend klein. Akzeptiert man diese Aussage, so folgt aus der Taylorreihenentwicklung. Es gilt

$$f(a + h) - f(a) = \underset{=0}{\nabla f(a)h} + \frac{1}{2} \langle h, Hf(a)h \rangle + R_2(a + h, a) \geq \frac{\lambda}{4} \|h\|^2 + R_2(a + h, a).$$

Der Restterm beinhaltet Potenzen größer 2 von  $h$  und geht damit für  $h \rightarrow 0$  schneller gegen Null als die anderen Terme niedrigerer Ordnung.<sup>2</sup> Damit ist für hinreichend kleines  $h$  die rechte Seite strikt positiv. Es gilt also

$$f(a + h) - f(a) > 0$$

für hinreichend kleine  $h$ , weswegen  $a$  lokale Minimalstelle von  $f$  sein muss. □

<sup>2</sup>Man sagt auch: Der Restterm ist von Ordnung  $\mathcal{O}(h^2)$ . Diese  $\mathcal{O}$  ist ein sogenanntes *Landau-Symbol*.

Bemerkung: Zur Erinnerung: Eine Matrix symmetrische  $A \in M(n, n)$  ist positiv definit, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

1. Es gilt  $\langle x, Ax \rangle > 0$  für alle  $x \in \mathbb{R}^n$  mit  $x \neq 0$ .
2. Alle Eigenwerte der Matrix sind positiv (insbesondere reell).

Eine positiv semi-definite Matrix hat entsprechend nicht-negative Eigenwerte.

Für negativ (semi-)definite Matrizen gelten die entsprechenden Aussagen mit umgedrehten Relationszeichen.

### Beispiele

1. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := x^2 + y^2 + xy + 1$ . Wir suchen die lokalen Extrempunkte von  $f$ . Dazu berechnen wir

$$\frac{\partial f}{\partial x} = 2x + y \quad \text{und} \quad \frac{\partial f}{\partial y} = 2y + x.$$

Nun muss also gemäß (2.32)  $2x + y = 2y + x = 0$  sein. Dies ist nur für den Punkt  $(0, 0)$  erfüllt, der also der einzige *kritische Punkt* ist. Wir bilden die Hesse-Matrix von  $f$

$$Hf(x, y) = Hf(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Sie ist konstant, insbesondere ist  $Hf(0, 0) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ . Entweder weiß man nun, dass strikt diagonal dominanten Matrizen positiv definit sind, wenn ihre Diagonalelemente positiv sind, oder man überprüft, dass

$$\langle h, Hf(x, y)h \rangle = 2h_1^2 + h_1h_2 + h_2h_1 + h_2^2 = (h_1 + h_2)^2 + h_1^2 + h_2^2 > 0$$

für alle  $h = (h_1, h_2) \neq 0$ . Also ist die Hesse-Matrix stets positiv definit, insbesondere ist  $(0, 0)$  ein Minimalpunkt von  $f$ .

2. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^3 + 3y - y^3 - 3x$ . Die partiellen Ableitungen von  $f$  lauten

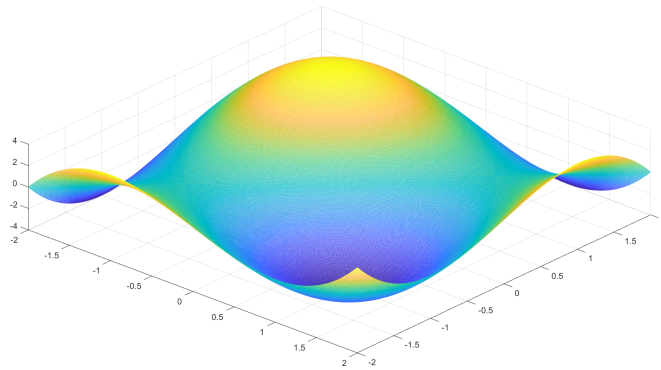
$$\frac{\partial f}{\partial x} = 3x^2 - 3 \quad \text{und} \quad \frac{\partial f}{\partial y} = -3y^2 + 3.$$

Dieses quadratische Gleichungssystem in zwei Variablen hat die vier Nullstellen  $(-1, -1)$ ,  $(-1, 1)$ ,  $(1, -1)$  und  $(1, 1)$ . Die Hesse-Matrix von  $f$  lautet

$$Hf(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & -6y \end{pmatrix}.$$

Setzen wir die vier kritischen Punkte ein, finden wir direkt, dass die Punkte  $(-1, -1)$  und  $(1, 1)$  Sattelpunkte von  $f$  sein müssen. An  $(1, -1)$  ist die Matrix positiv definit und  $f$  nimmt somit an  $(1, -1)$  ein lokales Minimum an. Analog ist  $Hf(-1, 1)$  negativ definit, sodass  $f$  an dieser Stelle ein lokales Maximum annimmt.

Man kann die kritischen Punkte im Plot rechts gut erkennen. Es fällt insbesondere auf, dass die Sattelpunkte aus  $x$ -Richtung eine andere Charakterisierung haben als aus  $y$ -Richtung: Liefere man von der Minimalstelle gerade auf die Sattelpunkte zu, so befände sich an dieser Stelle lokal der höchste Punkt des Wegs. Von der Maximalstelle aus kommend sind die Sattelpunkte jedoch lokale Minimalstellen. Diese ambivalente Charakterisierung ist typisch für Sattelpunkte.



**Bemerkung 2.23:** Für zweimal stetig differenzierbare Funktionen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gilt:

1. Die Hesse-Matrix  $Hf(x, y)$  an  $(x, y)$  ist genau dann positiv definit, wenn  $\det Hf(x, y) > 0$  und  $\partial_{xx}f(x, y) > 0$ .
2. Die Hesse-Matrix  $Hf(x, y)$  an  $(x, y)$  ist genau dann negativ definit, wenn  $\det Hf(x, y) > 0$  und  $\partial_{xx}f(x, y) < 0$ .
3. Ist die Determinante der Hesse-Matrix an  $(x, y)$ , so hat sie dort einen Sattelpunkt.

## 2.8 Der Satz von der Umkehrabbildung und der Satz über implizite Funktionen

Aus der linearen Algebra wissen wir, wann und wie wir lineare Gleichungssysteme  $Ax = 0$  mit  $m$  Gleichungen und  $n$  Unbekannten lösen können. Falls  $n \geq m$  ist, werden zum Beispiel mit Hilfe des Gauß-Verfahrens die Gleichungen sukzessive nach den einzelnen Variablen aufgelöst. Außerdem kennen wir die Struktur der Lösungsmenge des Gleichungssystems: Im homogenen Fall besitzt sie die Struktur eines linearen Unterraums, im nicht-homogenen Fall die eines affinen Unterraums. Die gleichen Fragen nach der Lösbarkeit lassen sich auch für nichtlineare Gleichungssysteme stellen. Hier ist jedoch zunächst völlig unklar, ob ein System mit  $m$  nichtlinearen Funktionen und  $n$  Unbekannten überhaupt Lösungen besitzt, wie die Struktur der Lösungsmenge aussieht und von wie vielen Parametern die Lösungsmenge abhängt. Um diese Fragen zu beantworten, versuchen wir ebenso wie im linearen Fall, Gleichungen nach verschiedenen Variablen aufzulösen.

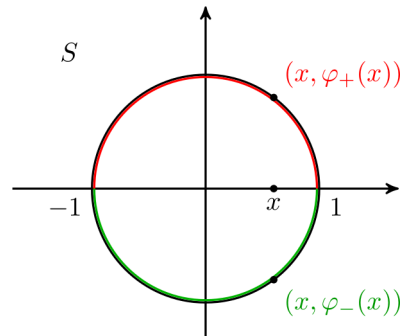
Betrachten wir beispielsweise die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) = x^2 + y^2 - 1$ . Die Nullstellenmenge  $S$  dieser Funktion beschreibt den Kreis (genauer: den Rand des Kreises) mit Radius 1 um den Nullpunkt in der reellen Zahlenebene. Nun lässt sich die Gleichung  $x^2 + y^2 - 1 = 0$  nach  $y$  auflösen. Dabei erhalten wir

$$y = \pm\sqrt{1-x^2}.$$

Definieren wir also die Funktionen  $\varphi_+, \varphi_- : [-1, 1] \rightarrow \mathbb{R}$  mit

$$\varphi_+(x) := +\sqrt{1-x^2} \quad \text{und} \quad \varphi_-(x) := -\sqrt{1-x^2},$$

so beschreiben die Graphen dieser beiden Funktionen jeweils einen Teil der Lösungsmenge  $S$ . Beide Funktionen  $\varphi_+$  und  $\varphi_-$  sind *lokale* Auflösungen der Gleichung  $f(x) = 0$ . Eine globale Auflösung gibt es hier jedoch nicht.



Bevor wir zur entscheidenden Aussage dieses Abschnitts, dem Satz über Implizite Funktionen, kommen, der diese Beobachtung mathematisch beweist, führen wir zunächst wichtige Definitionen ein.

**Definition 2.24 (Diffeomorphismus):** Seien  $U, V \subset \mathbb{R}^n$  zwei Gebiete. Eine Abbildung  $f : U \rightarrow V$  heißt *Diffeomorphismus zwischen  $U$  und  $V$* , wenn  $f$  eine stetig differenzierbare Abbildung ist, deren Umkehrabbildung  $f^{-1} : V \rightarrow U$  existiert und ebenfalls stetig differenzierbar ist. Die Funktion  $f$  heißt  *$C^k$ -Diffeomorphismus zwischen  $U$  und  $V$* , wenn  $f$  und  $f^{-1}$  beide  $k$ -mal stetig differenzierbar sind.

Diffeomorphismen heißen auch *Koordinatentransformationen*.

**Satz 2.25:** Seien  $U, V \subset \mathbb{R}^n$  und  $f : C^k(U, V)$  ein  $C^k$ -Diffeomorphismus. Dann ist die Jacobi-Matrix  $J_f(x)$  an jeder Stelle  $x \in U$  invertierbar. Für die Jacobi-Matrix der inversen Funktion  $f^{-1} : V \rightarrow U$  gilt mit  $y = f(x) \in V$  die Identität

$$(J_f^{-1}(y)) = (J_f(x))^{-1}. \tag{2.37}$$

*Beweis.* Da  $f^{-1} \circ f = \text{id}_U$  und die Jacobi-Matrix der Identität die Einheitsmatrix ist, gilt

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = J(f^{-1} \circ f) = J(f^{-1}(f(x)))J_f(x)$$

nach Kettenregel. Die Jacobi-Matrix der Umkehrabbildung  $f^{-1}$  ist nach Voraussetzung invertierbar. Die Aussage folgt nun aus der Eindeutigkeit der Inversen einer Matrix.  $\square$

**Bemerkung 2.26:** Äquivalent zur Invertierbarkeit der Jacobi-Matrix  $J_f(x)$  von  $f$  an  $x$  ist die Aussage, dass das Differential  $df(x)$  von  $f$  in  $x$  also für alle  $x \in U$  ein Isomorphismus ist. Der Satz 2.25 liefert eine Möglichkeit zur Berechnung der Ableitung der inversen Abbildung, ohne die inverse Abbildung selbst bestimmen zu müssen.

**Beispiel (Koordinatentransformation in der Ebene)** Seien die Menge  $U := (0, \infty) \times (-\pi, \pi)$  und  $V := \mathbb{R}^2 \setminus \{(x, 0), x \leq 0\}$  gegeben.<sup>3</sup> Wir definieren  $f : U \rightarrow V$  als die Abbildung, die

<sup>3</sup> $V$  ist also die reelle Zahlenebene ohne den nicht-positiven Teil der  $x$ -Achse.

Polar- in kartesische Koordinaten transformiert, also

$$f(r, \varphi) := \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}. \quad (2.38)$$

Die Umkehrabbildung zu  $f$  ist durch  $f^{-1} : V \rightarrow U$  mit

$$f^{-1}(x, y) := \begin{pmatrix} r \\ \varphi \end{pmatrix} := \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan\left(\frac{y}{x}\right) \end{pmatrix} \quad (2.39)$$

gegeben. In der Tat ist  $f^{-1}$  aufgrund der Einschränkung in  $V$  wohldefiniert. Beide Abbildungen sind beliebig oft stetig differenzierbar, sodass  $f$  ein  $\mathcal{C}^\infty$ -Diffeomorphismus ist.

Betrachten wir nun die Ableitungen. Die Jacobi-Matrix von  $f$  lautet

$$J_f(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Ihre Inverse lässt sich mithilfe der bekannten Regel für 2x2-Matrizen bestimmen. Sie lautet

$$(J_f(r, \varphi))^{-1} = \frac{1}{r} \begin{pmatrix} r \cos \varphi & r \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}.$$

Berechnen wir die Ableitung der Umkehrabbildung  $f^{-1}$ , finden wir

$$J(f^{-1}(x, y)) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ -\frac{y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Die Matrizen  $J(f^{-1}(x, y))$  und  $(J_f(r, \varphi))^{-1}$  stimmen also tatsächlich überein.

**Satz 2.27 (Satz von der Umkehrabbildung):** Sei  $\Omega \subset \mathbb{R}^n$  und  $f \in \mathcal{C}^k(\Omega)$ , wobei die Jacobi-Matrix  $J_f(a)$  von  $f$  an einem Punkt  $a \in \Omega$  invertierbar sein soll. Dann gibt es eine Umgebung  $U \subset \Omega$  von  $a$  und eine Umgebung  $V$  von  $f(a)$ , sodass die Restriktion von  $f$  auf  $U$

$$f|_U : U \rightarrow V \quad \text{mit } f|_U(x) := f(x)$$

ein  $\mathcal{C}^k$ -Diffeomorphismus zwischen  $U$  und  $V$  ist.

*Beweis.* Der Beweis lässt sich in der angegebenen Literatur nachlesen. Der Satz von der Umkehrabbildung (inversen Funktion) ist ein Korollar des nachfolgenden Satzes über implizite Funktionen 2.28.  $\square$

**Beispiel** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  durch

$$f(x, y) := \begin{pmatrix} x + y + \exp(x) \\ x + y + \exp(y) \end{pmatrix}$$

definiert. Die Jacobi-Matrix von  $f$  lautet

$$J_f(x, y) := \begin{pmatrix} 1 + \exp(x) & 1 \\ 1 & 1 + \exp(y) \end{pmatrix}$$

und ist somit am Punkt  $(0, 0)$  invertierbar. Da  $f \in \mathcal{C}^\infty(\mathbb{R}^2)$ , existiert nach Satz 2.27 eine Umgebung  $U$  um den Ursprung und eine Umgebung  $V$  um  $f(0, 0) = (1, 1)$ , sodass die Einschränkung von  $f|_U : U \rightarrow V$  ein  $\mathcal{C}^\infty$ -Diffeomorphismus ist. Nach Satz 2.25 gilt außerdem

$$J(f^{-1}(1, 1)) = (J_f(0, 0))^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Obwohl wir also dieses Wissen über die Umkehrabbildung der Einschränkung von  $f$  auf  $U$  haben, ist eine explizite Darstellung dieser Umkehrfunktion nicht bekannt.

**Satz 2.28 (Satz über Implizite Funktion (IFT)):** Sei  $\Omega \subset \mathbb{R}^p \times \mathbb{R}^m$  offen und die Funktion  $f : \Omega \rightarrow \mathbb{R}^m$   $k$ -mal stetig differenzierbar. Sei  $(\bar{x}, \bar{y}) \in \Omega$  eine Nullstelle der Funktion  $f$ , und es gelte die Bedingung

$$0 \neq \det \left( \frac{\partial f_i}{\partial y_k}(\bar{x}, \bar{y}) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq m}}. \quad (2.40)$$

Die letzten  $m$  Spalten der Jacobi-Matrix seien also linear-unabhängig. Dann gibt es Umgebungen  $U \subset \Omega$  von  $(\bar{x}, \bar{y})$  und  $V \subset \mathbb{R}^p$  von  $\bar{x}$  und eine eindeutige Funktion  $\varphi \in \mathcal{C}^k(V, \mathbb{R}^m)$ , sodass

$$f(x, y) = 0 \iff y = \varphi(x) \quad (2.41)$$

für alle  $(x, y) \in U$  gilt. Das Gleichungssystem  $f(x, y) = 0$  lässt sich also lokal eindeutig nach  $y$  auflösen. Die Lösungsmenge von  $f$  wird lokal durch den Graphen von  $\varphi$  beschrieben. Außerdem gilt für die partiellen Ableitungen von  $\varphi$  und alle  $x \in V$

$$\left( \frac{\partial \varphi_l}{\partial x_i}(x) \right)_{\substack{1 \leq l \leq m \\ 1 \leq i \leq p}} = - \left( \frac{\partial f_k}{\partial y_l}(x, \varphi(x)) \right)_{\substack{1 \leq k \leq m \\ 1 \leq l \leq m}}^{-1} \cdot \left( \frac{\partial f_k}{\partial x_i}(x, \varphi(x)) \right)_{\substack{1 \leq k \leq m \\ 1 \leq i \leq p}}. \quad (2.42)$$

*Beweis.* Der Beweis lässt sich in der angegebenen Literatur nachlesen. □

**Bemerkung 2.29:** Durch  $f(x, y) = 0$  ist die lokale Auflösungsfunktion  $\varphi$  implizit gegeben. Sie lässt sich jedoch nicht immer explizit angeben. Nichtsdestotrotz liefert das Theorem eine fundamentale Möglichkeit, die Existenz von Lösungen nichtlinearer Gleichungssysteme nachzuweisen.

**Beispiel** Sei  $f : \Omega \subset \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$  die Funktion mit

$$f(x_1, x_2, y) := x_1^3 + 4x_1x_2 + y^2 - 3x_2y^2 - 3.$$

Durch scharfes Hinsehen finden wir die Nullstelle  $(1, 1, 1)$  von  $f$ . Die partielle Ableitung von  $f$  nach  $y$  ist durch

$$\frac{\partial f}{\partial y}(x_1, x_2, y) = 2y - 6x_2y$$

gegeben, sodass  $\frac{\partial f}{\partial y}(1, 1, 1) = -4 \neq 0$ . Damit lässt sich die Funktion  $f$  nach Satz 2.28 lokal eindeutig nach  $y$  auflösen. Es existieren also Umgebungen  $U$  von  $(1, 1, 1)$  und  $V$  von  $(1, 1)$  sowie eine eindeutige Funktion  $\varphi : V \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ , sodass

$$f(x_1, x_2, \varphi(x_1, x_2)) = 0$$

für alle  $(x_1, x_2) \in V$  gilt. Berechnen wir die partiellen Ableitungen von  $f$  an  $(1, 1, 1)$  bezüglich der beiden  $x$ -Komponenten, so finden wir

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_1, x_2, y) = 3x_1^2 + 4x_2 &\implies \frac{\partial f}{\partial x_1}(1, 1, 1) = 7 \\ \frac{\partial f}{\partial x_2}(x_1, x_2, y) = 4x_1 - 3y^2 &\implies \frac{\partial f}{\partial x_2}(1, 1, 1) = 1. \end{aligned}$$

Damit ist die Ableitung von  $\varphi$  gemäß (2.42) an  $(1, 1)$  gegeben durch

$$\left( \frac{\partial \varphi}{\partial x_1}(x_1, x_2), \frac{\partial \varphi}{\partial x_2}(x_1, x_2) \right) = - \left( \frac{\partial f}{\partial y}(1, 1, 1) \right)^{-1} \left( \frac{\partial f}{\partial x_1}(1, 1, 1), \frac{\partial f}{\partial x_2}(1, 1, 1) \right) = \left( \frac{7}{4}, \frac{1}{4} \right).$$

Bemerkung: Die Ableitung von  $\varphi$  existiert nicht nur an  $(1, 1)$ , sondern für alle Punkte in einer Umgebung um  $(1, 1)$ .

## Parametrisierung von Lösungsmannigfaltigkeiten

**Definition 2.30:** Sei  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$   $k$ -mal stetig differenzierbar und  $1 \leq m < n$ . Die Nullstellenmenge

$$N := \{x \in \Omega, f(x) = 0\} \tag{2.43}$$

heißt *Lösungsmannigfaltigkeit* von  $f$ , falls  $N \neq \emptyset$  ist und die Jacobi-Matrix  $J_f(x)$  an jedem  $x \in N$  vollen Zeilenrang hat. Die Gradienten der Komponentenfunktionen  $\nabla f_1(x), \dots, \nabla f_m(x)$  müssen also für alle  $x \in N$  linear unabhängig sein.

Sei nun  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine Funktion mit Lösungsmannigfaltigkeit  $N$  und  $c \in N$ . Dann hat  $J_f(c)$  vollen Zeilenrang. Da Spaltenrang gleich Zeilenrang ist, gibt es also  $m$  linear unabhängige Spalten in der Jacobi-Matrix  $J_f(c)$ . Wir können ohne Beschränkung der Allgemeinheit annehmen, das seien die letzten  $m$  Spalten der Matrix (andernfalls sortieren wir um). Dann schreiben wir die Lösungsmannigfaltigkeit als

$$N = \{(x, y) \in \mathbb{R}^{n-m} \times \mathbb{R}^m, f(x, y) = 0\},$$



wobei  $n - m \geq 1$ . Nach Definition ist dann

$$\det \left( \frac{\partial f_i}{\partial y_k}(c) \right) \neq 0.$$

Der Satz über Implizite Funktionen liefert uns nun die lokale Parametrisierung der Lösungsmannigfaltigkeit: Wir können das Gleichungssystem  $f(x, y) = 0$  in einer Umgebung  $W$  von  $c$  nach den letzten  $m$  Variablen  $y_1, \dots, y_m$  eindeutig auflösen, sodass

$$\begin{aligned} y_1 &= \varphi_1(x_1, \dots, x_{n-m}) \\ &\vdots \\ y_m &= \varphi_m(x_1, \dots, x_{n-m}), \end{aligned}$$

wobei  $\varphi_1, \dots, \varphi_m$   $C^k$ -Funktionen in einer Umgebung  $U$  von  $x$  sind.

## 2.9 Lokale Extrema unter Nebenbedingungen

Wenn an die Suche nach lokalen Extremstellen Nebenbedingungen gestellt werden, ist klar, dass die in Satz 2.22 gegebenen hinreichenden und notwendigen Bedingungen nicht ohne Weiteres übertragbar sind. Beispielsweise nimmt die konvexe Funktion  $f(x) = x$  auf  $[1, 2]$  ihr striktes globales Minimum in  $x = 1$  an. Dabei gilt weder  $f'(1) = 0$  noch  $f''(1) > 0$ . Die Suche nach lokalen Extremalstellen unter Nebenbedingungen unterscheidet sich also grundlegend von der unrestringierten Suche.

Im Folgenden sei  $\Omega \subset \mathbb{R}^n$  ein Gebiet und  $f_1, \dots, f_m \in C^1(\Omega)$  und  $g \in C^1(\Omega)$ , wobei  $m < n$ .

**Definition 2.31:** Definiere

$$N := \{x \in \Omega, f_i(x) = 0 \text{ für alle } i = 1, \dots, m\}$$

die gemeinsame Lösungsmenge der  $f_i$ . Ein Punkt  $x^* \in N$  heißt *lokale Minimalstelle* (*lokale Maximalstelle*) von  $g$  über  $N$ , wenn für alle  $x \in N \cap K_\varepsilon(x^*)$

$$g(x^*) \leq g(x) \quad (\text{beziehungsweise } g(x^*) \geq g(x)) \quad (2.44)$$

gilt. Die Gesamtmenge dieser lokalen Minimal-/Maximalstellen heißen *lokale Extremalstellen* von  $g$  unter den Nebenbedingungen  $f_1 = \dots = f_m = 0$ .

**Satz 2.32 (Notwendige Bedingung für Extremalität):** Sei  $N$  eine Lösungsmannigfaltigkeit und habe  $g$  an der Stelle  $z^* \in N$  ein lokales Extremum unter den Nebenbedingungen  $f_1, \dots, f_m = 0$ . Dann gibt es eindeutig bestimmte  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  (sogenannte *Lagrange-Multiplikatoren*), sodass

$$\nabla g(z^*) = \sum_{j=1}^m \lambda_j \nabla f_j(z^*). \quad (2.45)$$

Die Funktion  $g - \sum_{j=1}^m \lambda_j \nabla f_j$  hat an  $z^*$  einen kritischen Punkt.

*Beweis.* Da  $m < n$  ist, können wir die alle  $z \in \Omega$  wieder aufspalten in die ersten  $n - m$  und die letzten  $m$  Komponenten, sodass

$$N = \{(x, y), f_i(x, y) = 0\},$$

wobei mit  $x = (x_1, \dots, x_{n-m})$  die ersten  $n - m$  und mit  $y = (y_1, \dots, y_m)$  die letzten  $m$  Komponenten von  $z$  bezeichnet werden. Da  $N$  nach Voraussetzung eine Lösungsmannigfaltigkeit ist, können wir wie im Anwendungsbeispiel des letzten Abschnitts vorgehen. Für  $z^* = (x^*, y^*)$  existieren also, als Konsequenz des Satzes über Implizite Funktionen 2.28, zwei Umgebungen  $U \subset \mathbb{R}^{n-m} \times \mathbb{R}^m$  von  $(x^*, y^*)$  und  $V \subset \mathbb{R}^{n-m}$  von  $x^*$  sowie eine Funktion  $\varphi \in \mathcal{C}^1(V, \mathbb{R}^m)$ , sodass

$$f(x, y) = 0 \iff y = \varphi(x)$$

für alle  $(x, y) \in U$ . Bemerke, dass außerdem  $y^* = \varphi(x^*)$  gilt.

Wir betrachten nun die Funktion  $\bar{g} : V \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}$  mit  $\bar{g}(x) := g(x, \varphi(x))$ . Nun hat  $g$  nach Voraussetzung ein lokales Extremum in  $(x^*, y^*)$ , weswegen

$$\bar{g}(x^*) = g(x^*, \varphi(x^*)) = g(x^*, y^*) \leq g(x, y) = g(x, \varphi(x)) = \bar{g}(x)$$

für alle  $(x, y) \in N \cap K_\varepsilon((x^*, y^*))$  gilt. Also hat  $\bar{g} = g(\cdot, \varphi(\cdot))$  ein lokales Extremum in  $x^*$ . Insbesondere verschwinden die partiellen Ableitungen von  $\bar{g}$  an  $x^*$ . Es gilt somit nach Kettenregel

$$0 = \frac{\partial \bar{g}}{\partial x_i}(x^*) = \frac{\partial g}{\partial x_i}(x^*, y^*) + \sum_{j=1}^m \frac{\partial g}{\partial x_{n-m+j}}(x^*, y^*) \frac{\partial \varphi_j}{\partial x_i}(x^*)$$

für alle  $i = 1, \dots, n - m$ .

Des Weiteren sei  $\bar{f}_i : V \rightarrow \mathbb{R}$  für  $i = 1, \dots, m$  definiert durch  $\bar{f}_i(x) := f_i(x, \varphi(x))$ . Dann gilt  $\bar{f}_i = 0$  nach Definition der Lösungsmannigfaltigkeit. Die Funktion  $\bar{f}_i$  ist also konstant, weswegen ihre partiellen Ableitungen null sind. Es gilt also

$$0 = \frac{\partial \bar{f}_k}{\partial x_i}(x^*) = \frac{\partial f_k}{\partial x_i}(x^*, y^*) + \sum_{j=1}^m \frac{\partial f_k}{\partial x_{n-m+j}}(x^*, y^*) \frac{\partial \varphi_j}{\partial x_i}(x^*)$$

für alle  $i = 1, \dots, n - m$  und alle  $k = 1, \dots, m$ . Für die Vektoren

$$v_i := e_i + \sum_{j=1}^m \frac{\partial \varphi_j}{\partial x_i}(x^*) e_{n-m+j}, \quad i = 1, \dots, n - m,$$

gilt also nach obigen Überlegungen

$$0 = \langle v_i, \nabla g(x^*, y^*) \rangle \quad \text{und} \quad 0 = \langle v_i, \nabla f_k(x^*, y^*) \rangle$$

für  $k = 1, \dots, m$ . Die Vektoren  $\nabla g(x^*, y^*)$  und  $\nabla f_1(x^*, y^*), \dots, \nabla f_m(x^*, y^*)$  liegen also im Annullator  $W^\perp$  von  $W := \text{span}\{v_1, \dots, v_{n-m}\}$ , wobei

$$W^\perp := \{w \in \mathbb{R}^n, \langle w, v \rangle = 0 \text{ für alle } v \in W\}.$$

Aus der Definition der  $v_i$  folgt direkt ihre lineare Unabhängigkeit, sodass  $\dim U = n - m$ . Deswegen ist  $\dim U^\perp = n - \dim U = m$ . Da nach Voraussetzung und Definition einer Lösungsmannigfaltigkeit die  $m$  Vektoren  $\nabla f_i(x^*, y^*)$  linear unabhängig sind, bilden sie eine Basis von  $U^\perp$ . Wir können also jeden Vektor aus  $U^\perp$  eindeutig als Linearkombination der  $\nabla f_i$  darstellen. Das gilt insbesondere für  $\nabla g(x^*, y^*) \in U^\perp$ , also

$$\nabla g(x^*, y^*) = \sum_{k=1}^m \lambda_k \nabla f_k(x^*, y^*)$$

mit eindeutig bestimmten  $\lambda_1, \dots, \lambda_m$ . □

**Satz 2.33 (Hinreichende Bedingungen für Extremalität):** Seien Funktionen  $f_1, \dots, f_m$  und  $g$  nun  $\mathcal{C}^2$ -Funktionen auf  $\Omega$ . Sei außerdem  $N$  wieder eine gemeinsame Lösungsmannigfaltigkeit der  $f_i$ . Angenommen, es existieren Lagrange-Multiplikatoren  $\lambda_1, \dots, \lambda_m$  und eine Funktion  $G : \Omega \rightarrow \mathbb{R}^n$  mit

$$G(z) := g(z) - \sum_{j=1}^m \lambda_j f_j(z), \tag{2.46}$$

die an einem Punkt  $z^* \in \Omega$  folgende Bedingungen erfüllt:

1. Es gilt

$$\nabla G(z^*) = 0; \tag{2.47}$$

2. für alle Vektoren  $v \in \mathbb{R}^n \setminus \{0\}$ , die senkrecht auf den Gradienten  $\nabla f_i(z^*)$  stehen, also  $v \in \text{span}\{\nabla f_i(z^*), i = 1, \dots, m\}^\perp$ , gilt

$$\langle v, D^{(2)}G(z^*)v \rangle > 0 \quad (\text{beziehungsweise } \langle v, D^{(2)}G(z^*)v \rangle < 0) \tag{2.48}$$

Dann ist  $z^* \in N$  eine lokale Minimalstelle (beziehungsweise lokale Maximalstelle) von  $g$  unter den Nebenbedingungen  $f_1 = \dots = f_m = 0$ .

**Beispiel** Gegeben sei die Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $g(x, y) := xy + 14$ . Wir suchen

$$\min_{(x,y) \in \mathbb{R}^2} g(x, y) \\ \text{u. d. N. } (x, y) \in \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - 18 = 0\}$$

Die Nebenbedingungen werden also durch die Nullstellenmenge  $N$  der Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := x^2 + y^2 - 18$  beschrieben. Bemerke, dass die Funktion  $g$  eine lokale Extremalstelle an  $(0, 0)$  hat, da  $\nabla g(0, 0) = 0$ . Diese lokale Extremalstelle liegt jedoch nicht in  $N$ .

Nehmen wir an, dass eine Stelle  $(x^*, y^*) \in N$  existiert, an der  $g$  ein lokales Extremum hat. Dann muss an dieser Stelle Bedingung (2.45) erfüllt sein. Bestimmen wir also zunächst die Gradienten von  $g$  und  $f$ . Wir haben

$$\nabla g(x, y) := \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{und} \quad \nabla f(x, y) := \begin{pmatrix} 2x \\ 2y \end{pmatrix},$$

sodass an  $(x^*, y^*)$

$$\nabla g(x^*, y^*) = \lambda \nabla f(x^*, y^*) \implies \begin{cases} y^* &= 2\lambda x^* \\ x^* &= 2\lambda y^* \end{cases}$$

gelten muss. Bemerke, dass  $\lambda \in \mathbb{R}$  liegt. Aus beiden Bedingungen folgt

$$\lambda = \frac{y^*}{2x^*} = \frac{x^*}{2y^*} \implies (x^*)^2 = (y^*)^2$$

am möglichen Extrempunkt  $(x^*, y^*)$ . Da der Punkt auch Nullstelle von  $f$  sein muss, muss also

$$0 = x^2 + y^2 - 18 = 2x^2 - 18 \iff x^2 = 9$$

sein. Diese Bedingung wird von den vier kritischen Punkten

$$z_1 = (-3, -3), \quad z_2 = (-3, 3), \quad z_3 = (3, -3), \quad \text{und} \quad z_4 = (3, 3)$$

erfüllt. Je nach Wahl ist dann  $\lambda = \frac{1}{2}$  oder  $\lambda = -\frac{1}{2}$ . Insbesondere sehen wir, beispielsweise für  $z_3 = (3, -3)$ , dass

$$\nabla g(3, -3) = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 6 \\ -6 \end{pmatrix} = \lambda \nabla f(3, -3).$$

Überprüfen wir nun Bedingungen (2.47) und (2.48). Sei dafür  $G := g - \lambda f$ . Dann ist

$$\nabla G(x, y) = \begin{pmatrix} y - \lambda 2x \\ x - \lambda 2y \end{pmatrix},$$

sodass tatsächlich  $\nabla G(z_i) = 0$  für  $i = 1, 2, 3, 4$ .

Zur Überprüfung von (2.48) bestimme die Hesse-Matrix von  $G$  als

$$HG(x, y) = \begin{pmatrix} -2\lambda & 1 \\ 1 & -2\lambda \end{pmatrix}.$$

Betrachten wir den Punkt  $z_1 = (-3, -3)$ . Hier ist  $\lambda = \frac{1}{2}$ . Zunächst müssen wir die Vektoren bestimmen, die senkrecht auf dem Gradienten von  $f$  an  $z_1$  stehen. Nun ist

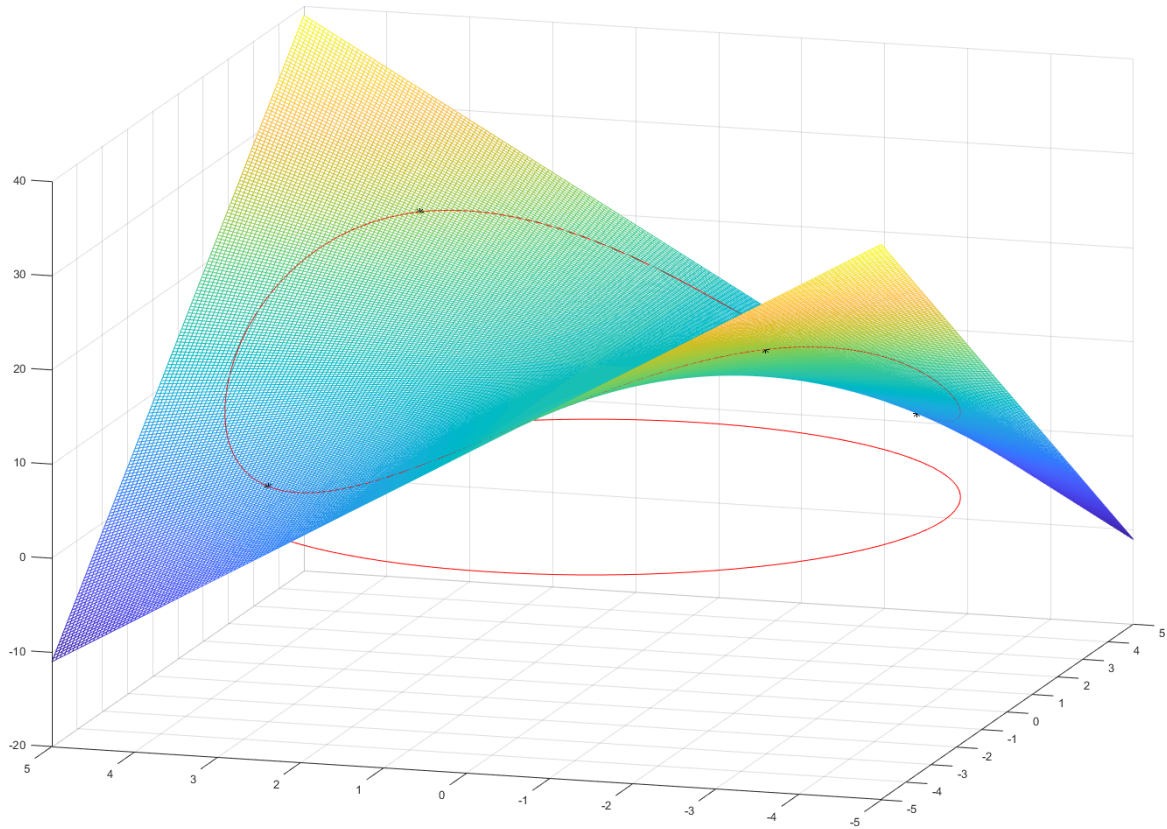
$$0 = \langle v, \nabla f(-3, -3) \rangle \iff -6v_1 + -6v_2 = 0 \iff v_1 = -v_2.$$

Da nun für  $v = (v_1, -v_1)^T \neq 0$

$$\langle v, HG(z_1)v \rangle = \begin{pmatrix} v_1 & -v_1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ -v_1 \end{pmatrix} = -4v_1^2 < 0$$

gilt, ist  $z_1 = (-3, -3)$  eine lokale Maximalstelle von  $g$  unter den Nebenbedingungen  $N$ .

Auf analogem Weg findet man, dass auch  $z_4$  eine lokale Maximalstelle ist, und dass  $z_2$  und  $z_3$  lokale Minimalstellen von  $g$  unter den Nebenbedingungen  $N$  sind. Die Funktionswerte an den Maximalstellen betragen  $g(-3, -3) = g(3, 3) = 23$ , an den Minimalstellen  $g(-3, 3) = g(3, -3) = 5$ .



Im Bild ist der Graph der Funktion  $g$  geplottet. Die durch die Nullstellenmenge von  $f$  gegebene Restriktionsmenge ist rot dargestellt. Die zulässige Menge des Optimierungsproblems ist also die Menge aller Punkte  $(x, y, z) \in \mathbb{R}^3$ , die im Graphen von  $g$  liegen, und deren ersten beiden Komponenten eine Nullstelle von  $f$  bilden. Im Plot entspricht das der roten Linie im Graphen von  $g$ . Die Extrempunkte des restringierten Optimierungsproblems sind gekennzeichnet.

# Kapitel 3

## Integralrechnung in $\mathbb{R}^n$

In diesem Kapitel schauen wir uns Verallgemeinerungen des Riemann-Integrals für Funktionen in mehreren Variablen an. In der Schule wurde das Integral anschaulich als Flächeninhalt der Fläche unterhalb eines Funktionsgraphen eingeführt. Wir möchten deswegen versuchen, einen Integralbegriff einzuführen, der in natürlicher Weise die Begriffe „Flächeninhalt“ und „Volumen“ verallgemeinert. Insbesondere soll die Berechnung des Integrals von elementargeometrischen Objekte (wie etwa Quadraten, Kreisen, Kegeln, Kugeln, Zylindern, ...) den gleichen Wert liefern wie die Berechnung des Flächeninhalts/Volumens der Objekte nach den aus der Schule bekannten Formeln.

Um dem gerecht zu werden, werden folgende Bedingungen an einen Integralbegriff gestellt:

1. Der Integralbegriff soll *monoton* sein: Ist  $A$  also eine Teilmenge von  $B$ , so fordern wir, dass das Integral über  $A$  kleiner oder gleich dem Integral über  $B$  ist.
2. Das Integral über ein Objekt soll *translationsinvariant* sein. Man muss das Objekt also verschieben können, ohne dass sich das Integral ändert.
3. Der Integralbegriff soll in natürlicher Weise die Berechnungsformeln für den Flächeninhalt von Rechtecken und Quadern in 3 Dimensionen verallgemeinern.
4. Das Integral soll  $\sigma$ -additiv sein. Das bedeutet, das Integral über die Vereinigung abzählbar vieler paarweise disjunkter Mengen entspricht der Summe der Integrale über die einzelnen Mengen. Nimmt man also beispielsweise  $n$  zweidimensionale Objekte, die sich nicht überlappen, soll der Flächeninhalt der Vereinigung all dieser Mengen der Summe der Flächeninhalte der einzelnen Objekte entsprechen.

Es lässt sich nun recht einfach beweisen, dass sich solch ein Integralbegriff nicht wohldefinieren lässt. Es gibt also keine Funktion  $\mu : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$  mit den vier genannten Eigenschaften. Der entscheidende Punkt ist hier die Wahl des Definitionsbereiches: Einen Integralbegriff auf beliebigen Teilmengen der reellen Zahlen schaffen zu wollen, ist zu ambitioniert. Wir müssen uns also auf bestimmte Teilmengen des  $\mathbb{R}^n$  beschränken. Betrachten wir nur bestimmte beschränkte Teilmengen des  $\mathbb{R}^n$ , führt das auf den Begriff des sogenannten *Jordan-Volumens*. Diesem historisch lange Zeit etablierten und geometrisch leicht zugänglichen Integralbegriff steht der moderne Begriff des *Lebesgue-Integrals* gegenüber. Letzterer ermöglicht die sinnvolle Definition eines Integrals auf einer deutlich größeren Klasse von Mengen und erlaubt hierbei insbesondere auch unbeschränkte Mengen. Auch besitzt das Lebesgue-Integrals bessere

Eigenschaften als das Riemann-Integral. Die Einführung ist jedoch technisch viel schwieriger und aufwändiger. Deswegen beschäftigen uns hier mit der ersten, ursprünglichen Variante, dem Jordan-Volumen.

### 3.1 Das Riemann-Integral auf Quadern

Führen wir also zunächst einen Integralbegriff in  $n$  Dimensionen ein, der die Flächeninhaltsberechnungen über Rechtecken/Volumenberechnungen über Quadern in 3 Dimensionen verallgemeinert.

**Definition 3.1:** Seien eine natürliche Zahl  $n \geq 1$  und reelle Zahlen  $a_1, \dots, a_n, b_1, \dots, b_n$  mit  $a_i \leq b_i$  für alle  $i = 1, \dots, n$  gegeben. Als *abgeschlossenen  $n$ -dimensionalen Quader* bezeichnen wir die Menge

$$Q := [a_1, b_1] \times \cdots \times [a_n, b_n] := \{x \in \mathbb{R}^n, a_i \leq x_i \leq b_i \text{ für alle } i = 1, \dots, n\} \quad (3.1)$$

und als *offenen  $n$ -dimensionalen Quader* bezeichnen wir die Menge

$$Q^0 := (a_1, b_1) \times \cdots \times (a_n, b_n) := \{x \in \mathbb{R}^n, a_i < x_i < b_i \text{ für alle } i = 1, \dots, n\}. \quad (3.2)$$

Das Volumen der Quader ist durch

$$\text{vol}(Q) := \text{vol}(Q^0) := \prod_{i=1}^n |b_i - a_i|$$

definiert.

**Definition 3.2:** Sei  $Q$  ein offener oder abgeschlossener Quader. Für jedes  $k \in \{1, \dots, n\}$  heißt

$$Z_k := \{a_k = x_{k,0} < \cdots < x_{k,A(k)} = b_k\} \quad (3.3)$$

Zerlegung von  $[a_k, b_k]$  und wir bezeichnen das  $n$ -Tupel von Mengen

$$Z := (Z_1, \dots, Z_n) \quad (3.4)$$

Zerlegung des Quaders  $Q$ . Sei  $\alpha = (\alpha_1, \dots, \alpha_n)$  ein  $n$ -Tupel von Indizes mit  $\alpha_k \in \{1, \dots, A(k)\}$  für alle  $k = 1, \dots, n$ . Dann ist

$$Q_\alpha := (x_{1,\alpha_1-1} - x_{1,\alpha_1}) \times \cdots \times (x_{n,\alpha_n-1} - x_{n,\alpha_n}) \quad (3.5)$$

der zu  $\alpha$  gehörende offene Teilquader von  $Q$ . Um zu kennzeichnen, dass der Quader  $Q_\alpha$  zur Zerlegung  $Z$  des Quaders  $Q$  gehört, schreiben wir

$$Q_\alpha \sqsubset Z.$$

Bemerke, dass die Wahl der  $\alpha_k$  von der Wahl der Zerlegung  $Z$  abhängt.

**Bemerkung 3.3:** Die Teilquader einer Zerlegung sind nach Definition disjunkt. Deswegen gilt

$$\text{vol}(Q) = \sum_{Q_\alpha \sqsubset Z} \text{vol}(Q_\alpha). \quad (3.6)$$

**Definition 3.4:** Sei  $f : Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine beschränkte Funktion auf dem Quader  $Q$ . Sei weiterhin eine Zerlegung  $Z$  von  $Q$  gegeben. Dann definiert

$$U(f; Z) := \sum_{Q_\alpha \sqsubset Z} \text{vol}(Q_\alpha) \inf\{f(x), x \in Q_\alpha\} \quad (3.7)$$

die *Untersumme* von  $f$  über  $Q$  mit Zerlegung  $Z$ , und mit

$$O(f; Z) := \sum_{Q_\alpha \sqsubset Z} \text{vol}(Q_\alpha) \sup\{f(x), x \in Q_\alpha\} \quad (3.8)$$

die *Obersumme* von  $f$  über  $Q$  mit Zerlegung  $Z$ .

Das *untere Integral* von  $f$  über  $Q$  ist nun durch

$$\underline{I}(f) := \sup\{U(f; Z), Z \text{ ist eine Zerlegung von } Q\}, \quad (3.9)$$

das *obere Integral* von  $f$  über  $Q$  durch

$$\bar{I}(f) := \inf\{O(f; Z), Z \text{ ist eine Zerlegung von } Q\} \quad (3.10)$$

definiert.

Die Funktion  $f$  heißt *Riemann-integrierbar*, falls  $\underline{I}(f) = \bar{I}(f)$ . Wir schreiben dann

$$\int_Q f \, \text{dvol} := \underline{I}(f) = \bar{I}(f) \quad (3.11)$$

und nennen diesen Wert das *Riemann-Integral* von  $f$  über  $Q$ .

**Lemma 3.5:** Sei  $Q$  ein Quader,  $Z = (Z_1, \dots, Z_n)$  eine Zerlegung von  $Q$  und  $Z' = (Z'_1, \dots, Z'_n)$  eine Verfeinerung von  $Z$ . Jedes  $Z'_i$  ist also eine Verfeinerung von  $Z_i$  für alle  $i = 1, \dots, n$ . Dann gilt

$$U(f; Z) \leq U(f; Z') \leq O(f; Z') \leq O(f; Z).$$

Daraus folgt für beschränkte Funktionen  $f : Q \rightarrow \mathbb{R}$  insbesondere

$$\underline{I}(f) = \bar{I}(f). \quad (3.12)$$

Beschränkte Funktionen sind also Riemann-integrierbar über Quadern.

### Eigenschaften des Riemann-Integrals

Das Riemann-Integral besitzt folgende Eigenschaften:



**Linearität** Sei  $Q \subset \mathbb{R}^n$  ein Quader und seien  $f, g : Q \rightarrow \mathbb{R}$  zwei Riemann-integrierbare reellwertige Funktionen, sowie  $\lambda, \mu \in \mathbb{R}$ . Dann ist auch die Funktion  $\lambda f + \mu g$  Riemann-integrierbar. Es gilt

$$\int_Q (\lambda f + \mu g) \, d\text{vol} = \lambda \int_Q f \, d\text{vol} + \mu \int_Q g \, d\text{vol}. \quad (3.13)$$

**Monotonie** Besitzt eine Funktion  $f$  nur nicht-negative Funktionswerte, so ist das Integral über einem beliebigen Quader nicht-negativ: Für beliebiges  $Q \subset \mathbb{R}^n$  und  $f : Q \rightarrow \mathbb{R}$  gilt

$$f \geq 0 \implies \int_Q f \, d\text{vol} \geq 0. \quad (3.14)$$

Wird eine Funktion durch eine andere dominiert, so gilt das auch für die Integrale der Funktionen über beliebigen Quadern: Für beliebiges  $Q \subset \mathbb{R}^n$  und  $f, g : Q \rightarrow \mathbb{R}$  gilt

$$f \leq g \implies \int_Q f \, d\text{vol} \leq \int_Q g \, d\text{vol}. \quad (3.15)$$

**Dreiecksungleichung** Sei  $f : Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  Riemann-integrierbar. Dann ist auch  $|f|$  Riemann-integrierbar. Es gilt

$$\left| \int_Q f \, d\text{vol} \right| \leq \int_Q |f| \, d\text{vol}. \quad (3.16)$$

Als nächstes geben wir Kriterien an, mit denen die Riemann-Integrierbarkeit bestimmter Funktionen überprüft werden kann. Die Kriterien sind also gewisse Charakterisierungen der Integrierbarkeit.

**Lemma 3.6:** Sei  $f : Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine beschränkte Funktion. Dann ist  $f$  genau dann Riemann-integrierbar, wenn für beliebiges  $\varepsilon > 0$  eine Zerlegung  $Z$  von  $Q$  existiert mit

$$O(f; Z) - U(f; Z) \leq \varepsilon. \quad (3.17)$$

Dieses Kriterium heißt auch *Riemannsches Integrierbarkeitskriterium*.

**Lemma 3.7:** Sei  $Q \subset \mathbb{R}^n$  abgeschlossen (und damit auch kompakt). Dann ist jede stetige Funktion  $f : Q \rightarrow \mathbb{R}$  Riemann-integrierbar.

**Definition 3.8:** Eine Teilmenge  $N \subset \mathbb{R}^n$  heißt *Nullmenge*, falls für jedes  $\varepsilon > 0$  eine abzählbare Menge von offenen Quadern  $\{Q_k \subset \mathbb{R}^n, k \in \mathbb{N}\}$  gibt, sodass

$$N \subset \bigcup_{k=1}^{\infty} Q_k \quad \text{und} \quad \sum_{k=1}^{\infty} \text{vol } Q_k < \varepsilon.$$

**Lemma 3.9:** Sei  $f : Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  eine beschränkte Funktion. Dann ist  $f$  genau dann Riemann-integrierbar, wenn die Menge

$$N := \{x \in Q, f \text{ ist unstetig in } x\} \subset \mathbb{R}^n$$

eine Nullmenge ist. In diesem Fall heißt  $f$  fast überall stetig.

Bemerkung: Die Verwendung der Beschreibung *fast überall* im Kontext der Integralrechnung ist mathematisch klar gefasst: Sei  $x \in \mathbb{R}^n$ . Dann gilt eine Aussage  $A(x)$  fast überall in  $\mathbb{R}^n$ , wenn die Menge

$$N := \{x \in X, A(x) \text{ gilt nicht}\}$$

eine Nullmenge ist. Die Aussage  $A$  gilt also überall, außer an endlich/abzählbar vielen Punkten.

**Beispiele** Es folgen ein paar Beispiele für Nullmengen.

- Die Menge  $\{x\} \subset \mathbb{R}^n$  ist eine Nullmenge für beliebiges  $x \in \mathbb{R}^n$ . Die Menge  $\{x\}$  ist nämlich in beliebig kleinen Quadern  $Q \subset \mathbb{R}^n$  enthalten.
- Jede abzählbare Menge von  $\mathbb{R}^n$  ist eine Nullmenge, so zum Beispiel auch  $\mathbb{Q}^n$ .
- Jeder Unterraum  $\mathbb{R}^k \subset \mathbb{R}^n$  mit  $0 \leq k < n$  ist eine Nullmenge des  $\mathbb{R}^n$ .
- Jede Teilmenge einer Nullmenge ist eine Nullmenge. Auch abzählbare Vereinigungen von Nullmengen sind Nullmengen.

## 3.2 Das Riemann-Integral von Funktionen über Jordan-messbaren Mengen

**Definition 3.10:** Eine Menge  $B \subset \mathbb{R}^n$  heißt *Jordan-messbar*, falls es einen Quader  $Q \subset \mathbb{R}^n$  mit  $B \subset Q$  gibt, sodass die charakteristische Funktion

$$\chi_B(x) = \begin{cases} 1, & \text{falls } x \in B \\ 0 & \text{sonst} \end{cases} \quad (3.18)$$

auf  $Q$  Riemann-integrierbar ist.

Ist  $B$  Jordan-messbar, so ist das *Jordan-Volumen* (auch: *Inhalt* oder *Jordan-Maß*) von  $B$  durch

$$\text{vol}(B) := \int_B \chi_B \, d\text{vol} \quad (3.19)$$

definiert.

**Bemerkung 3.11:** Das Volumen  $\text{vol}(B)$  ist wohldefiniert und hängt nicht von der Wahl des Quaders  $Q$  mit  $B \subset Q$  ab. (Bemerge, dass dies streng genommen bewiesen werden muss!)

**Lemma 3.12:** Eine Menge  $B \subset \mathbb{R}^n$  ist genau dann Jordan-messbar, wenn sie beschränkt und ihr Rand  $\partial B$  eine Nullmenge ist.

Falls zwei Mengen  $B_1, B_2 \subset \mathbb{R}^n$  Jordan-messbar sind, so ist auch  $B_1 \cup B_2$  sowie  $B_1 \setminus B_2$  Jordan-messbar.

**Definition 3.13:** Sei  $B \subset \mathbb{R}^n$  eine beschränkte Jordan-messbare Menge und  $f : B \rightarrow \mathbb{R}$  eine Funktion. Dann heißt  $f$  Riemann-integrierbar über  $B$ , falls es einen abgeschlossenen Quader  $Q \subset \mathbb{R}^n$  mit  $B \subset Q$  gibt, sodass die Funktion  $\chi_B f : Q \rightarrow \mathbb{R}$  mit

$$(\chi_B \cdot f)(x) := \begin{cases} f(x), & \text{falls } x \in B \\ 0, & \text{falls } x \in Q \setminus B \end{cases} \quad (3.20)$$

auf  $Q$  Riemann-integrierbar ist. Wir schreiben in diesem Fall  $f \in \mathcal{R}(B)$  und nennen

$$\int_B f \, d\text{vol} := \int_Q \chi_B \cdot f \, d\text{vol} \quad (3.21)$$

das Riemann-Integral von  $f$  über  $B$ .

**Bemerkung 3.14:** Analog zum Integral der charakteristischen Funktion über  $B$  hängt auch dieser Integralbegriff nicht von der Wahl des Quaders  $Q \subset \mathbb{R}^n$  mit  $B \subset Q$  ab.

Die im vorigen Abschnitt besprochenen Eigenschaften des Riemann-Integrals (Linearität, Monotonie, Dreiecksungleichung) übertragen sich auf das Integral von Riemann-integrierbaren Funktionen über Jordan-messbaren Mengen  $B \subset \mathbb{R}^n$ .

**Satz 3.15:** Sei  $B \subset \mathbb{R}^n$  Jordan-messbar und  $f : B \rightarrow \mathbb{R}$  beschränkt. Dann ist  $f$  genau dann Riemann-integrierbar, wenn  $f$  auf  $B$  fast überall stetig ist

**Lemma 3.16:** Seien  $B_1, B_2 \subset \mathbb{R}^n$  zwei Jordan-messbare Mengen und  $f : B_1 \cup B_2 \rightarrow \mathbb{R}$  eine beschränkte Riemann-integrierbare Funktion. Dann sind auch die Einschränkungen  $f|_{B_1}$ ,  $f|_{B_2}$  und  $f|_{B_1 \cap B_2}$  Riemann-integrierbar. Es gilt

$$\int_{B_1 \cup B_2} f \, d\text{vol} = \int_{B_1} f \, d\text{vol} + \int_{B_2} f \, d\text{vol} - \int_{B_1 \cap B_2} f \, d\text{vol}. \quad (3.22)$$

*Beweis.* Für die charakteristischen Funktionen gilt

$$\chi_{B_1 \cup B_2} = \chi_{B_1} + \chi_{B_2} - \chi_{B_1 \cap B_2}.$$

Die Aussage folgt dann aus der Linearität des Riemann-Integrals. □

**Definition 3.17:** Eine Menge  $N \subset \mathbb{R}^n$  heißt Jordan-Nullmenge, falls für jedes  $\varepsilon > 0$  eine endliche Menge von offenen Quadern  $Q_1, \dots, Q_K \subset \mathbb{R}^n$  gibt, sodass

$$N \subset \bigcup_{i=1}^K Q_i \quad \text{und} \quad \sum_{i=1}^K \text{vol } Q_i < \varepsilon.$$

**Satz 3.18:** Sei  $N \subset \mathbb{R}^n$ . Dann sind folgende Aussagen äquivalent:

1.  $N$  ist eine Jordan-Nullmenge.
2.  $N$  ist Jordan-messbar und  $\text{vol}(N) = 0$ .
3. Der Abschluss  $\bar{N} = N \cup \partial N$  ist beschränkt und eine Nullmenge.

Bemerkung: Es ist aus der Definition sofort klar, dass jede Jordan-Nullmenge eine Nullmenge ist. Allerdings sind nicht alle Nullmengen Jordan-Nullmengen. Nullmengen müssen nicht einmal Jordan-messbar sein. Zum Beispiel ist die Menge  $[0, 1] \cap \mathbb{Q} \subset \mathbb{R}$  aller rationalen Zahlen zwischen 0 und 1 abzählbar, also insbesondere eine Nullmenge. Sie ist aber keine Jordan-Nullmenge, da ihr Abschluss  $\overline{[0, 1] \cap \mathbb{Q}} = [0, 1]$  keine Nullmenge ist. Die Menge ist nach der Charakterisierung aus Lemma 3.12 nicht Jordan-messbar.

### 3.3 Der Satz von Fubini

Das Ziel dieses Abschnitts ist der Beweis, dass Integrale von gewissen Funktionen in  $\mathbb{R}^n$  sukzessive nach allen Komponenten integriert werden können, und dass die Reihenfolge der sukzessiven Integration keine Rolle spielt.

Um zu veranschaulichen, was damit gemeint ist, sei zunächst  $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$ . Sei außerdem  $f : Q \rightarrow \mathbb{R}$  stetig. Wir fixieren zunächst die erste Variable von  $f$ , betrachten also die parameterabhängige Funktion  $f_x : [c, d] \rightarrow \mathbb{R}$  mit  $f_x(y) := f(x, y)$  und betrachten das Integral über diese Funktion

$$\int_c^d f_x(y) \, dy = \int_c^d f(x, y) \, dy.$$

Die Frage, die sich unmittelbar stellt, ist, inwieweit der Wert dieses Integrals von  $x$  abhängt. Ändert sich der Wert des Integrals also auch nur gering, wenn  $x$  gering gestört wird ( $\rightarrow$  *stetige Abhängigkeit von  $x$* )? Ziel des Abschnitts ist es zu zeigen, dass das Integral der Funktion  $I_f : [a, b] \rightarrow \mathbb{R}$  mit

$$I_f(x) := \int_c^d f_x(x, y) \, dy = \int_c^d f(x, y) \, dy$$

gleich dem Integral von  $f$  über den Quader ist. Es soll also gezeigt werden, dass

$$\int_Q f \, d\text{vol} = \int_a^b I_f(x) \, dx = \int_a^b \int_c^d f(x, y) \, dy \, dx = \int_c^d \int_a^b f(x, y) \, dx \, dy$$

gilt. Ist dies der Fall, so schreiben wir statt  $d\text{vol}$  auch  $dy \, dx$  beziehungsweise  $dx \, dy$ . Im Falle  $n$ -dimensionaler Variablen schreiben wir  $d^n x$  statt  $dx_1 \dots dx_n$ .

**Satz 3.19:** Sei  $Q \subset \mathbb{R}^n$  ein abgeschlossener (und daher kompakter) Quader und  $\Omega \subset \mathbb{R}^m$  ein Gebiet. Sei zudem  $f : Q \times \Omega$  in allen Komponenten stetig. Dann ist auch die Funktion

$$F(y) := \int_Q f(x, y) \, d^n x \tag{3.23}$$

stetig in  $\Omega$ .

Ist  $f$  zudem überall stetig partiell differenzierbar in  $y$ , existieren also die Ableitungen  $\frac{\partial f}{\partial y_k}(x, y)$  in ganz  $Q \times \Omega$  und sind stetig für alle  $k = 1, \dots, m$ , so ist  $F$  eine  $C^1$ -Funktion auf  $\Omega$ . Außerdem gilt dann

$$\frac{\partial F}{\partial y_k}(y) = \int_Q \frac{\partial f}{\partial y_k}(x, y) \, d^n x. \quad (3.24)$$

*Beweis.* Wir fixieren  $y \in \Omega$  und wählen  $r > 0$ , sodass  $K := \overline{K_r(y)} \subset \Omega$ . Da  $\Omega \times K$  kompakt ist, ist  $f$  hier gleichmäßig stetig. Also gibt es für alle  $\varepsilon > 0$  ein  $\delta \in (0, r)$ , sodass für alle  $z, \tilde{z} \in K$

$$\|z - \tilde{z}\| < \delta \implies \|f(x, z) - f(x, \tilde{z})\| < \varepsilon$$

gilt, wobei  $x \in Q$  beliebig ist. Aus der Dreiecksungleichung für Integrale folgt

$$|F(z) - F(\tilde{z})| = \left| \int_Q f(x, z) - f(x, \tilde{z}) \, d^n x \right| \leq \int_Q |f(x, z) - f(x, \tilde{z})| \, d^n x < \varepsilon \operatorname{vol}(Q).$$

Damit ist die Stetigkeit von  $F$  gezeigt.

Sei nun  $f$  stetig partiell differenzierbar. Wir betrachten wieder nur den Fall  $m = 1$ , da der allgemeine Fall aus der Aussage für die Komponentenfunktionen folgt. Damit reduziert sich die zu betrachtende Menge auf  $K = [y - r, y + r] \subset \Omega$ . Sei nun  $\rho < r$  gewählt. Dann gibt es nach dem Mittelwertsatz für differenzierbare Funktionen für jedes  $x \in Q$  und  $\tilde{y} \in K$  mit  $|\tilde{y} - y| < \rho$  ein  $\theta \in (y - \rho, y + \rho)$ , sodass

$$\frac{\partial f}{\partial y}(x, \theta) = \left| \frac{f(x, \tilde{y}) - f(x, y)}{\tilde{y} - y} \right|.$$

Insbesondere ist

$$\begin{aligned} \left| \frac{F(\tilde{y}) - F(y)}{\tilde{y} - y} - \int_Q \frac{\partial f}{\partial y}(x, y) \, d^n x \right| &= \left| \int_Q \frac{f(x, \tilde{y}) - f(x, y)}{\tilde{y} - y} - \frac{\partial f}{\partial y}(x, y) \, d^n x \right| \\ &= \left| \int_Q \frac{\partial f}{\partial y}(x, \theta) - \frac{\partial f}{\partial y}(x, y) \, d^n x \right| \end{aligned}$$

Bilden wir den Limes  $\tilde{y} \rightarrow y$  auf beiden Seiten, so muss auch  $\theta \rightarrow y$  gehen. Wegen der Stetigkeit der partiellen Ableitungen verschwindet daher der Integrand auf der rechten Seite, mithin das Integral. Also muss auch die linke Seite verschwinden. Nach Definition und Eindeutigkeit des Differentials muss deswegen

$$\frac{dF}{dy}(y) = F'(y) = \int_Q \frac{\partial f}{\partial y}(x, y) \, d^n x.$$

Mit dem ersten Teil des Beweises folgt sogar die Stetigkeit der Ableitung von  $F$ .  $\square$

Bemerkung: Die Bedeutung der Aussage aus Satz 3.19 kann nicht genug betont werden. Es ist nämlich per se überhaupt nicht klar, dass Integration und Differentiation so einfach vertauscht werden dürfen.

**Satz 3.20 (Fubini):** Sei  $f : [a, b] \times [c, d] \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Dann gilt

$$\int_{[a,b] \times [c,d]} f(x, y) d^2(x, y) = \int_a^b \left( \int_c^d f(x, y) dy \right) dx = \int_c^d \left( \int_a^b f(x, y) dx \right) dy. \quad (3.25)$$

Analog gilt für kompakte Quader  $Q_1 \subset \mathbb{R}^p$  und  $Q_2 \subset \mathbb{R}^q$  und jedes stetige  $f : Q_1 \times Q_2 \rightarrow \mathbb{R}$

$$\int_{Q_1 \times Q_2} f(x, y) d^{p+q}(x, y) = \int_{Q_1} \left( \int_{Q_2} f(x, y) d^q y \right) d^p x = \int_{Q_2} \left( \int_{Q_1} f(x, y) d^p x \right) d^q y. \quad (3.26)$$

*Beweisidee.* Um die Aussage zu beweisen, betrachten wir zunächst die Funktion  $f_x : Q_2 \rightarrow \mathbb{R}$  mit  $f_x(y) := f(x, y)$ . Diese ist für alle  $x \in Q_1$  stetig, also Riemann-integrierbar. Insbesondere ist nach Satz 3.19 die Funktion  $F : Q_1 \rightarrow \mathbb{R}$

$$F(x) := \int_{Q_2} f(x, y) dy$$

Riemann-integrierbar und es gilt

$$\int_{Q_1} F(x, y) dx = \int_{Q_1} \int_{Q_2} f(x, y) dy dx.$$

Analog lässt sich die Aussage für  $f_y : Q_1 \rightarrow \mathbb{R}$  mit  $f_y(x) := f(x, y)$  und  $G : Q_2 \rightarrow \mathbb{R}$  mit

$$G(y) := \int_{Q_1} f(x, y) dx$$

zeigen. □

**Bemerkung 3.21:** Ist  $f$  nicht stetig sondern lediglich Riemann-integrierbar, so folgt die Riemann-Integrierbarkeit von  $f_x$  und  $f_y$  nicht. Sie muss stattdessen zusätzlich vorausgesetzt werden. Auch kann dann Satz 3.19 nicht verwendet werden, um die Riemann-Integrierbarkeit von  $F$  und/oder  $G$  zu zeigen. Diese lässt sich folgendermaßen zeigen: Wählen wir Zerlegungen  $Z_1$  von  $Q_1$  und  $Z_2$  von  $Q_2$ , so ist  $Z = (Z_1, Z_2)$  eine Zerlegung von  $Q_1 \times Q_2$ . Dann gilt

$$\underline{U}(f; Z) \leq U(F; Z_1)$$

und analog

$$\underline{O}(f; Z) \geq O(F; Z_1).$$

Da  $f$  auf  $Q_1 \times Q_2$  integrierbar ist, folgt nun, dass

$$\int_{Q_1 \times Q_2} f(x, y) d^{p+q}(x, y) = \sup_Z U(f; Z) = \inf_Z O(f; Z),$$

weswegen auch  $\sup U(F; Z_1) = \inf O(F; Z_1)$  sein muss. Also ist  $F$  Riemann-integrierbar. Analog funktioniert der Beweis der Riemann-Integrierbarkeit von  $G$ . Der Satz bliebe also auch unter den genannten schwächeren Voraussetzungen gültig.

**Korollar 3.22:** Sei  $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$  ein  $n$ -dimensionaler Quader und  $f : Q \rightarrow \mathbb{R}$  Riemann-integrierbar. Dann gilt

$$\int_Q f \, d\text{vol} = \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \dots, x_n) \, dx_1 \cdots dx_n. \quad (3.27)$$

**Korollar 3.23:** Sei  $D \subset \mathbb{R}^{n-1}$  eine Jordan-messbare Menge. Für beschränkte und stetige Funktionen  $h_-, h_+ : D \rightarrow \mathbb{R}$  mit  $h_- \leq h_+$  sei die Jordan-messbare Menge  $B$  definiert durch

$$B := \{(x, y) \in D \times \mathbb{R}, h_-(x) \leq y \leq h_+(x)\} \subset \mathbb{R}^n.$$

Das ist also die Menge, die auf  $D$  zwischen den Graphen der Funktionen  $h_-$  und  $h_+$  liegt. Dann gilt für eine Riemann-integrierbare Funktion  $f : B \rightarrow \mathbb{R}$

$$\int_B f(x, y) \, dx \, dy = \int_D \int_{h_-(x)}^{h_+(x)} f(x, y) \, dy \, dx, \quad (3.28)$$

wobei hier  $dx = dx_1 \cdots dx_{n-1}$ .

### Beispiele

1. Sei die Funktion  $f : [0, \pi]^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := \sin(x + y)$  gegeben. Sie ist stetig und daher Riemann-integrierbar. Dann ist

$$G(y) = \int_0^\pi \sin(x + y) \, dx = -\cos(x + y) \Big|_0^\pi = \cos y - \cos(y + \pi) = 2 \cos y$$

und deswegen

$$\int_{[0, \pi]^2} \sin(x + y) \, dx \, dy = \int_{[0, \pi]} G(y) \, dy = \int_0^\pi 2 \cos(y) \, dy = 2 \sin(y) \Big|_0^\pi = 0.$$

2. Wir betrachten die Menge  $B := \{(x, y) \in [0, 1]^2, x^2 \leq y \leq \sqrt{x}\}$ , die von den beiden Parabeln  $y = x^2$  und  $x = y^2$  begrenzt wird und oberhalb der  $x$ -Achse liegt. Wir möchten den Flächenschwerpunkt von  $B$  berechnen. Er entspricht dem Massenmittelpunkt eines physikalischen Körpers, der aus homogenem Material besteht. Die Masse ist also gleichmäßig auf  $B$  verteilt. In diesem Fall ist die Masse gleich dem Flächeninhalt, der sich als

$$\text{vol}(B) = \int_B d\text{vol} = \int_0^1 \sqrt{x} - x^2 \, dx = \frac{2x^{3/2} - x^3}{3} \Big|_0^1 = \frac{1}{3}$$

berechnet. Der Schwerpunkt  $S = (x_S, y_S)$  hat nach Definition die Koordinaten

$$x_S := \frac{1}{\text{vol}(B)} \int_B x \, d(x, y) \quad \text{und} \quad y_S := \frac{1}{\text{vol}(B)} \int_B y \, d(x, y).$$

Wir berechnen also

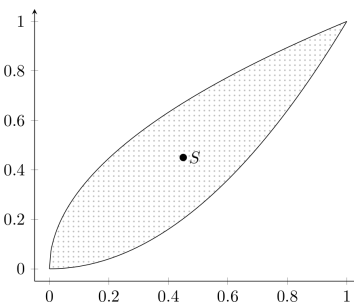
$$\begin{aligned} x_S &= \frac{1}{\text{vol}(B)} \int_B x \, d(x, y) = 3 \int_0^1 \int_{x^2}^{\sqrt{x}} x \, dy \, dx \\ &= 3 \int_0^1 [xy]_{x^2}^{\sqrt{x}} \, dx \\ &= 3 \int_0^1 x^{3/2} - x^3 \, dx = 3 \left[ \frac{2x^{5/2}}{5} - \frac{x^4}{4} \right]_0^1 = \frac{9}{20} \end{aligned}$$

sowie

$$\begin{aligned} y_S &= \frac{1}{\text{vol}(B)} \int_B y \, dx \, dy = 3 \int_0^1 \int_{x^2}^{\sqrt{x}} y \, dy \, dx \\ &= 3 \int_0^1 \left[ \frac{y^2}{2} \right]_{x^2}^{\sqrt{x}} \, dx \\ &= \frac{3}{2} \int_0^1 x - x^4 \, dx = \frac{3}{2} \left[ \frac{x^2}{2} - \frac{x^5}{5} \right]_0^1 = \frac{9}{20}. \end{aligned}$$

Der Flächenschwerpunkt liegt also bei  $S = (\frac{9}{20}, \frac{9}{20})$ .

Bemerke, dass die Grenzen für die  $y$ -Komponenten in  $B$  von  $x$  abhängen. Deswegen integrieren wir hier zunächst nach  $y$ . Alternativ ließe sich natürlich auch  $x$  in Abhängigkeit von  $y$  beschreiben, wobei dann zunächst nach  $x$  gelöst werden müsste. Die Reihenfolge der Integration spielt also tatsächlich keine Rolle; die Integrations-schranken müssen aber adäquat gewählt werden.



3. Wir möchten das Volumen der dreidimensionalen Einheitskugel berechnen. Dafür betrachten wir die Menge  $A := \{x \in \mathbb{R}^3, \|x\|^2 \leq 1\}$  mit Rand  $\partial A := \{x \in \mathbb{R}^3, \|x\| = 1\}$ . Es gilt

$$A \subset Q_1 := [-1, 1] \times [-1, 1] \times [-1, 1].$$

Die Menge  $A$  ist also beschränkt; ihr Rand ist eine Nullmenge. Insbesondere ist  $A$  Jordanmessbar.

Nun gilt

$$\text{vol}(A) = \int_Q \chi_A(x) \, d^3x = \int_A 1 \, d^3x = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 \chi_A \, dx \, dy \, dz.$$

aufgrund der Definition des Volumens und des Satzes von Fubini 3.20. Die Frage ist nun, wie wir die charakteristische Funktion  $\chi_A$  charakterisieren können. Dazu machen wir uns klar, dass die erste Variable  $x$  unabhängig von den anderen beiden Variablen  $y$  und  $z$  zwischen den Grenzen  $-1$  und  $1$  integriert werden kann. Für die zweite Variable  $y$  gilt nun nach Satz des Pythagoras in zwei Dimensionen, dass der entsprechende Punkt  $(x, y, z)$  nur dann innerhalb von  $A$  liegt, wenn  $y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}]$ . Analog muss für



fest gewählte  $x$  und  $y$  die dritte Variable  $z \in [-\sqrt{1-x^2-y^2}, \sqrt{1-x^2-y^2}]$  sein, damit der Punkt  $(x, y, z)$  in der Kugel liegt. Somit ist das Volumen durch

$$\text{vol}(A) = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} 1 \, dz \, dy \, dx$$

gegeben. Wir berechnen also

$$\begin{aligned} \text{vol}(A) &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} 1 \, dz \, dy \, dx \\ &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 2\sqrt{1-x^2-y^2} \, dy \, dx \end{aligned}$$



An dieser Stelle führen wir eine Variablensubstitution durch: Wir ersetzen  $y$  durch  $y(\varphi) := \sqrt{1-x^2} \sin(\varphi)$ . Das ist ein  $\mathcal{C}^1$ -Diffeomorphismus, wobei  $\frac{dy}{d\varphi}(\varphi) = \sqrt{1-x^2} \cos(\varphi)$ . Das neue Integrationsintervall ist nun  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ , da  $y(\pm\frac{\pi}{2}) = \pm\sqrt{1-x^2}$ . Damit ist

$$\begin{aligned} \text{vol}(A) &= 2 \int_{-1}^1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-x^2 - \left(\sqrt{1-x^2} \sin(\varphi)\right)^2} \underbrace{\sqrt{1-x^2} \cos(\varphi) \, d\varphi}_{=dy} \, dx \\ &= 2 \int_{-1}^1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{(1-x^2)(1-\sin^2(\varphi))} \sqrt{1-x^2} \cos(\varphi) \, d\varphi \, dx \\ &= 2 \int_{-1}^1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\sqrt{1-x^2}\right)^2 \cos^2(\varphi) \, d\varphi \, dx \\ &= \int_{-1}^1 (1-x^2) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1 \, d\varphi \, dx = \left[1 - \frac{x^3}{3}\right]_{-1}^1 \pi \, dx = \frac{4\pi}{3}. \end{aligned}$$

Das bestätigt die bereits aus der Schule bekannte Formel für das Volumen von Kugeln.

Bemerke: Die Gleichheit aus der vorletzten und der letzten Zeile folgt aus

$$(\cos(\varphi) \sin(\varphi))' = \cos^2(\varphi) - \sin^2(\varphi).$$

Die beidseitige Integration im Intervall  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  liefert

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(\varphi) \, d\varphi = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2(\varphi) \, d\varphi.$$

Aus  $\sin^2(\varphi) = 1 - \cos^2(\varphi)$  folgt nun, dass

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(\varphi) \, d\varphi = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2} \, d\varphi.$$

### Volumenberechnung von Rotationskörpern im $\mathbb{R}^3$

Sei  $r : [a, b] \rightarrow [0, R]$  eine stetige Funktion. Wir betrachten in der  $xz$ -Ebene die Fläche zwischen Graph der Funktion  $r$  und der  $x$ -Achse, also

$$\{(x, y, z) \in \mathbb{R}^3, x \in [a, b], y = 0, z \in [0, r(x)]\}.$$

Diese Fläche rotieren wir um die  $x$ -Achse, wobei der Rotationskörper

$$K := \{(x, y, z) \in \mathbb{R}^3, x \in [a, b], y^2 + z^2 \leq r(x)^2\}$$

entsteht. Bemerke, dass für jedes fixierte  $x$  der Schnitt von  $K$  mit der durch  $x$  und parallel zur  $yz$ -Achse verlaufenden Ebene ein Kreis mit Zentrum  $(x, 0, 0)$  und Radius  $r(x)$  ist. Da die Funktionswerte von  $r$  nach oben durch  $R$  beschränkt sind, gilt  $|y| \leq R$  für die  $y$ -Komponente aller Punkte  $(x, y, z) \in K$ .

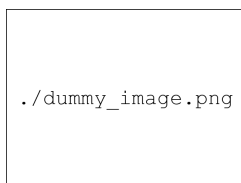
Wir definieren die Funktion  $f : [a, b] \times [-R, R] \rightarrow \mathbb{R}$  durch

$$f(x, y) := \begin{cases} \sqrt{r(x)^2 - y^2}, & \text{für } |y| \leq r(x) \\ 0, & \text{sonst, also für } |y| \in (r(x), R] \end{cases}.$$

Diese Funktion beschreibt also die Oberfläche des Rotationskörpers oberhalb der  $xy$ -Ebene: Für jeden Punkt  $(x, y)$  innerhalb der Schnittmenge von Rotationskörper und  $xy$ -Ebene gibt  $f$  die  $z$ -Koordinate zurück, sodass  $(x, y, z) \in \partial K$ . Damit ist klar, dass das Volumen des Rotationskörpers zweimal der Fläche zwischen dem Graphen von  $f$  und der  $xy$ -Ebene entspricht: Aufgrund der Rotation um die  $x$ -Achse fungiert die  $xy$ -Ebene wie eine Spiegelungsebene; sie teilt  $K$  in zwei kongruente Hälften. Es gilt also

$$\text{vol}(K) = 2 \int_a^b \int_{-R}^R f(x, y) \, dy \, dx = 2 \int_a^b \int_{-r(x)}^{r(x)} \sqrt{r(x)^2 - y^2} \, dy \, dx = \pi \int_a^b r(x)^2 \, dx.$$

Der Leser mache sich zur Übung die Gültigkeit der letzten Gleichheit klar. Als Hilfestellung sei erneut darauf hingewiesen, dass für fixiertes  $x$  der Graph von  $f$  einen Halbkreis mit Radius  $r(x)$  beschreibt.



#### Übungen

Zur Übung berechne man

- das Volumen des Körpers  $K_1$ , der durch Rotation der Fläche

$$\{(x, y, z) \in \mathbb{R}^3, x \in [0, 8], y \in [0, \sqrt[3]{x}], z = 0\}$$

und die  $x$ -Achse entsteht.

**Lösung:** Man setze hierfür  $r(x) = \sqrt[3]{x}$ . Dann ist  $y \in [0, 2]$  (also  $R = 2$ ) und wir definieren

$$f(x, y) := \begin{cases} \sqrt{r(x)^2 - y^2}, & \text{für } |y| \leq r(x) \\ 0, & \text{sonst, also für } |y| \in (r(x), 2] \end{cases}.$$

Es ergibt sich, dass

$$\text{vol}(K_1) = \pi \int_0^8 \sqrt[3]{x} \, dx = \frac{96\pi}{5}.$$

2. das Volumen einer Kugel  $B_R(0)$  mit Radius  $R$  und Zentrum im Ursprung als Volumen eines Rotationskörpers.

**Lösung:** Hierbei ist dann  $r(x) = \sqrt{R^2 - x^2}$ , sodass

$$\text{vol}(B) = \pi \int_{-R}^R (R^2 - x^2) \, dx = \frac{4\pi}{3} R^3.$$

### 3.4 Das Riemann-Integral stetiger Funktionen über offenen Mengen

Wir möchten nun die Definition der Integrierbarkeit von Funktionen über Jordan-messbare Mengen auf Funktionen auf beliebigen offenen Mengen ausweiten. Das schließt also ausdrücklich auch unbeschränkte Mengen ein.

**Definition 3.24:** Sei  $\Omega \subset \mathbb{R}^n$  eine beliebige offene nicht-leere Teilmenge des  $\mathbb{R}^n$ . Sei weiterhin eine Folge von kompakten Quadern  $(Q_i)_{i \in \mathbb{N}}$  mit  $\text{int } Q_i \cap \text{int } Q_j = \emptyset$  für alle  $i \neq j$  gegeben, die  $\Omega$  überdecken, also

$$\Omega = \bigcup_{i=1}^{\infty} Q_i. \tag{3.29}$$

Dann heißt eine solche Zerlegung *Quaderzerlegung von  $\Omega$* , wenn jede kompakte Teilmenge  $K \subset \Omega$  bereits von endlich vielen dieser Quader überdeckt wird.

**Satz 3.25:** Für jede beliebige offene nicht-leere Menge  $\Omega \subset \mathbb{R}^n$  gibt es eine solche Quaderzerlegung.

**Definition 3.26:** Sei  $\Omega \subset \mathbb{R}^n$  eine offene nicht-leere Menge mit der Quaderzerlegung  $(Q_i)_{i \in \mathbb{N}}$ . Dann definieren wir das  *$n$ -dimensionale Volumen von  $\Omega$*  durch

$$\text{vol}(\Omega) := \begin{cases} \sum_{i=1}^{\infty} \text{vol}(Q_i), & \text{falls die Reihe konvergiert;} \\ \infty & \text{andernfalls.} \end{cases} \tag{3.30}$$

Insbesondere folgt aus der Definition des Integrals über Quader (vgl. Definition 3.10)

$$\text{vol}(\Omega) = \int_{\Omega} 1 \, d^n x. \tag{3.31}$$

Bemerke: Die Definition des Volumens ist unabhängig von der Wahl der Quaderzerlegung.

**Bemerkung 3.27:** Auch unbeschränkte Mengen können ein endliches Volumen haben. Beispielsweise gilt für das Volumen der Menge  $\Omega := \{(x, y) \in \mathbb{R}^2, x > 1, 0 < y < \frac{1}{x^2}\}$

$$\text{vol}(\Omega) = \int_{\Omega} 1 \, d^2(x, y) = \int_1^{\infty} \int_0^{1/x^2} 1 \, dy \, dx = \int_1^{\infty} \frac{1}{x^2} \, dx = 1.$$

Das Beispiel dient lediglich der Illustration. Die Wohldefiniertheit des Integrals auf der rechten Seite ist nämlich bisher noch überhaupt nicht klar, und wird erst in den folgenden Abschnitten sichergestellt.

Wir haben jetzt die Definition des Volumens mithilfe von Quaderzerlegungen auf unbeschränkte Mengen erweitert. Kommen wir nun zur Integrierbarkeit von Funktionen über offenen (und potentiell unbeschränkten) Mengen. Dazu nutzen wir wieder die Quaderzerlegungen.

**Definition 3.28:** Sei  $\Omega \subset \mathbb{R}^n$  und  $f : \Omega \rightarrow \mathbb{R}$  stetig. Dann heißt  $f$  *Riemann-integrierbar über  $\Omega$* , wenn es eine Konstante  $M > 0$  gibt, sodass für alle  $k \in \mathbb{N}$  eine Auswahl endlich vieler kompakter und disjunkter Quader  $Q_1, \dots, Q_k \subset \Omega$  existiert mit

$$\sum_{i=1}^k \int_{Q_i} |f(x)| \, d^n x \leq M. \quad (3.32)$$

Ist  $f$  eine über  $\Omega$  Riemann-integrierbare Funktion und  $(Q_i)_{i \in \mathbb{N}}$  eine beliebige Quaderzerlegung von  $\Omega$ , so definieren wir das *Integral von  $f$  über  $\Omega$*  durch

$$\int_{\Omega} f := \int_{\Omega} f(x) \, d^n x := \lim_{k \rightarrow \infty} \sum_{i=1}^k \int_{Q_i} f(x) \, d^n x = \sum_{i=1}^{\infty} \int_{Q_i} f(x) \, d^n x < \infty. \quad (3.33)$$

**Satz 3.29:** Unter der Voraussetzung (3.32) ist der Grenzwert in (3.33) wohldefiniert und insbesondere unabhängig von der gewählten Quaderzerlegung.

*Beweis.* Zur Erinnerung: Eine Reihe ist absolut konvergent, wenn die Summe der Beträge der Folgenglieder konvergiert. Deswegen folgt die absolute Konvergenz der Reihe aus

$$\sum_{i=1}^k \left| \int_{Q_i} f(x) \, d^n x \right| \leq \sum_{i=1}^k \int_{Q_i} |f(x)| \, d^n x \leq M$$

für beliebiges  $k \in \mathbb{N}$ . □

**Bemerkung 3.30:** Ist  $f$  auf einem kompakten Quader  $Q$  stetig, so ist  $f$  über  $Q$  integrierbar und es gilt nach Definition 3.28

$$\int_Q f = \int_{\text{int } Q} f. \quad (3.34)$$

Der Vergleich mit Definition 3.13, insbesondere (3.21), zeigt, dass die Definition 3.28 als Verallgemeinerung des Integrals von Funktionen über Jordan-messbaren Mengen gesehen werden kann. Sie ist außerdem die direkte Verallgemeinerung der aus der Schule bekannten Definition des „uneigentlichen“ Integrals über offene Intervalle in  $\mathbb{R}$ .

**Satz 3.31:** Sei  $\Omega \subset \mathbb{R}^n$  offen und beschränkt. Jede beschränkte stetige Funktion  $f : \Omega \rightarrow \mathbb{R}$  ist über  $\Omega$  Riemann-integrierbar.

*Beweis.* Da  $\Omega$  beschränkt ist, gibt es einen kompakten Quader  $Q$  mit endlichem Inhalt und  $\Omega \subset Q$ . Außerdem existiert nach Voraussetzung eine Konstante  $c > 0$ , sodass  $|f(x)| \leq c$  für alle  $x \in \Omega$ . Sei  $N \in \mathbb{N}$  beliebig und wähle  $N$  kompakte und disjunkte Quader  $Q_1, \dots, Q_N \subset \Omega$ . Dann gilt

$$\sum_{k=1}^N \int_{Q_k} |f(x)| \, d^n x \leq c \sum_{k=1}^N \text{vol}(Q_k) \leq \text{vol}(Q) < \infty.$$

Damit ist  $f$  nach Definition Riemann-integrierbar über  $\Omega$ . □

### 3.4.1 Eigenschaften des Riemann-Integrals über offene Mengen

Das Riemann-Integral über offene Mengen  $\Omega \subset \mathbb{R}^n$  besitzt folgende Eigenschaften:

**Linearität** Seien  $f, g : \Omega \rightarrow \mathbb{R}$  zwei Riemann-integrierbare Funktionen über  $\Omega$ , sowie  $\lambda, \mu \in \mathbb{R}$ . Dann ist auch die Funktion  $\lambda f + \mu g$  Riemann-integrierbar. Es gilt

$$\int_{\Omega} (\lambda f + \mu g)(x) \, d^n x = \lambda \int_{\Omega} f(x) \, d^n x + \mu \int_{\Omega} g(x) \, d^n x. \quad (3.35)$$

**Monotonie** Seien  $f, g : \Omega \rightarrow \mathbb{R}$  zwei Riemann-integrierbare Funktionen und gelte  $f \leq g$  auf  $\Omega$ . Dann gilt

$$\int_{\Omega} f(x) \, d^n x \leq \int_{\Omega} g(x) \, d^n x. \quad (3.36)$$

**Dreiecksungleichung** Ist  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  über  $\Omega$  Riemann-integrierbar, so ist auch  $|f|$  über  $\Omega$  Riemann-integrierbar. Es gilt

$$\left| \int_{\Omega} f(x) \, d^n x \right| \leq \int_{\Omega} |f(x)| \, d^n x. \quad (3.37)$$

**Integrierbarkeit auf Teilmengen** Ist  $f : \Omega \rightarrow \mathbb{R}$  über  $\Omega$  Riemann-integrierbar, so ist  $f$  auch über jeder offenen Teilmenge  $\tilde{\Omega} \subset \Omega$  Riemann-integrierbar. Offensichtlich gilt dann

$$\int_{\tilde{\Omega}} |f(x)| \, d^n x \leq \int_{\Omega} |f(x)| \, d^n x \quad (3.38)$$

### 3.4.2 Kriterien für die Integrierbarkeit von Funktionen über offenen Mengen

#### 3.4.2.1 Zerlegungsbarkeitskriterium

Eine Funktion  $f : \Omega \rightarrow \mathbb{R}$  ist genau dann über  $\Omega$  Riemann-integrierbar, wenn beide Funktionen

$$f_+ := \frac{1}{2}(|f| + f) \quad \text{und} \quad f_- := \frac{1}{2}(|f| - f)$$

über  $\Omega$  Riemann-integrierbar sind. Es gilt

$$\int_{\Omega} f(x) \, d^n x = \int_{\Omega} f_+(x) \, d^n x - \int_{\Omega} f_-(x) \, d^n x. \quad (3.39)$$

#### 3.4.2.2 Majorantenkriterium

**Satz 3.32 (Majorantensatz):** Ist  $f : \Omega \rightarrow \mathbb{R}$  stetig und existiert eine Riemann-integrierbare Funktion  $g : \Omega \rightarrow \mathbb{R}$  mit  $|f| \leq g$  in  $\Omega$ , so ist  $f$  über  $\Omega$  integrierbar.

#### 3.4.2.3 Alternative Charakterisierung über Ausschöpfungen

Die Riemann-Integrierbarkeit von stetigen Funktionen über offenen Mengen lässt sich auch über sogenannte Ausschöpfungen mit offenen Mengen charakterisieren. Je nach Anwendungsfall ist zum Nachweis der Integrierbarkeit einer Funktion der Zugang über offene Mengen (Ausschöpfungen) oder abgeschlossene Mengen (Quaderzerlegung) intuitiver. Insbesondere rechtfertigt die nachfolgende Charakterisierung die Schreibweise aus Bemerkung 3.27.

**Definition 3.33:** Die offene und nicht-leere Menge  $\Omega \subset \mathbb{R}^n$  wird von den offenen Mengen  $(\Omega_i)_{i \in \mathbb{N}}$  ausgeschöpft, falls  $\Omega_i \subset \Omega_{i+1}$  für alle  $i \in \mathbb{N}$  und

$$\Omega = \bigcup_{i=1}^{\infty} \Omega_i, \quad (3.40)$$

gilt, und die Menge  $\bar{\Omega}_i \subset \Omega$  für alle  $i \in \mathbb{N}$  eine kompakte Teilmenge ist.

**Beispiel** Die Menge  $\mathbb{R}^2$  wird von den offenen Mengen  $\Omega_k := (-k, k) \times (-k, k)$  ausgeschöpft.

Allgemeiner: Sei  $(Q_i)_{i \in \mathbb{N}}$  eine Quaderzerlegung von  $\Omega$  und setze

$$\Omega_N := \text{int} \left( \bigcup_{i=1}^N Q_i \right).$$

Dann ist  $(\Omega_N)$  eine Ausschöpfung von  $\Omega$ .

**Satz 3.34:** Sei  $\Omega \subset \mathbb{R}^n$  und  $(\Omega_i)_{i \in \mathbb{N}}$  eine Ausschöpfung von  $\Omega$ . Dann ist eine stetige Funktion  $f : \Omega \rightarrow \mathbb{R}$  genau dann Riemann-integrierbar über  $\Omega$ , wenn die Folge der Integrale

$$\int_{\Omega_i} |f(x)| \, d^n x \tag{3.41}$$

beschränkt ist. In diesem Fall gilt, unabhängig von der Wahl der Ausschöpfung,

$$\int_{\Omega} f(x) \, d^n x = \lim_{i \rightarrow \infty} \int_{\Omega_i} f(x) \, d^n x. \tag{3.42}$$

*Beweis.* Ist  $f$  stetig und integrierbar auf  $\Omega$ , so ist

$$\int_{\Omega} |f(x)| \, dx < \infty.$$

Wegen der Eigenschaft der Integrierbarkeit auf Teilmengen ist dann für beliebiges  $i \in \mathbb{N}$  auch

$$\int_{\Omega_i} |f(x)| \, dx < \infty,$$

womit die Beschränktheit der Integrale gezeigt ist.

Sei nun für die Rückrichtung  $f$  stetig und die Folge der Integrale (3.41) beschränkt. Sei o. B. d. A.  $f \geq 0$ .<sup>1</sup> Aus der Monotonie des Integrals bezüglich Teilmengenbeziehungen, also

$$\int_{\Omega_i} f \, dx \leq \int_{\Omega_{i+1}} f \, dx$$

und aus der Beschränktheit der rechten Seite für alle  $i \in \mathbb{N}$  folgt die Existenz eines  $\alpha \geq 0$  mit

$$\lim_{i \rightarrow \infty} \int_{\Omega_i} f \, dx =: \alpha.$$

Wir bezeichnen die (nach Satz 3.25 existente) Quaderzerlegung von  $\Omega$  mit  $(Q_j)$ . Angenommen  $f$  sei nicht integrierbar über  $\Omega$ . Dann existiert nach Definition 3.28 für alle  $M > 0$  eine Auswahl von  $k_M$  disjunkten Quadern,  $k_M \in \mathbb{N}$ , sodass

$$M < \sum_{j=1}^{k_M} \int_{Q_j} f(x) \, dx.$$

Aufgrund der Eigenschaft der Ausschöpfung und der Offenheit vom  $\Omega$  existiert aber für jede Auswahl endlich vieler Quader aus der Quaderzerlegung ein  $\Omega_i$ , sodass  $\bigcup Q_j \subset \Omega_i$ . Wegen der Monotonie des Integrals bezüglich Teilmengenbeziehungen folgt damit

$$M < \sum_{j=1}^{k_M} \int_{Q_j} f(x) \, dx \leq \int_{\Omega_i} f(x) \, dx \leq \alpha.$$

---

<sup>1</sup>Bemerke zunächst, dass aus der Beschränktheit der Integrale von  $f$  über  $\Omega_i$  die Beschränktheit der Integrale von  $f_+$  und  $f_-$  über  $\Omega_i$  folgt (vgl. Zerlegbarkeitskriterium oben). Zudem lässt sich der Fall  $f \leq 0$  wegen  $-f_+ = f_-$  direkt auf den Fall  $f \geq 0$  zurückführen. Mit (3.39) folgt nun der allgemeine Fall.

Die rechte Seite ist unabhängig von  $M$  beschränkt. Da  $M$  aber beliebig groß werden kann, ergibt sich ein Widerspruch. Also war die Annahme falsch und  $f$  ist auf  $\Omega$  integrierbar.

Wir müssen nun noch die Beziehung (3.42) zeigen. Dazu überlegen wir uns, dass es für jedes  $\Omega_i$  aufgrund der Kompaktheit des Abschlusses  $\overline{\Omega}_i$  eine Auswahl von endlich vielen Quadern mit  $\Omega_i \subset \bigcup Q_j$  gibt. Die Vereinigung dieser endlich vielen Quader ist aber nun ihrerseits Teilmenge eines  $\Omega_k$  für  $k \in \mathbb{N}$ . Die Eigenschaft der Quaderzerlegung  $\Omega$  komplett auszufüllen, und eine Anwendung des Sandwicharguments liefern (3.42).  $\square$

**Beispiele** Sei  $n \in \mathbb{N}$ . Wir wollen wissen, ob das Integral

$$\int_{\mathbb{R}^n} \exp(-\|x\|_2^2) \, d^n x \quad (*)$$

existiert. Zunächst stellen wir fest, dass die Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(x) := \exp(-\|x\|_2^2)$  auf ganz  $\mathbb{R}^n$  stetig ist. Sei  $\Omega_k$  definiert als

$$\Omega_k := \{x \in \mathbb{R}^n, -k < x_i < k \text{ für alle } i = 1, \dots, n\}.$$

Dann bilden die  $\Omega_k$  eine Ausschöpfung von  $\mathbb{R}^n$ . Es gilt nach dem Satz von Fubini 3.20

$$\begin{aligned} \int_{\Omega_k} \exp(-\|x\|_2^2) \, d^n x &= \int_{-k}^k \cdots \int_{-k}^k \left( \int_{-k}^k \exp(-x_1^2) \cdots \exp(-x_n^2) \, dx_n \right) dx_{n-1} \cdots dx_1 \\ &= \prod_{i=1}^n \int_{-k}^k \exp(-x_i^2) \, dx_i \\ &= \left( \int_{-k}^k \exp(-t^2) \, dt \right)^n \leq \left( \int_{-\infty}^{\infty} \exp(-t^2) \, dt \right)^n =: c. \end{aligned}$$

Das Integral auf der rechten Seite existiert. Deswegen ist das Integral auf der linken Seite unabhängig von  $k$  durch  $c$  beschränkt. Damit sind die Voraussetzungen für Satz 3.34 erfüllt, und wir dürfen schreiben

$$\int_{\mathbb{R}^n} \exp(-\|x\|_2^2) \, d^n x = \lim_{k \rightarrow \infty} \int_{\Omega_k} \exp(-\|x\|_2^2) \, d^n x \leq \lim_{k \rightarrow \infty} \left( \int_{-k}^k \exp(-t^2) \, dt \right)^n = c.$$

Also existiert das Integral (\*).

Sei nun  $f : \mathbb{R}^3 \rightarrow \mathbb{R} \setminus \{0\}$  gegeben durch

$$f(x) := \frac{x_1 \cdot x_2 \cdot x_3}{\|x\|_2^5}.$$

Diese Funktion hat einen Pol im Ursprung. Wir möchten diese Funktion über der Menge  $D := (0, 1)^3$  integrieren. Zunächst stellen wir fest, dass  $f$  nach Satz 3.31 auf  $D$  integrierbar ist. Das Integral ist also endlich. Wir definieren

$$\Omega_{\varepsilon, \delta} := \{x \in \mathbb{R}^3, \varepsilon < x_i < \delta \text{ für } i = 1, 2, 3\} \subset D,$$



sodass die Mengen  $\Omega_{\varepsilon,\delta}$  für  $\varepsilon = \frac{1}{k}$  und  $\delta = 1 - \frac{1}{k}$  eine Ausschöpfung von  $D$  bilden. Wegen der Integrierbarkeit von  $f$  auf Teilmengen und Satz 3.34 gilt

$$\int_D f \, d^3x = \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 1} \int_{\Omega_{\varepsilon,\delta}} f \, d^3x.$$

Berechnen wir also das Integral von  $f$  über  $\Omega_{\varepsilon,\delta}$  mithilfe des Satzes von Fubini 3.20:

$$\int_{\Omega_{\varepsilon,\delta}} f(x) \, d^3x = \int_{\varepsilon}^{\delta} \int_{\varepsilon}^{\delta} \int_{\varepsilon}^{\delta} \frac{x_1 x_2 x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \, dx_3 \, dx_2 \, dx_1$$

Nun ist  $f$  auf dem Definitionsbereich stetig, weswegen der Grenzwert für  $\delta \rightarrow 1$  definiert ist. Es gilt

$$\begin{aligned} \lim_{\delta \rightarrow 1} \int_{\Omega_{\varepsilon,\delta}} f(x) \, d^3x &= \int_{\varepsilon}^1 \int_{\varepsilon}^1 \int_{\varepsilon}^1 \frac{x_1 x_2 x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \, dx_3 \, dx_2 \, dx_1 \\ &= \frac{1}{3} \int_{\varepsilon}^1 \int_{\varepsilon}^1 \left[ -\frac{x_1 x_2}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} \right]_{x_3=\varepsilon}^{x_3=1} \, dx_2 \, dx_1 \\ &= \frac{1}{3} \int_{\varepsilon}^1 \int_{\varepsilon}^1 -\frac{x_1 x_2}{(x_1^2 + x_2^2 + 1^2)^{3/2}} + \frac{x_1 x_2}{(x_1^2 + x_2^2 + \varepsilon^2)^{3/2}} \, dx_2 \, dx_1 \\ &= \frac{1}{3} \int_{\varepsilon}^1 \left[ \frac{x_1}{(x_1^2 + x_2^2 + 1^2)^{1/2}} - \frac{x_1}{(x_1^2 + x_2^2 + \varepsilon^2)^{1/2}} \right]_{x_2=\varepsilon}^{x_2=1} \, dx_1 \\ &= \frac{1}{3} \int_{\varepsilon}^1 \frac{x_1}{(x_1^2 + 2)^{1/2}} - \frac{2x_1}{(x_1^2 + 1^2 + \varepsilon^2)^{1/2}} + \frac{x_1}{(x_1^2 + 2\varepsilon^2)^{1/2}} \, dx_1 \\ &= \frac{1}{3} \left[ (x_1^2 + 2)^{1/2} - 2(x_1^2 + 1^2 + \varepsilon^2)^{1/2} + (x_1^2 + 2\varepsilon^2)^{1/2} \right]_{x_1=\varepsilon}^{x_1=1} \\ &= \frac{1}{3} (\sqrt{3} + 3\sqrt{1 + 2\varepsilon^2} - 3\sqrt{\varepsilon^2 + 2} - \sqrt{3\varepsilon^2}) \end{aligned}$$

Betrachten wir nun den Grenzwert für  $\varepsilon \rightarrow 0$ , so finden wir, dass

$$\int_D f \, d^3x = \frac{1}{3} (\sqrt{3} + 3 - 3\sqrt{2}).$$

### 3.4.2.4 Kriterium für die Integration über $\mathbb{R}^2$

Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Außerdem setzen wir voraus, dass

- a) zu jedem kompakten Intervall  $I$  eine Riemann-integrierbare Majorante  $g_I : \mathbb{R} \rightarrow \mathbb{R}$  existiert, sodass

$$|f(x, y)| \leq g_I(y) \tag{3.43}$$

für alle  $x \in I$  und alle  $y \in \mathbb{R}$  gilt.

b) die Funktion  $|F| : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$|F|(x) := \int_{-\infty}^{\infty} |f(x, y)| \, dy \quad (3.44)$$

Riemann-integrierbar über ganz  $\mathbb{R}$  ist.

Unter diesen Voraussetzungen sind die Funktionen  $F : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$F(x) := \int_{-\infty}^{\infty} f(x, y) \, dy$$

und  $|F|$  stetig. Außerdem gilt

$$\int_{\mathbb{R}^2} f(x, y) \, d^2(x, y) = \int_{-\infty}^{\infty} F(x) \, dx = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(x, y) \, dy \right) \, dx. \quad (3.45)$$

**Satz 3.35 (Fubini in 2D):** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Falls sowohl  $f$  als auch  $\tilde{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $\tilde{f}(x, y) := f(y, x)$  das Kriterium für die Integration über  $\mathbb{R}^2$  erfüllen, dann kann die Integrationsreihenfolge vertauscht werden. Dann gilt also

$$\int_{\mathbb{R}^2} f(x, y) \, d^2(x, y) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(x, y) \, dy \right) \, dx = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(x, y) \, dx \right) \, dy. \quad (3.46)$$

**Beispiel** Zunächst betrachten wir ein Gegenbeispiel für die Integration über  $\mathbb{R}^2$ . Theorem 3.35 fordert unmissverständlich, dass die Funktion  $f$  in beiden Variablen die Kriterien aus Abschnitt 3.4.2.4 erfüllen muss. Um die Notwendigkeit dieser Forderung einzusehen, betrachten wir folgendes Beispiel. Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $f(x, y) := \exp(-|y|(1+x^2))$  definiert. Dann gilt für die kompakten Intervalle  $I_n := [-n, n]$  und alle  $x \in I_n$

$$|f(x, y)| = \exp(-|y|) \exp(-|y|x^2) \leq \exp(-|y|) =: g_{I_n}(y).$$

Insbesondere ist also das Kriterium a) erfüllt. Außerdem ist die Funktion

$$\begin{aligned} F(x) = |F|(x) &= \int_{-\infty}^{\infty} |f(x, y)| \, dy = \int_{-\infty}^{\infty} \exp(-|y|(1+x^2)) \, dy \\ &= 2 \int_0^{\infty} \exp(-y(1+x^2)) \, dy \\ &= 2 \left[ -\frac{1}{1+x^2} \exp(-y(1+x^2)) \right]_{y=0}^{y=\infty} \\ &= \lim_{y \rightarrow \infty} -\frac{2}{1+x^2} \exp(-y(1+x^2)) + \frac{2}{1+x^2} = \frac{2}{1+x^2}. \end{aligned}$$

über  $\mathbb{R}$  integrierbar. Das erfüllt Kriterium b). Auf der anderen Seite ist  $f(x, 0) = 1$ , sodass

$$|\tilde{F}(0)| = \int_{-\infty}^{\infty} |f(x, 0)| \, dx = \int_{-\infty}^{\infty} 1 \, dx.$$

Die Integrationsvariablen dürfen also nicht vertauscht werden.

### 3.4.3 Transformationssatz

Wir kommen nun noch zu einem wichtigen Satz für die Berechnung von Integralen. Mit seiner Hilfe kann das Riemann-Integral über „komplizierte Mengen“ auf das Integral über „einfachere“ Mengen, beispielsweise Würfel, zurückgeführt werden. Dazu verwendet man die bereits kennengelernte Koordinatentransformation.

**Satz 3.36:** Seien  $U, V \subset \mathbb{R}^n$  offene Mengen und  $\varphi : U \rightarrow V$  ein  $\mathcal{C}^1$ -Diffeomorphismus. Dann gilt für alle stetigen Funktionen  $f \in \mathcal{C}(V)$

$$\int_V f(y) \, d^n y = \int_{\varphi(U)} f(y) \, d^n y = \int_U f(\varphi(x)) \cdot |\det J\varphi(x)| \, d^n x. \quad (3.47)$$

**Beispiel** Sei  $0 < a < b$  und

$$D := \{(x, y) \in \mathbb{R}^2, a \leq \sqrt{x^2 + y^2} \leq b, x, y \geq 0\}.$$

Die Menge  $D$  beschreibt also den Teil der Kreisscheibe mit innerem Radius  $a$  und äußerem Radius  $b$ , der im ersten Quadranten liegt. Wir möchten das Integral

$$\int_D \ln(x^2 + y^2) \, d(x, y)$$

berechnen. Wir verwenden hierzu den  $\mathcal{C}^1$ -Diffeomorphismus  $\varphi : \Omega := [a, b] \times [0, \frac{\pi}{2}] \rightarrow D$  aus Abschnitt 2.8 mit

$$\varphi(r, \theta) := \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}. \quad (3.48)$$

Dann ist die Jacobi-Matrix von  $\varphi$  durch

$$J\varphi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix};$$

ihre Determinante ist  $r$ . Damit ist

$$\begin{aligned} \int_D f(x, y) \, d(x, y) &= \int_{\Omega} \ln(r^2) |\det J\varphi(r, \theta)| \, d(r, \theta) \\ &= \int_a^b \int_0^{\pi/2} 2r \ln(r) \, d\theta \, dr \\ &= \frac{\pi}{2} \int_a^b 2r \ln(r) \, dr \\ &= \frac{\pi}{2} \left( \left[ r^2 \ln(r) \right]_a^b - \int_a^b r \, dr \right) = \pi \left( \frac{b^2 \ln(b) - a^2 \ln(a)}{2} - \frac{b^2 - a^2}{4} \right). \end{aligned}$$

Dabei wurden in der letzten Zeile die aus Analysis I bekannte Formel für die partielle Integration verwendet.

### Geometrische Interpretation der Determinante

Sei  $P$  das von den linear unabhängigen Vektoren  $a_1, \dots, a_n \in \mathbb{R}^n$  und dem Stützvektor  $a \in \mathbb{R}^n$  aufgespannte Parallelepipid, also

$$P = a + \left\{ \sum_{i=1}^n \lambda_i a_i, 0 \leq \lambda_i \leq 1 \right\}.$$

Für  $a = 0 \in \mathbb{R}^n$  heißt  $P$  auch Parallelotop. Dann ist das Volumen von  $P$  gegeben durch

$$\text{vol}(P) = |\det(a_1, \dots, a_n)|. \quad (3.49)$$

Zum Nachweis sei  $A$  die Matrix, die die  $a_i$  als Spaltenvektoren enthält. Wir definieren die Funktion  $\varphi : Q \rightarrow P$  durch  $\varphi(x) := a + Ax$ , wobei  $Q$  den  $n$ -dimensionalen Einheitswürfel bezeichnet. Da die Matrix  $A$  regulär ist, ist  $\varphi$  ein Diffeomorphismus, insbesondere  $\varphi(Q) = P$ . Nach Transformationsformel gilt dann

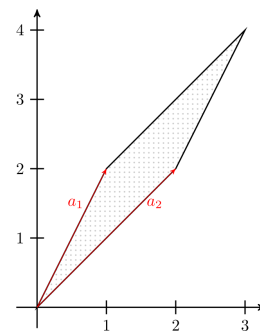
$$\text{vol}(P) = \int_P 1 \, dy = \int_Q 1 \cdot |\det(A)| \, dx = \det(A) \cdot \text{vol}(Q) = \det(A).$$

**Beispiel** Sei  $n = 2$ , also  $Q$  der Würfel mit den Eckpunkten  $(0, 0)$ ,  $(1, 0)$ ,  $(1, 1)$  und  $(0, 1)$ . Außerdem sei  $a_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  und  $a_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ . Wir definieren  $\varphi : Q \rightarrow P$  durch

$$\varphi(x) = Ax = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Dann ist

$$\text{vol}(P) = \int_P 1 \, dy = \int_Q |\det(A)| \, dx = 2.$$



**Bemerkung 3.37:** Für eine affine Bewegung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $T(x) := a + Ax$  mit  $\det A = 1$  und jedes  $\Omega \subset \mathbb{R}^n$

$$\text{vol}(\Omega) = \text{vol}(T(\Omega)). \quad (3.50)$$

# Kapitel 4

## Grundlagen der Vektoranalysis

In diesem Kapitel betrachten wir Integrale von Vektorfeldern  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  über Kurven und Flächen und leiten damit fundamentale Sätze der Vektoranalysis wie den Satz von Gauß und den Satz von Stokes her.

### 4.1 Kurvenintegrale

Genauso, wie sich Funktionen entlang von Wegen differenzieren lassen, können sie natürlich auch entlang von Wegen integriert werden. Im nachfolgenden sind die Begriffe Kurve und Wege gleichbedeutend.

**Definition 4.1:** Eine  $\mathcal{C}^1$ -Kurve  $\alpha : I \rightarrow \mathbb{R}^n$  mit  $\alpha(t) := (\alpha_1(t), \dots, \alpha_n(t))^T$  ist *regulär*, wenn der Tangentenvektor  $\dot{\alpha}(t) \neq 0$  für alle  $t \in I$ .

Analog definieren wir  $\mathcal{C}^k$ -Kurven mit  $k \in \mathbb{N}$ .

**Beispiele** Die archimedische Spirale  $\alpha : [0, T] \rightarrow \mathbb{R}^2$  mit

$$\alpha(t) := \begin{pmatrix} t \cos(t) \\ t \sin(t) \end{pmatrix}$$

ist regulär für beliebiges  $T \in \mathbb{R}$ .

Die Schraubenlinie  $\beta : [0, T] \rightarrow \mathbb{R}^3$  mit

$$\beta(t) := \begin{pmatrix} r \cos(t) \\ r \sin(t) \\ ht \end{pmatrix}$$

ist regulär für beliebiges  $T \in \mathbb{R}$ , falls  $r^2 + h^2 > 0$ .

Die Zykloide  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  mit

$$\gamma(t) := \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix}$$

ist nicht regulär, da  $\dot{\alpha}(t) = (1 - \cos(t), \sin(t)) = (0, 0)$  für  $t = 2\pi n$ ,  $n \in \mathbb{N}$ .

**Definition 4.2:** Eine Menge  $C \subset \mathbb{R}^n$  heißt  $\mathcal{C}^k$ -Kurvenstück, wenn  $C$  das Bild einer injektiven regulären  $\mathcal{C}^k$ -Kurve  $\alpha : [a, b] \rightarrow \mathbb{R}^n$  ist. Für  $C$  heißt eine jede solche Kurve  $\alpha$  mit  $\mathcal{C}^k$ -Parametrisierung von  $C$ .

**Bemerkung 4.3:** Zu je zwei  $\mathcal{C}^k$ -Parametrisierungen  $\alpha : [a, b] \rightarrow \mathbb{R}^n$  und  $\beta : [c, d] \rightarrow \mathbb{R}^n$  eines Kurvenstücks  $C \subset \mathbb{R}^n$  existiert ein  $\mathcal{C}^k$ -Diffeomorphismus  $\varphi : [a, b] \rightarrow [c, d]$ , sodass

$$\alpha(t) = \beta \circ \varphi(t) \quad (4.1)$$

für alle  $t \in [a, b]$  gilt.

**Definition 4.4:** Für ein Kurvenstück  $C \subset \mathbb{R}^n$  definieren wir die *Länge*  $L(C)$  von  $C$  als Supremum der Länge aller in  $C$  einbeschriebenen Sehnenpolynome.

**Satz 4.5:** Für jede  $\mathcal{C}^1$ -Parametrisierung  $\alpha : [a, b] \rightarrow \mathbb{R}^n$  eines Kurvenstücks  $C$  gilt

$$L(C) := L(\alpha) := \int_a^b \|\dot{\alpha}(t)\| dt. \quad (4.2)$$

**Bemerkung 4.6:** Die Aussage des Satzes 4.5 ist unabhängig von der gewählten Parametrisierung. Ist nämlich  $\beta : [c, d] \rightarrow \mathbb{R}^n$  eine andere Parametrisierung von  $C$  und  $\varphi$  der entsprechende  $\mathcal{C}^1$ -Diffeomorphismus (vgl. Bemerkung 4.3), dann gilt

$$L(\alpha) = \int_a^b \|\dot{\alpha}(t)\| dt = \int_a^b \|\dot{\beta}(h(t))\dot{h}(t)\| dt = \int_a^b \|\dot{\beta}(h(t))\| \|\dot{h}(t)\| dt = \int_c^d \|\dot{\beta}(s)\| ds = L(\beta).$$

**Beispiele** Sei  $\alpha$  der Graph einer Funktion  $y : [a, b] \rightarrow \mathbb{R}$ , also  $\alpha(t) := (t, y(t))$ . Dann ist

$$L(\alpha) := \int_a^b \sqrt{1 + (\dot{y}(t))^2} dt.$$

Seien  $x_0, y_0 \in \mathbb{R}$  gegeben und definiere  $\alpha : [a, b] \rightarrow \mathbb{R}^2$  durch

$$\alpha(t) := \begin{pmatrix} x_0 + r(t) \cos(\omega(t)) \\ y_0 + r(t) \sin(\omega(t)) \end{pmatrix},$$

wobei  $r, \omega : [a, b] \rightarrow \mathbb{R}$  gegebene Radius- und Winkelfunktionen sind. Dann ist

$$L(\alpha) = \int_a^b \sqrt{(\dot{r}(t))^2 + r^2(t)(\dot{\omega}(t))^2} dt.$$

Insbesondere beschreibt für  $r = 1$  und  $\omega = \text{id}$  die Funktion  $\alpha$  einen Kreisabschnitt. Für das Intervall  $[a, b] = [0, 2\pi)$  ist dann

$$L(\alpha) = \int_0^{2\pi} r dt = 2\pi r$$

genau der Kreisumfang.

### 4.1.1 Skalare Kurvenintegrale

**Definition 4.7:** Sei  $C$  ein Kurvenstück und  $f : C \rightarrow \mathbb{R}$  eine Funktion, sodass für mindestens eine Parametrisierung  $\alpha$  von  $C$  die Abbildung  $f \circ \alpha$  stetig ist. Das *skalare Kurvenintegral* (*Wegintegral*) von  $f$  über  $C$  ist definiert als

$$\int_C f \, ds := \int_a^b f(\alpha(t)) \|\dot{\alpha}(t)\| \, dt. \quad (4.3)$$

Hierbei bezeichnet  $ds$  das Integral über das skalare Bogenelement.

Das Integral ist unabhängig von der Wahl der Parametrisierung (vgl. Bemerkung 4.6).

**Beispiel** Sei eine Kurve  $C \subset \mathbb{R}^2$  parametrisiert durch  $\alpha : [0, \pi/2] \rightarrow \mathbb{R}^2$  gegeben durch

$$\alpha(t) := (-3 \cos^3(t), 3 \sin^3(t)).$$

Dann ist die Ableitung gegeben als

$$\dot{\alpha}(t) = (-9 \cos^2(t) \sin t, 9 \sin^2(t) \cos(t))$$

und somit  $\|\dot{\alpha}(t)\| = 9 \sin(t) \cos(t)$ . Integrieren wir die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit  $f(x, y) := 1 + \frac{y}{3}$  über  $C$ , ergibt sich

$$\begin{aligned} \int_C f \, ds &= \int_0^{\pi/2} \left( 1 + \frac{3 \sin^3(t)}{3} \right) \cdot 9 \sin(t) \cos(t) \, dt \\ &= 9 \int_0^{\pi/2} (\sin(t) + \sin^4(t)) \cos(t) \, dt = 9 \left[ \frac{\sin^2(t)}{2} + \frac{\sin^5(t)}{5} \right]_0^{\pi/2} = \frac{63}{5} \end{aligned}$$

### 4.1.2 Vektorielle Kurvenintegrale

Eine stetige Abbildung  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  heißt *Vektorfeld*. Vektorfelder sind beispielsweise Geschwindigkeitsfelder von Gasen oder Flüssigkeiten, Gravitationsfelder, elektrische oder magnetische Felder und so weiter. Sie weisen jedem Punkt im Raum einen Vektor im Raum zu. Betrachten wir beispielsweise die Strömung in einem Fluss, gibt das Vektorfeld an, wohin und mit welcher Geschwindigkeit sich das an einem Ort  $x$  befindliche Wasserteilchen bewegt.

Ist ein Vektorfeld, also alle seine Komponenten,  $k$ -mal stetig differenzierbar, so sprechen wir auch von einem  $\mathcal{C}^k$ -Vektorfeld.

**Definition 4.8:** Sei  $\alpha : [a, b] \rightarrow \mathbb{R}^n$  eine reguläre  $\mathcal{C}^1$ -Kurve mit Bild in  $\Omega$  und  $v : \Omega \rightarrow \mathbb{R}^n$  ein Vektorfeld. Dann definieren wir das *Wegintegral von  $v$  über den von  $\alpha$  parametrisierten Weg* durch

$$\int_\alpha v \cdot ds := \int_a^b \langle v(\alpha(t)), \dot{\alpha}(t) \rangle \, dt. \quad (4.4)$$

### Physikalische Entsprechung

- Beschreibt das Vektorfeld ein Kraftfeld, entspricht das Wegintegral der Arbeit.
- Beschreibt das Vektorfeld ein Geschwindigkeitsfeld, entspricht das Wegintegral der Zirkulation.
- Beschreibt das Vektorfeld eine elektrische Feldstärke, entspricht das Wegintegral der elektrischen Spannung.
- Beschreibt das Vektorfeld eine Temperaturänderung, entspricht das Wegintegral der Temperatur.

Sei also beispielsweise ein Massepunkt im Ursprung  $(0, 0, 0)$  gegeben. Er erzeugt ein Gravitationsfeld  $K : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}$ , das durch

$$K(x) := -\frac{x}{\|x\|^3} = -(x_1^2 + x_2^2 + x_3^2)^{-3/2} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

gegeben ist. Wird ein zweiter Massepunkt der Masse 1 längs einer Kurve  $\alpha : [a, b] \rightarrow \mathbb{R}^3 \setminus \{0\}$  bewegt, so ist die an ihm geleistete Arbeit genau das Wegintegral

$$\int_{\alpha} K \, ds = - \int_a^b \frac{x_1(t)\dot{x}_1(t) + x_2(t)\dot{x}_2(t) + x_3(t)\dot{x}_3(t)}{(x_1^2(t) + x_2^2(t) + x_3^2(t))^{3/2}} dt.$$

**Bemerkung 4.9:** Sei  $C \subset \Omega$  ein  $\mathcal{C}^1$ -Kurvenstück in  $\Omega$  und seien  $\alpha, \beta$  zwei Parametrisierungen von  $C$ . Dann gilt die Kettenregel

$$\int_{\alpha} v \, ds = \int_{\alpha} \langle v(\beta \circ \varphi(t)), \dot{\beta}(\varphi(t))\dot{\varphi}(t) \rangle dt = \begin{cases} \int_{\beta} v \, ds, & \text{falls } \dot{\varphi} > 0 \\ - \int_{\beta} v \, ds, & \text{falls } \dot{\varphi} < 0 \end{cases} \quad (4.5)$$

für die Parametertransformation  $\varphi$  (vgl. Bemerkung 4.6).

**Definition 4.10:** Sei  $C$  ein  $\mathcal{C}^1$ -Kurvenstück. Als *Orientierung von  $C$*  bezeichnen wir eine Festlegung des Durchlaufsinns. Sie besteht darin, eine bestimmte  $\mathcal{C}^1$ -Parametrisierung  $\alpha$  festzulegen und dann jede weitere  $\mathcal{C}^1$ -Parametrisierung  $\beta$  von  $C$  als (ebenfalls) positive Parametrisierung von  $C$  zu bezeichnen, falls  $\beta = \alpha \circ \varphi$  mit  $\dot{\varphi} > 0$ .

Bei der Änderung der Parametrisierung von  $\alpha$  zu  $\tilde{\beta} := \alpha \circ \varphi$  mit  $\dot{\varphi} < 0$  sprechen wir von einer Änderung des Durchlaufsinns.



**Beispiel** Sei ein Vektorfeld in  $\mathbb{R}^2$  gegeben durch  $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $v(x, y) := (x^2, xy)^T$ . Weiterhin seien  $\gamma_1, \bar{\gamma}_1 : [0, 1] \rightarrow \mathbb{R}^2$  zwei Wege, wobei  $\gamma_1(t) := (t, t)^T$  und  $\bar{\gamma}_1(t) = (1-t, 1-t)^T$ . Dann parametrisieren beide Wege die Strecke (das Kurvenstück) von  $(0, 0)$  nach  $(1, 1)$ . Wir legen die Orientierung über  $\gamma_1$  fest, sodass also  $\gamma_1$  den positiven Richtungssinn bestimmt. Dementsprechend besitzt  $\bar{\gamma}_1$  negativen Richtungssinn. Es gilt nämlich

$$\bar{\gamma}_1 = \gamma_1 \circ h$$

für  $h : \mathbb{R} \rightarrow \mathbb{R}$  mit  $h(t) := 1 - t$  und  $\dot{h} = -1$ .

Berechnen wir nun das Integral des Vektorfeldes über das Kurvenstück mithilfe der beiden Parametrisierungen und der Kettenregel (4.5), so ergibt sich

$$\int_{\gamma_1} v \, ds = \int_0^1 (t^2 + t^2) \, dt = \frac{2}{3} \quad \text{und} \quad \int_{\bar{\gamma}_1} v \, ds = \int_0^1 -2(1-t)^2 \, dt = -\frac{2}{3}$$

Vergleichen wir nun noch das Integral des Vektorfeldes  $v$  entlang  $\gamma_1$  mit dem Integral des Vektorfeldes entlang  $\gamma_2 : [0, 1] \rightarrow \mathbb{R}^2$ , wobei

$$\gamma_2 := \begin{cases} (t, 0) & \text{für } t \in [0, 1), \\ (1, t-1) & \text{für } t \in [1, 2] \end{cases} \quad (4.6)$$

Dann gilt

$$\int_{\gamma_2} v \, ds = \int_0^1 t^2 \, dt + \int_1^2 t-1 \, dt = \frac{1}{3} + \left[ \frac{t^2}{2} - t \right]_1^2 = \frac{5}{6}.$$

Das Wegintegral von  $v$  von  $(0, 0)$  nach  $(1, 1)$  ist also nicht unabhängig vom gewählten Weg.

Bemerkung: Im Allgemeinen gilt für einen stückweise glatten Weg  $\alpha = \alpha_1 \cup \dots \cup \alpha_n$

$$\int_{\alpha} v \, ds = \sum_{k=1}^n \int_{\alpha_k} v \, ds. \quad (4.7)$$

**Definition 4.11:** Ein Vektorfeld  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  heißt *konservativ (oder exakt)*, falls das Integral von  $v$  über einen stückweisen glatten Weg  $\gamma$  nur von den Endpunkten und nicht vom Weg selbst abhängt.

Für  $x_0, x_1 \in \Omega$  und  $\gamma_1$  ein Weg von  $x_0$  nach  $x_1$  und  $\gamma_2$  ein Weg von  $x_1$  nach  $x_0$  gilt also

$$\int_{\gamma_1} v \, ds = - \int_{\gamma_2} v \, ds \implies \int_{\gamma_1} v \, ds + \int_{\gamma_2} v \, ds = 0. \quad (4.8)$$

**Lemma 4.12:** Sei  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  ein Vektorfeld. Dann ist  $v$  genau dann konservativ, wenn für jede geschlossene stückweise glatte Kurve  $\gamma \subset \Omega$  gilt

$$\int_{\gamma} v \, ds = 0. \quad (4.9)$$

**Beispiele** Das Gravitationsfeld  $K(x) := -\frac{x}{\|x\|^3}$  ist konservativ.

Als Gegenbeispiel betrachte einen in der  $z$ -Ebene liegenden Draht, durch den ein konstanter Strom fließt. Dieser erzeugt ein Magnetfeld  $H : \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3, z = 0\} \rightarrow \mathbb{R}^3$ , das durch

$$H(x, y, z) := \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}. \quad (4.10)$$

definiert ist. Dieses Magnetfeld ist nicht konservativ. Betrachten wir beispielsweise den geschlossenen Weg  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3$  mit  $\gamma(t) := (\sin t, \cos t, 0)^T$ , so ist

$$\begin{aligned} \int_{\gamma} H \, ds &= \int_0^{2\pi} \langle H(\gamma(t)), \dot{\gamma}(t) \rangle \, dt \\ &= \int_0^{2\pi} \frac{1}{\sin^2 t + \cos^2 t} \left\langle \begin{pmatrix} -\cos t \\ \sin t \\ 0 \end{pmatrix}, \begin{pmatrix} \cos t \\ -\sin t \\ 0 \end{pmatrix} \right\rangle \, dt \\ &= \int_0^{2\pi} -1 \, dt = -2\pi \neq 0. \end{aligned}$$

**Satz 4.13:** Ein Vektorfeld  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  ist genau dann konservativ, wenn es eine  $\mathcal{C}^1$ -Funktion  $u : \Omega \rightarrow \mathbb{R}$  gibt, sodass

$$v = \nabla u. \quad (4.11)$$

Für jede stückweise glatte Kurve  $\gamma \subset \Omega$  von  $x_0$  nach  $x_1$  gilt also

$$\int_{\gamma} v \, ds = u(x_1) - u(x_0). \quad (4.12)$$

*Beweisidee.* Sei  $u$  gegeben, sodass  $v = \nabla u$ . Dann folgt für eine  $\mathcal{C}^1$ -Kurve  $\gamma : [a, b] \rightarrow \Omega$  nach Kettenregel

$$\frac{d}{dt} u(\gamma(t)) = \langle \nabla u(\gamma(t)), \dot{\gamma}(t) \rangle = \langle v(\gamma(t)), \dot{\gamma}(t) \rangle.$$

Integriert man beide Seiten, so folgt die zu zeigende Aussage aus dem Hauptsatz der Differential- und Integralrechnung. Die Rückrichtung lässt sich in der angegebenen Literatur finden.  $\square$

**Bemerkung 4.14:** Existiert für ein Vektorfeld  $v$  eine Funktion  $u$  mit  $v = \nabla u$ , so heißt  $u$  auch Stammfunktion oder *Potential des Vektorfeldes*  $v$ . Die Funktion  $u$  ist bis auf eine additive Konstante eindeutig. In der Physik wird das Potential häufig auch mit negativem Vorzeichen definiert, also  $v = -\nabla u$ .

### 4.1.3 Potentiale von Vektorfeldern

Es stellt sich die Frage, unter welchen Voraussetzungen Vektorfelder Potentiale besitzen. Ein notwendiges Kriterium für die Existenz eines Potentials sind beispielsweise folgende Integrabilitätsbedingungen.

**Lemma 4.15:** Besitzt das  $\mathcal{C}^1$ -Vektorfeld  $v : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  auf einem Gebiet  $\Omega$  ein Potential  $u$ , so muss für alle  $i, k = 1, \dots, n, i \neq k$ ,

$$\frac{\partial v_i}{\partial x_k} = \frac{\partial v_k}{\partial x_i} \quad (4.13)$$

gelten.

*Beweis.* Unter den gegebenen Voraussetzungen ist  $u$  zweimal stetig differenzierbar. Nach dem Satz von Schwarz (2.10) gilt dann

$$\frac{\partial v_i}{\partial x_k} = \frac{\partial}{\partial x_k} \frac{\partial u}{\partial x_i} = \frac{\partial}{\partial x_i} \frac{\partial u}{\partial x_k} = \frac{\partial v_k}{\partial x_i}. \quad \square$$

**Bemerkung 4.16:** Die Bedingung (4.13) ist nicht hinreichend. Betrachte beispielsweise das in (4.10) definierte Magnetfeld  $H : \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3, z = 0\} \rightarrow \mathbb{R}^3$ . Wie wir bereits gesehen haben, ist es nicht konservativ, besitzt also keine Stammfunktion. Es ist jedoch differenzierbar, und es gilt

$$\frac{\partial H_1}{\partial y} = \frac{\partial}{\partial y} \left( -\frac{y}{x^2 + y^2} \right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x} \left( \frac{x}{x^2 + y^2} \right) = \frac{\partial H_2}{\partial x}$$

sowie

$$\frac{\partial H_1}{\partial z} = \frac{\partial H_3}{\partial x} = 0 = \frac{\partial H_3}{\partial y} = \frac{\partial H_2}{\partial z}.$$

Auch physikalisch ist das durchaus sinnvoll: Es lässt sich als Übungsaufgabe leicht nachweisen, dass das Integral über eine den Draht umlaufende Kreislinie nicht verschwindet. Das Magnetfeld kann daher kein Potential besitzen.

Vektorfelder mit der Eigenschaft (4.13) heißen auch *rotationsfrei*.

Um auch hinreichende Bedingungen für die Existenz eines Potentials für ein Vektorfeld angeben zu können, benötigen wir strengere Vorgaben an das zugrundeliegende Gebiet  $\Omega$ .

**Definition 4.17:** Ein Gebiet  $\Omega \subset \mathbb{R}^n$  heißt *sternförmig*, wenn es einen Punkt  $x_0 \in \Omega$  gibt, sodass mit jedem  $x \in \Omega$  auch die Strecke  $\lambda x_0 + (1 - \lambda)x \in \Omega$  für alle  $\lambda \in (0, 1)$ .

**Definition 4.18:** Ein Gebiet  $\Omega \subset \mathbb{R}^n$  heißt *einfach*, wenn es das Bild eines sternförmigen Gebiets unter einem  $\mathcal{C}^2$ -Diffeomorphismus ist. Es muss also einen  $\mathcal{C}^2$ -Diffeomorphismus von einem sternförmigen Gebiet nach  $\Omega$  geben.

### Beispiele

1. Für jeden Punkt  $x_0 \in \mathbb{R}^n$  sei  $K_\delta(x_0)$  die Kugel um  $x_0$  mit Radius  $\delta$ . Diese Kugel ist sternförmig.
2. Jeder Quader  $Q = Q_1 \times Q_2 \times \cdots \times Q_n \subset \mathbb{R}^n$  ist sternförmig.
3. Ebenso sind herzförmige Gebiete sternförmig.
4. Die Menge  $K^* := \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 1\} \setminus \{(x, 0), x > 0\}$  ist sternförmig.
5. Die Menge  $\Omega := \{(x, y) \in \mathbb{R}^2, r < \sqrt{x^2 + y^2} < R\} \setminus \{(x, 0), x \geq 0\}$  ist einfach.
6. Die Menge  $\mathbb{R}^2 \setminus \{(t \cos t, t \sin t) \in \mathbb{R}^2, t \in \mathbb{R}\}$ , also  $\mathbb{R}^2$  ohne die archimedische Spirale, ist ebenfalls einfach.

**Bemerkung 4.19:** In einem einfachen Gebiet lässt sich jede geschlossene Kurve stetig auf einen Punkt zusammenziehen, ohne dass das Gebiet verlassen werden muss

**Satz 4.20:** Sei  $\Omega \subset \mathbb{R}^n$  ein einfaches Gebiet. Dann besitzt das Vektorfeld  $v : \Omega \rightarrow \mathbb{R}^n$  ein Potential, wenn es die Integrabilitätsbedingungen (4.13) erfüllt.

Bemerkung: Im Fall des oben betrachteten Magnetfelds  $H$  (vgl. (4.10)) ist die Definitionsmenge  $\{(x, y, z) \in \mathbb{R}^3, z \neq 0\}$  nicht einfach.

#### 4.1.3.1 Berechnung des Potentials in $\mathbb{R}^2$

Die nachfolgenden Rechenschritte zur Bestimmung eines Potentials werden zunächst formal durchgeführt. Es wird also nicht in jedem Schritt geprüft, ob der Schritt tatsächlich durchführbar ist. Am Ende dieses Formalismus steht ein Kandidat, und es muss nachgeprüft werden, dass der Kandidat tatsächlich ein Potential des Vektorfelds ist.

Sei  $\Omega$  einfach und  $v : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$  ein  $\mathcal{C}^1$ -Vektorfeld mit

$$v(x, y) = \begin{pmatrix} p(x, y) \\ q(x, y) \end{pmatrix}.$$

Wir nehmen an, dass  $\frac{\partial p}{\partial y}(x, y) = \frac{\partial q}{\partial x}(x, y)$ , sodass dann nach Satz 4.20 ein Potential  $u$  existiert. Zur Bestimmung eines Potentials halten wir zunächst  $y$  fest und integrieren die Funktion  $x \mapsto p(x, y)$ . Das liefert eine Stammfunktion  $P$  mit  $\partial_x P = p$ . Wir machen den Ansatz

$$u(x, y) := P(x, y) + \tilde{Q}(x, y),$$

wobei  $\tilde{Q}$  noch zu bestimmen ist. Leiten wir nun  $u$  nach  $y$  ab, so muss

$$\partial_y u(x, y) = \partial_y P(x, y) + \partial_y \tilde{Q}(x, y) \stackrel{!}{=} q(x, y)$$

gelten nach Definition des Potentials. Es ist also

$$\partial_y \tilde{Q}(x, y) = q(x, y) - \partial_y P(x, y).$$

Da die rechte Seite aufgrund der Integrabilitätsbedingung von  $x$  unabhängig ist, muss auch die linke Seite von  $x$  unabhängig sein.<sup>1</sup> Damit ergibt sich also

$$u(x, y) = P(x, y) + Q(y).$$

Nun lässt sich eine Stammfunktion für

$$Q'(y) = \partial_y P(x, y) - q(x, y)$$

bestimmen.

**Beispiel** Sei  $\Omega \subset \mathbb{R}^2$  ein einfaches Gebiet und  $v : \Omega \rightarrow \mathbb{R}^2$  durch

$$v(x, y) := \begin{pmatrix} 2x \cos(y) \\ -x^2 \sin(y) \end{pmatrix}$$

gegeben. Offensichtlich sind die Integrabilitätsbedingungen erfüllt. Es existiert also ein Potential. Zur Berechnung halten wir zunächst  $y$  fest und bestimmen als Stammfunktion für  $x \mapsto 2x \cos(y)$

$$P(x, y) = x^2 \cos(y).$$

Wir setzen damit für das Potential  $u(x, y) := x^2 \cos(y) + Q(y)$  an, sodass

$$Q(y) := u(x, y) - x^2 \cos(y).$$

Als Ableitung von  $Q$  ergibt sich

$$Q'(y) = \partial_y u(x, y) - \partial_y P(x, y) = -x^2 \sin(y) + x^2 \sin(y) = 0.$$

Damit ist also  $u(x, y) := x^2 \cos(y) + a$  mit  $a \in \mathbb{R}$ . In der Tat ist damit  $\nabla u = v$ .

#### 4.1.3.2 Berechnung des Potentials in $\mathbb{R}^3$

Die Vorgehensweise in höheren Dimensionen ist im Prinzip ganz analog. Zunächst schreiben wir das  $\mathcal{C}^1$ -Vektorfeld  $v : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$  als

$$v(x, y, z) = \begin{pmatrix} p(x, y, z) \\ q(x, y, z) \\ r(x, y, z) \end{pmatrix},$$

---

<sup>1</sup>Um das einzusehen, leite man die Gleichung nach  $x$  ab.

wobei wieder die Voraussetzungen von Satz 4.20 erfüllt seien. Wir halten  $(y, z)$  fest und bestimmen eine Stammfunktion der Abbildung  $x \mapsto p(x, y, z)$ . Diese bezeichnen wir mit  $P(x, y, z)$ . Erneut setzen wir das Potential als

$$u(x, y, z) := P(x, y, z) + Q(y, z)$$

an, wobei wie zuvor aus den Integrierbarkeitsbedingungen folgt, dass der zusätzliche Term  $Q$  tatsächlich unabhängig von  $x$  ist. Außerdem muss aus dem gleichen Grund

$$\partial_y Q(y, z) = q(x, y, z) - \partial_y P(x, y, z) \quad \text{sowie} \quad \partial_z Q(y, z) = r(x, y, z) - \partial_z P(x, y, z).$$

Die Bestimmung von  $Q$  erfolgt nun ganz genauso wie im zweidimensionalen Fall.

**Beispiel** Gegeben sei das Vektorfeld  $v : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$v(x, y, z) := \begin{pmatrix} p(x, y, z) \\ q(x, y, z) \\ r(x, y, z) \end{pmatrix} := \begin{pmatrix} 2x \cos(y) - 2z^3 \\ 3 + 2y \exp(z) - x^2 \sin(y) \\ y^2 \exp(z) - 6xz^2 \end{pmatrix}.$$

Die Integrierbarkeitsbedingungen sind erfüllt, da

$$\begin{aligned} \partial_y p(x, y, z) &= -2x \sin(y) = \partial_x q(x, y, z), \\ \partial_z p(x, y, z) &= -6z^2 = \partial_x r(x, y, z) \end{aligned}$$

und

$$\partial_z q(x, y, z) = 2y \exp(z) = \partial_y r(x, y, z).$$

Zur Bestimmung des Potentials halten wir zunächst  $y$  und  $z$  fest und integrieren  $p$  nach  $x$ . Das führt auf  $P(x, y, z) = x^2 \cos(y) - 2xz^3$ . Wir setzen daher das Potential mit

$$u(x, y, z) = x^2 \cos(y) - 2xz^3 + Q(y, z)$$

an. Nun ist

$$\begin{aligned} \partial_y Q(y, z) &= q(x, y, z) - \partial_y P(x, y, z) \\ &= 3 + 2y \exp(z) - x^2 \sin(y) + x^2 \sin(y) = 3 + 2y \exp(z). \end{aligned}$$

Wir halten nun  $z$  fest, integrieren  $\partial_y Q(y, z)$  nach  $y$  und erhalten

$$Q(y, z) = 3y + y^2 \exp(z) + R(z),$$

wobei, wie wie zuvor gezeigt werden kann, dass  $R$  tatsächlich von  $y$  unabhängig ist. Leiten wir nun

$$u(x, y, z) = \underbrace{x^2 \cos(y) - 2xz^3}_{=P} + \underbrace{3y + y^2 \exp(z) + R(z)}_{=Q}$$

nach  $z$  ab, so folgt

$$\partial_z u(x, y, z) = -6xz^2 + y^2 \exp(z) + R'(z) \stackrel{!}{=} r(x, y, z),$$

weswegen

$$R'(z) = y^2 \exp(z) - 6xz^2 + 6xz^2 - y^2 \exp(z) = 0$$

und damit  $R$  eine Konstante ist. Da das Potential ohnehin nur bis auf eine (additive) Konstante eindeutig ist, ergibt sich für das Potential  $u$  des Vektorfeldes  $v$  die Beziehung

$$u(x, y, z) = x^2 \cos(y) - 2xz^3 + 3y + y^2 \exp(z).$$

Die Differenzierung dieses Potentials zeigt, dass in der Tat  $\nabla u = v$  gilt.

## 4.2 Oberflächenintegrale

**Definition 4.21:** Sei  $U \subset \mathbb{R}^2$  ein Gebiet und  $\Phi : U \rightarrow \mathbb{R}^3$  eine injektive Abbildung mit

$$\Phi(u_1, u_2) := \begin{pmatrix} \Phi_1(u_1, u_2) \\ \Phi_2(u_1, u_2) \\ \Phi_3(u_1, u_2) \end{pmatrix}. \quad (4.14)$$

Dann heißt  $\Phi$  eine  $\mathcal{C}^k$ -Flächenparametrisierung, wenn die partiellen Ableitungen

$$\frac{\partial}{\partial u_1} \Phi(u_1, u_2) \quad \text{und} \quad \frac{\partial}{\partial u_2} \Phi(u_1, u_2) \quad (4.15)$$

für jedes  $u = (u_1, u_2) \in U$  linear unabhängig sind.

**Definition 4.22:** Für ein Gebiet  $U$ , eine Flächenparametrisierung  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , festes  $u \in U$  und  $\varepsilon > 0$  sind die Abbildungen  $\gamma_1, \gamma_2 : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^3$  mit

$$\gamma_1(t) := \Phi(u_1 + t, u_2) \quad \text{sowie} \quad \gamma_2(t) := \Phi(u_1, u_2 + t) \quad (4.16)$$

$\mathcal{C}^1$ -Kurven in  $\Phi(U) \subset \mathbb{R}^3$ . Für die Ableitungen dieser Kurven gilt

$$\frac{d}{dt} \gamma_1(0) = \frac{d}{dt} \Phi(u_1 + t, u_2) \Big|_{t=0} = \frac{\partial}{\partial u_1} \Phi(u_1, u_2) \quad (4.17)$$

und

$$\frac{d}{dt} \gamma_2(0) = \frac{d}{dt} \Phi(u_1, u_2 + t) \Big|_{t=0} = \frac{\partial}{\partial u_2} \Phi(u_1, u_2) \quad (4.18)$$

Deswegen nennen wir  $\frac{\partial}{\partial u_1} \Phi$  und  $\frac{\partial}{\partial u_2} \Phi$  die Tangentenvektoren von  $\Phi$  an  $u$ .

**Bemerkung 4.23:** Aus der linearen Unabhängigkeit der partiellen Ableitungen folgt, dass das Bild  $\Phi(U)$  tatsächlich zweidimensional ist

**Definition 4.24:** Sei  $U$  ein Gebiet. Eine Menge  $M \subset \mathbb{R}^3$  heißt  $\mathcal{C}^k$ -Flächenstück, wenn sie Bild einer  $\mathcal{C}^k$ -Flächenparametrisierung  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  ist, deren Umkehrabbildung  $\Phi^{-1} : M \rightarrow U$  stetig ist. Jede solche Abbildung  $\Phi$  heißt *Parametrisierung des Flächenstücks*  $M$ .

Bemerke, dass ein Flächenstück per Definition offen sein muss.

**Beispiele**

1. Für drei linear unabhängige Vektoren  $a, a_1, a_2 \in \mathbb{R}^3$  lässt sich die von diesen Vektoren aufgespannte Ebene im dreidimensionalen Raum über  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\Phi(\lambda, \mu) := a + \lambda a_1 + \mu a_2 \tag{4.19}$$

parametrisieren.

2. Die obere beziehungsweise untere Hemisphäre der dreidimensionalen Einheitskugel lässt sich über  $\Phi : K_1(0) \rightarrow \mathbb{R}^3$  mit

$$\Phi(u_1, u_2) := \begin{pmatrix} u_1 \\ u_2 \\ \pm \sqrt{1 - \|u\|_2^2} \end{pmatrix} \tag{4.20}$$

parametrisieren.

3. Die Einheitskugel ohne Nullmeridian, also die Kugeloberfläche des Einheitskreises ohne den von  $(0, 0, 1)$  über  $(1, 0, 0)$  nach  $(0, 0, -1)$  verlaufenden Meridian, lässt sich durch  $\Phi : (0, 2\pi) \times (0, \pi) \rightarrow \mathbb{R}^3$  mit

$$\Phi(\varphi, \theta) := \begin{pmatrix} \cos(\varphi) \sin(\theta) \\ \sin(\varphi) \sin(\theta) \\ \cos(\theta) \end{pmatrix} \tag{4.21}$$

parametrisieren.

4. Der geschlitzte Mantel eines Kreiskegels lässt sich durch  $\Phi : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^3$  mit

$$\Phi(z, \varphi) := \begin{pmatrix} z \cos \varphi \\ z \sin \varphi \\ z \end{pmatrix} \tag{4.22}$$

parametrisieren.

Alle diese Beispiele sind Flächenstücke. Ebenso ist der Graph einer  $\mathcal{C}^k$ -Funktion  $f : U \rightarrow \mathbb{R}$  ein Flächenstück.

**Lemma 4.25:** Seien  $M \subset \mathbb{R}^3$  ein  $\mathcal{C}^k$ -Flächenstück und seien  $\Phi : U \rightarrow \mathbb{R}^3$  und  $\Psi : V \rightarrow \mathbb{R}^3$  zwei Parametrisierungen von  $M$ . Dann existiert ein  $\mathcal{C}^k$ -Diffeomorphismus  $h : U \rightarrow V$ , sodass  $\Phi = \Psi \circ h$ . Jedes solche  $h$  heißt *Parametrisierungstransformation* zwischen  $\Phi$  und  $\Psi$ .

**Beispiel** Sei  $S := \{x \in \mathbb{R}^3, \|x\| = 1\}$  die Einheitskugel und  $N = (0, 0, 1)$  der ausgezeichnete Nordpol. Die Stereographische Projektion ist die Abbildung  $\Phi : \mathbb{R}^2 \rightarrow S \setminus \{N\}$  mit

$$\Phi(u_1, u_2) := \frac{1}{\|u\|^2 + 1} \begin{pmatrix} 2u_1 \\ 2u_2 \\ \|u\|^2 - 1 \end{pmatrix}. \tag{4.23}$$

Sie besitzt die Umkehrabbildung  $\Phi^{-1} : S \setminus \{N\} \rightarrow \mathbb{R}^2$  mit

$$\Phi^{-1}(x, y, z) := \left( \frac{x}{1-z}, \frac{y}{1-z} \right). \tag{4.24}$$



### 4.2.1 Inhalte von Flächenstücken

Sei  $M \subset \mathbb{R}^3$  ein Flächenstück, das von  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  parametrisiert wird. Um eine Formel für den Flächeninhalt von  $M$  herzuleiten, fixieren wir einen Punkt  $u \in M$  und betrachten das von den Tangentialvektoren an  $M$  in  $u$  aufgespannte Parallelogramm. Die Fläche  $M$  kann lokal also durch dieses Parallelogramm approximiert werden. Im Grunde ist das nichts Anderes, als eine Taylorapproximation erster Ordnung an  $u$ , denn es gilt

$$\Phi(u+h) = \Phi(u) + h_1 \frac{\partial}{\partial u_1} \Phi(u) + h_2 \frac{\partial}{\partial u_2} \Phi(u)$$

mit  $h = (h_1, h_2)$  und  $h_1, h_2 \in (0, \varepsilon)$ .

Die Idee ist nun, den Flächeninhalt von  $M$  über die Flächeninhalte der Parallelogramme an allen  $u \in M$  anzunähern. Dazu benötigen wir zunächst eine Vorschrift, wie wir mithilfe der Parametrisierung  $\Phi$  den Flächeninhalt eines der Parallelogramme berechnen können. Wie haben bereits gesehen, dass die Tangentialebene von  $\Phi$  an einer Stelle  $u \in U$  von den Vektoren  $\partial_{u_1} \Phi(u)$  und  $\partial_{u_2} \Phi(u)$  aufgespannt wird. Außerdem erinnern wir uns, dass für zwei Vektoren  $a, b \in \mathbb{R}^3$  das Kreuzprodukt als

$$a \times b := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \quad (4.25)$$

definiert ist. Bezeichnen wir mit  $\varphi$  den Winkel zwischen  $a$  und  $b$ , so gilt zum einen

$$\|a \times b\| = \|a\|_2 \|b\|_2 \cdot \sin \varphi \quad (4.26)$$

und bekanntermaßen

$$|\langle a, b \rangle| = \|a\|_2 \|b\|_2 \cdot \cos \varphi. \quad (4.27)$$

Es lässt sich anschaulich leicht einsehen, dass die Formel (4.26) genau den Flächeninhalt des von  $a$  und  $b$  aufgespannten Parallelogramms angibt. Wir wenden diese Formel nun auf die Tangentialvektoren  $\partial_{u_1} \Phi(u)$  und  $\partial_{u_2} \Phi(u)$  an. Dafür setzen wir zunächst

$$G := \begin{pmatrix} \langle \partial_{u_1} \Phi, \partial_{u_1} \Phi \rangle & \langle \partial_{u_1} \Phi, \partial_{u_2} \Phi \rangle \\ \langle \partial_{u_2} \Phi, \partial_{u_1} \Phi \rangle & \langle \partial_{u_2} \Phi, \partial_{u_2} \Phi \rangle \end{pmatrix}, \quad (4.28)$$

sodass  $G = (g_{ij})$  mit  $g_{ij} = \langle \partial_i \Phi, \partial_j \Phi \rangle$ ,  $i, j \in \{1, 2\}$ . Die Gram-Matrix  $G$  ist abhängig von der gewählten Parametrisierung und abhängig vom konkreten Punkt  $u$ , und wir sehen, dass

$$\begin{aligned} \|\partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u)\|^2 &= \|\partial_{u_1} \Phi(u)\|^2 \|\partial_{u_2} \Phi(u)\|^2 \sin^2(\theta) \\ &= \|\partial_{u_1} \Phi(u)\|^2 \|\partial_{u_2} \Phi(u)\|^2 - |\langle \partial_{u_1} \Phi(u), \partial_{u_2} \Phi(u) \rangle|^2 = \det(G(u)), \end{aligned}$$

wobei  $\theta$  der Winkel zwischen den Tangentialvektoren beschreibt. Es gilt also die Beziehung

$$\sqrt{\det(G(u))} = \|\partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u)\|, \quad (4.29)$$

mit deren Hilfe wir den Flächeninhalt des von den Tangentialvektoren in  $u$  aufgespannten Parallelogramms berechnen können. Das motiviert die folgende Definition.

**Definition 4.26:** Sei  $M \subset \mathbb{R}^3$  ein Flächenstück und  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  eine Parametrisierung von  $M$ . Dann definieren wir den *Flächeninhalt von  $M$*  als

$$A(M) := \begin{cases} A(\Phi) := \int_U \sqrt{\det(G(u))} \, d^2u, & \text{falls dieser Wert existiert} \\ \infty & \text{sonst.} \end{cases} \quad (4.30)$$

Der Flächeninhalt von  $M$  entspricht also dem Grenzwert der Summe der Flächeninhalte der obigen Parallelelogramme.

Bemerkung: Die Gram-Matrix  $G$  ist symmetrisch und positiv definit. Da die partiellen Ableitungen von  $\Phi$  nach Definition des Flächenstücks linear unabhängig sind, ist die Determinante von  $G$  stets positiv.

**Lemma 4.27:** Die Definition 4.26 ist unabhängig von der gewählten Parametrisierung.

*Beweis.* Sei  $\Psi : V \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  eine weitere Parametrisierung von  $M$ . Dann existiert nach Lemma 4.25 ein Diffeomorphismus  $h : U \rightarrow V$  mit  $\Phi = \Psi \circ h$ . Aus der Kettenregel folgt dann

$$\begin{aligned} \partial_{u_1}\Phi(u) \times \partial_{u_2}\Phi(u) &= D\Psi(h(u)) \cdot \partial_1 h(u) \times D\Psi(h(u)) \cdot \partial_2 h(u) \\ &= \left( (\partial_{v_1}\Psi) \circ h(u) \right) \times \left( (\partial_{v_2}\Psi) \circ h(u) \right) \cdot \left( \partial_{u_1} h_1 \partial_{u_2} h_2 - \partial_{u_2} h_1 \partial_{u_1} h_2 \right)(u) \end{aligned}$$

nach kurzer aber intensiver Rechnung. Bezeichnen wir mit  $\bar{G} := (\bar{g}_{ij})$  mit  $\bar{g}_{ij} = \langle \partial_{v_1}\Psi, \partial_{v_2}\Psi \rangle$  für  $i, j \in \{1, 2\}$ , so ist obige Formel äquivalent zu

$$\sqrt{\det(G(u))} = \|\partial_{u_1}\Phi(u) \times \partial_{u_2}\Phi(u)\| = \|\det(Dh(u))\| \cdot \sqrt{\det(\bar{G} \circ h(u))}.$$

Aus dem Transformationssatz 3.36 folgt schließlich

$$\int_U \sqrt{\det(G(u))} \, d^2u = \int_U \sqrt{\det(\bar{G} \circ h(u))} \cdot \|\det Dh(u)\| \, d^2u = \int_V \sqrt{\det(\bar{G}(v))} \, d^2v. \quad \square$$

**Beispiel** Sei  $U = (0, 1) \times (0, 2\pi) \subset \mathbb{R}^2$  und  $\Phi : U \rightarrow \mathbb{R}^3$  gegeben durch

$$\Phi(r, \theta) := \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \\ r \end{pmatrix}$$

(vgl. mit der Parametrisierung eines geschlitzten Kreiskegels oben). Die partiellen Ableitungen von  $\Phi$  sind durch

$$\partial_r \Phi(r, \theta) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 1 \end{pmatrix} \quad \text{und} \quad \partial_\theta \Phi(r, \theta) = \begin{pmatrix} -r \sin(\theta) \\ r \cos(\theta) \\ 0 \end{pmatrix}$$

gegeben. Als Determinante der Matrix  $G$  bestimmen wir

$$\det G = \det \begin{pmatrix} \|\partial_r \Phi(r, \theta)\|^2 & \langle \partial_r \Phi(r, \theta), \partial_\theta \Phi(r, \theta) \rangle \\ \langle \partial_\theta \Phi(r, \theta), \partial_r \Phi(r, \theta) \rangle & \|\partial_\theta \Phi(r, \theta)\|^2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & r^2 \end{pmatrix} = 2r^2.$$

Somit ist  $\sqrt{\det(G(r, \theta))} = \|\partial_r \Phi(r, \theta) \times \partial_\theta \Phi(r, \theta)\| = r\sqrt{2}$ . Damit können wir den Flächeninhalt von  $\Phi(U)$  berechnen als

$$\int_U \sqrt{\det(G(r, \theta))} \, d^2(r, \theta) = \int_0^{2\pi} \int_0^1 r\sqrt{2} \, dr \, d\theta = \pi\sqrt{2}.$$

Zur Wiederholung mache man sich klar, warum die Anwendung des Satzes von Fubini hier gerechtfertigt ist.

**Beispiel** Sei  $S_r := \{(x, y, z) \in \mathbb{R}^3, \|(x, y, z)\|_2 = r\}$  die Kugel mit Radius  $r$  um den Ursprung. Diese Menge ist kein Flächenstück. Setzen wir  $N := \{(\cos \theta, 0, \sin \theta), 0 < \theta \leq \pi\}$  und definieren  $S_r^* := S_r \setminus N$ , so ist  $S_r^*$  ein Flächenstück. Da sich der Flächeninhalt durch das Entfernen von Nullmengen (endlich vielen Punkten, stückweise glatten Kurven, etc.) nicht ändert, ist

$$A(S_r^*) = A(S_r).$$

Da sich die Menge  $S_r^*$  durch Kugelkoordinaten parametrisieren lässt, können wir den Oberflächeninhalt von  $S_r$  berechnen. Sei dazu  $U := (0, 2\pi) \times (0, \pi)$  und  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

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

Die Ableitungen von  $\Phi$  sind gegeben durch

$$\partial_\varphi \Phi(\varphi, \theta) = \begin{pmatrix} -r \sin(\varphi) \sin(\theta) \\ r \cos(\varphi) \sin(\theta) \\ 0 \end{pmatrix} \quad \text{und} \quad \partial_\theta \Phi(\varphi, \theta) = \begin{pmatrix} r \cos(\varphi) \cos(\theta) \\ r \sin(\varphi) \cos(\theta) \\ -r \sin(\theta) \end{pmatrix}.$$

Wir zuvor berechnen wir die Determinante der Matrix  $G$  und finden  $\det(G(\varphi, \theta)) = r^4 \sin^2(\theta)$ . Somit berechnen wir den Oberflächeninhalt als

$$A(S_r) = A(S_r^*) = \int_U \sqrt{\det(G(\varphi, \theta))} \, d^2(\varphi, \theta) = \int_0^\pi \int_0^{2\pi} r^2 \sin(\theta) \, d\varphi \, d\theta = 4\pi r^2.$$

### 4.2.2 Skalare Oberflächenintegrale

Nachdem wir nun den Volumenbegriff auf parametrisierbare Flächen erweitert haben, können wir skalarwertige Funktionen über solchen Flächen betrachten.

**Definition 4.28:** Sei  $M \subset \mathbb{R}^3$  ein Flächenstück und  $f : M \rightarrow \mathbb{R}$  eine stetige Funktion. Die Fläche  $M$  sei wie zuvor durch  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  parametrisiert und wir bezeichnen mit  $G$  die Gram-Matrix aus (4.28). Dann ist das *skalare Oberflächenintegral von  $f$  über  $M$*  durch

$$\int_M f \, ds := \int_U f(\Phi(u)) \sqrt{\det(G(u))} \, d^2u. \quad (4.31)$$

definiert. Bemerke, dass das sogenannte *skalare Oberflächenelement*  $ds = \sqrt{\det(G(u))} \, d^2u$  nach Lemma 4.27 unabhängig von der gewählten Parametrisierung ist.

**Beispiel** Sei  $U := (0, 1) \times (0, 2\pi) \subset \mathbb{R}^2$  und betrachten wir die durch  $\Phi : U \rightarrow \mathbb{R}^3$  mit

$$\Phi(r, \theta) := \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ \theta \end{pmatrix}$$

parametrisierte Wendelfläche (Helikloid)  $S := \Phi(U)$ . Angenommen,  $S$  besitze die Massendichte  $m : \mathbb{R}^3 \rightarrow \mathbb{R}$  mit  $m(x, y, z) := \sqrt{x^2 + y^2}r$ . Dann lässt sich die gesamte Massendichte von  $S$  bestimmen durch

$$M(S) = \int_S 2\sqrt{x^2 + y^2} \, ds = \int_U 2r \sqrt{\det(G(r, \theta))} \, d^2(r, \theta).$$

Wir bestimmen daher die Determinante der Gram-Matrix  $G$  wie zuvor, indem wir  $\Phi$  partiell ableiten, und wir finden

$$\det(G(u)) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 + r^2 \end{pmatrix} = 1 + r^2.$$

Damit ergibt sich

$$M(S) = \int_0^{2\pi} \int_0^1 2r \sqrt{1 + r^2} \, dr \, d\theta = \frac{2}{3} \int_0^{2\pi} \left[ (1 + r^2)^{3/2} \right]_0^1 \, d\theta = \frac{4\pi}{3} (\sqrt{8} - 1).$$

**Definition 4.29:** Sei  $M \subset \mathbb{R}^3$  ein Flächenstück. Es seien zwei Parametrisierungen  $\Phi$  und  $\Psi$  von  $M$  gegeben und  $h$  die Parametrisierungstransformation, sodass  $\Phi = \Psi \circ h$ . Dann heißen  $\Phi$  und  $\Psi$  *gleich orientiert*, wenn

$$\det(Dh) > 0 \quad (4.32)$$

gilt. Wir können  $M$  *orientieren* und erhalten damit ein *orientiertes Flächenstück*, indem wir eine Parametrisierung auszeichnen. Alternativ kann man auch die Auszeichnung eines Einheitsnormalenfeldes  $n : M \rightarrow \mathbb{R}^3$  auf  $M$  angeben, wobei der Zusammenhang zwischen Normalenfeld und Parametrisierung  $\Phi$  für  $x := \Phi(u) \in M$  durch

$$n(x) = \pm \frac{\partial_{u_1} \Phi \times \partial_{u_2} \Phi}{\|\partial_{u_1} \Phi \times \partial_{u_2} \Phi\|}(u). \quad (4.33)$$

gegeben ist. Hieran lässt sich auch der Richtungssinn der Parametrisierung entsprechend der durch das Normalenfeld gegebenen Orientierung ablesen. Das Einheitsnormalenfeld ist ein normiertes Vektorfeld, das orthogonal auf der Tangentialebene an  $M$  in  $x$  steht.

### 4.2.3 Vektorielle Oberflächenintegrale

Betrachten wir zum Ende des Abschnitts über Oberflächenintegrale noch Vektorfelder über orientierten Flächenstücken.

**Definition 4.30:** Sei  $M \subset \mathbb{R}^3$  ein orientiertes Flächenstück und  $\Phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  eine Parametrisierung von  $M$  mit positivem (negativem) Richtungssinn. Für ein stetiges Vektorfeld  $v : M \rightarrow \mathbb{R}^3$  ist das *vektorielle Oberflächenintegral von  $v$  über  $M$*  durch

$$\int_M v \, \text{do} := \int_M v(x) \cdot \text{do}(x) := \pm \int_U \langle v \circ \Phi(u), \partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u) \rangle \, \text{d}^2u \quad (4.34)$$

definiert, sofern der Wert auf der rechten Seite existiert. Das Integral hat hierbei positives (negatives) Vorzeichen, wenn  $\Phi$  positiven (negativen) Richtungssinn besitzt.

**Bemerkung 4.31:** Mit  $\text{do}(x)$  bezeichnen wir das *vektorielle Oberflächenelement*. Die Schreibweise  $v(x) \cdot \text{do}(x)$  soll hierbei verdeutlichen, dass  $\text{do}(x)$  vektoriellen Charakter besitzt.

Formal gilt

$$\text{do}(x) = \pm \partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u) \, \text{d}^2u = \pm \frac{\partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u)}{\|\partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u)\|} \sqrt{\det(G(u))} \, \text{d}^2u = n \, \text{d}s \quad (4.35)$$

aufgrund der Definitionen von  $\sqrt{\det(G(u))}$  (vgl. (4.29)) und Normalenfeld (vgl. (4.33)). Daher ist

$$\int_M v(x) \cdot \text{do} = \int_M v \cdot n \, \text{d}s. \quad (4.36)$$

Das vektorielle Oberflächenintegral lässt sich also mithilfe des Einheitsnormalenfeldes auch als skalares Oberflächenintegral schreiben. Das skalare Oberflächenintegral über die Normalenkomponente von  $v$  wird auch *Fluss von  $v$  durch  $M$*  genannt.<sup>2</sup>

Als Erinnerung sei an das folgende nützliche Werkzeug zur Berechnung des vektoriellen Oberflächenintegrals erinnert: Für Vektoren  $a, b$  und  $c$  gilt  $\langle v, a \times b \rangle = \det(v, a, b)$ . Daraus folgt

$$\langle v \circ \Phi(u), \partial_{u_1} \Phi(u) \times \partial_{u_2} \Phi(u) \rangle = \det \begin{pmatrix} v_1 \circ \Phi(u) & v_2 \circ \Phi(u) & v_3 \circ \Phi(u) \\ \partial_{u_1} \Phi_1(u) & \partial_{u_1} \Phi_2(u) & \partial_{u_1} \Phi_3(u) \\ \partial_{u_2} \Phi_1(u) & \partial_{u_2} \Phi_2(u) & \partial_{u_2} \Phi_3(u) \end{pmatrix}. \quad (4.37)$$

#### Beispiele

1. Sei die Einheitskugel  $S := \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$  und eine skalarwertige Temperaturfunktion  $T : S \rightarrow \mathbb{R}$  mit  $T(x, y, z) := x^2 + y^2 + z^2$  gegeben. Als Orientierung sei festgelegt, dass das Normalenfeld nach außen zeigt. Da alle Elemente von  $S$  bereits Norm 1 haben, ist das Normalenfeld bereits das Einheitsnormalenfeld, und es ist durch  $n : S \rightarrow \mathbb{R}^3$  mit

<sup>2</sup>Gibt  $v$  beispielsweise die Geschwindigkeit und Richtung an, mit der eine Flüssigkeit über einer Oberfläche fließt, so entspricht dem Integral die Durchflussrate. Das Integral gibt also an, wie viel der Flüssigkeit pro Zeiteinheit über die Oberfläche fließt.

$n(x, y, z) := (x, y, z)$  gegeben. Die Änderung der Temperatur wird nun durch den Gradienten  $\nabla T$  gegeben. Integriert man das Vektorfeld  $v : S \rightarrow \mathbb{R}^3$  mit

$$v(x, y, z) := -\nabla T(x, y, z) = \begin{pmatrix} -2x \\ -2y \\ -2z \end{pmatrix}$$

über die Oberfläche  $S$ , so ergibt sich der Wärmefluss. Wir berechnen also

$$\begin{aligned} \int_S v \cdot do &= \int_S v \cdot n \, ds = \int_S \langle v(x, y, z), n(x, y, z) \rangle \, ds \\ &= \int_S \left\langle \begin{pmatrix} -2x \\ -2y \\ -2z \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\rangle \, ds \\ &= -2 \int_S x^2 + y^2 + z^2 \, ds = -2 \int_S 1 \, ds = -8\pi, \end{aligned}$$

da der Oberflächeninhalt einer Kugel mit Radius  $r$  durch  $4\pi r^2$  gegeben ist. Das negative Vorzeichen hier bedeutet, dass der Fluss entgegen der Richtung des Einheitsnormalenfeldes, also nach innen gerichtet ist. Das ist anschaulich auch klar, da die Temperatur im Ursprung 0 und auf dem Rand von  $S$  1 ist. Im Sinne des Konzentrationsausgleichs findet der Wärmefluss also in Richtung des Ursprungs statt.

2. Sei  $U := (0, 1) \times (-\pi, \pi)$  und  $M$  die Schraubenfläche, die durch  $\Phi : U \rightarrow \mathbb{R}^3$  mit

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

parametrisiert wird. Wir weisen ihr die durch das Einheitsnormalenfeld  $n : M \rightarrow \mathbb{R}^3$  mit

$$n\left(\frac{1}{2}, 0, 0\right) := \left(0, -\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)^T$$

gegebene Orientierung zu. Sei weiter  $v : M \rightarrow \mathbb{R}^3$  das Vektorfeld mit  $v(x, y, z) := (y, -x, z)^T$ . Wir möchten das vektorielle Oberflächenintegral von  $v$  über  $M$  berechnen. Dazu nutzen wir die Darstellung (4.37). Wir bestimmen zunächst die partiellen Ableitungen von  $\Phi$  als

$$\partial_r \Phi(r, \varphi) := \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix} \quad \text{und} \quad \partial_\varphi \Phi(r, \varphi) := \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \\ 1 \end{pmatrix}.$$

Wir Laplace-entwickeln die Determinante (4.37) nach der letzten Spalte und finden

$$\begin{aligned} \det \begin{pmatrix} r \sin \varphi & -r \cos \varphi & \varphi \\ \cos \varphi & \sin \varphi & 0 \\ -r \sin \varphi & r \cos \varphi & 1 \end{pmatrix} &= + \det \begin{pmatrix} r \sin \varphi & r \cos \varphi \\ \cos \varphi & \sin \varphi \end{pmatrix} + \varphi \begin{pmatrix} \cos \varphi & \sin \varphi \\ -r \sin \varphi & r \cos \varphi \end{pmatrix} \\ &= r(1 + \varphi). \end{aligned}$$

Damit ist dann

$$\begin{aligned} \int_M v \, d\mathbf{o} &= \pm \int_U \langle v \circ \Phi(r, \varphi), \partial_r \Phi(r, \varphi) \times \partial_\varphi \Phi(r, \varphi) \rangle \, d(r, \varphi) \\ &= \pm \int_{-\pi}^{\pi} \int_0^1 r(1 + \varphi) \, dr \, d\varphi = \pm \pi, \end{aligned}$$

wobei wir hier zunächst den Richtungssinn der Parametrisierung ignoriert haben. Das holen wir nun nach. Wir bestimmen also den Richtungssinn der Parametrisierung gemäß (4.33). Zunächst sehen wir direkt, dass für  $u := (r, \varphi) = (\frac{1}{2}, 0)$  gilt

$$\Phi(u) = \left( \frac{1}{2}, 0, 0 \right).$$

Das ist praktischerweise genau der Punkt, durch den die Orientierung unseres Normalenfeldes bestimmt wird. Wir bestimmen die partiellen Ableitungen von  $\Phi$  an  $u$  und finden

$$\partial_r \Phi(u) = (1, 0, 0)^T \quad \text{und} \quad \partial_\varphi \Phi(u) = \left( 0, \frac{1}{2}, 1 \right)^T.$$

Das Kreuzprodukt der beiden Vektoren ergibt

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1/2 \end{pmatrix};$$

die Norm dieses Vektors ist  $\frac{\sqrt{5}}{2}$ . Nun finden wir, dass

$$n\left(\frac{1}{2}, 0, 0\right) = \begin{pmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = +\frac{2}{\sqrt{5}} \begin{pmatrix} 0 \\ -1 \\ 1/2 \end{pmatrix},$$

sodass  $\Phi$  gemäß der durch  $n$  festgelegten Orientierung positiven Richtungssinn hat. Damit ist das vektorielle Oberflächenintegral von  $v$  über  $M$  durch

$$\int_M v \, d\mathbf{o} = \pi$$

gegeben.

### 4.3 Klassische Integralsätze

Zunächst lernen wir drei Differentialoperatoren kennen, die in der mathematischen Physik eine fundamentale Rolle spielen. Differentialoperatoren sind Abbildungen zwischen Funktionenräumen, die ganz allgemein einer stetig differenzierbaren Funktion eine andere Funktion zuweisen.

**Definition 4.32:** Sei  $\Omega \subset \mathbb{R}^n$  ein Gebiet und  $v = (v_1, \dots, v_n) : \Omega \rightarrow \mathbb{R}^n$  ein  $\mathcal{C}^1$ -Vektorfeld. Dann definieren wir die *Divergenz*  $\operatorname{div} v : \Omega \rightarrow \mathbb{R}$  von  $v$  als

$$\operatorname{div} v(x) := \frac{\partial v_1}{\partial x_1}(x) + \dots + \frac{\partial v_n}{\partial x_n}(x). \quad (4.38)$$

Der Divergenzoperator ist also ein Operator  $\operatorname{div} : \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n) \rightarrow \mathcal{C}^0(\mathbb{R}^n, \mathbb{R})$ , der einem Vektorfeld eine skalarwertige Funktion zuweist. Die Divergenz eines Vektorfeldes entspricht der Spur seiner Jacobi-Matrix.

**Definition 4.33:** Sei  $\Omega \subset \mathbb{R}^n$  ein Gebiet und  $u = (v_1, \dots, v_n) : \Omega \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Dann definieren wir den *Laplace Operator*  $\Delta u : \Omega \rightarrow \mathbb{R}$  von  $u$  als

$$\Delta u(x) := \operatorname{div}(\nabla u)(x) = \frac{\partial^2 u_1}{\partial x_1^2}(x) + \dots + \frac{\partial^2 u_n}{\partial x_n^2}(x). \quad (4.39)$$

Der Laplace-Operator ist also ein Operator  $\Delta : \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}^0(\mathbb{R}^n, \mathbb{R})$ . Der Laplace-Operator weist einer Funktion die Spur ihrer Hesse-Matrix zu.

**Definition 4.34:** Sei  $\Omega \subset \mathbb{R}^3$  ein Gebiet und  $v : \Omega \rightarrow \mathbb{R}^3$  ein  $\mathcal{C}^1$ -Vektorfeld in drei Dimensionen. Die *Rotation des Vektorfeldes*  $\operatorname{rot} v : \Omega \rightarrow \mathbb{R}^3$  ist als

$$\operatorname{rot} v(x) := \left( \frac{\partial v_3}{\partial x_2}(x) - \frac{\partial v_2}{\partial x_3}(x) \right) \hat{e}_1 + \left( \frac{\partial v_1}{\partial x_3}(x) - \frac{\partial v_3}{\partial x_1}(x) \right) \hat{e}_2 + \left( \frac{\partial v_2}{\partial x_1}(x) - \frac{\partial v_1}{\partial x_2}(x) \right) \hat{e}_3 \quad (4.40)$$

definiert. Hierbei sind  $\hat{e}_1, \hat{e}_2, \hat{e}_3$  die kanonischen Einheitsvektoren.

Bemerkung: Die Rotation eines Vektorfeldes  $v$  lässt sich formal auch als

$$\operatorname{rot} v = \det \begin{pmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{pmatrix} \quad (4.41)$$

schreiben. (Man entwickle die Determinante nach der ersten Zeile.)

**Bemerkung 4.35:** In der Physik findet man auch die (sehr formalen) Schreibweisen

$$\operatorname{div} v = \nabla \cdot v, \quad \Delta u = \nabla^2 u, \quad \text{und} \quad \operatorname{rot} v = \nabla \times v.$$

Sie helfen aber unter Umständen, sich die Definition der Operatoren zu merken. Insbesondere für die Rotation kann

$$\nabla \times v = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

hilfreich sein.



**Lemma 4.36:** *Es gelten folgende Rechenregeln.*

1. Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  eine skalarwertige  $\mathcal{C}^1$ -Funktion und  $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ein  $\mathcal{C}^1$ -Vektorfeld. Dann gilt die Kettenregel

$$\operatorname{div}(fv)(x) = \langle \nabla f(v), v(x) \rangle + f(x) \cdot \operatorname{div} v(x). \quad (4.42)$$

Im Fall  $n = 3$ , also  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  und  $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  gilt außerdem

$$\operatorname{rot}(fv)(x) = \nabla f(x) \times v(x) + f(x) \cdot \operatorname{rot} v(x). \quad (4.43)$$

2. Sei nun  $\Omega \subset \mathbb{R}^3$  ein Gebiet. Für skalarwertige  $\mathcal{C}^2$ -Funktion  $u : \Omega \rightarrow \mathbb{R}$  gilt

$$\operatorname{rot}(\nabla u)(x) = 0. \quad (4.44)$$

Analog gilt für ein  $\mathcal{C}^2$ -Vektorfeld  $v : \Omega \rightarrow \mathbb{R}^3$

$$\operatorname{div}(\operatorname{rot} v)(x) = 0. \quad (4.45)$$

3. Sind  $v, w : \Omega \rightarrow \mathbb{R}^3$  zwei  $\mathcal{C}^1$ -Vektorfelder, so gilt

$$\operatorname{div}(v \times w)(x) = \langle \operatorname{rot} v(x), w(x) \rangle - \langle v(x), \operatorname{rot} w(x) \rangle \quad (4.46)$$

sowie

$$\operatorname{rot}(v \times w)(x) = \operatorname{div} w(x) \cdot v(x) - \operatorname{div} v(x) \cdot w(x) + Dv(x) \cdot w(x) - Dw(x) \cdot v(x). \quad (4.47)$$

4. Weiterhin gilt für Vektorfelder  $v \in \mathcal{C}^2(\Omega, \mathbb{R}^3)$

$$\operatorname{rot}(\operatorname{rot} v)(x) = \nabla(\operatorname{div}(v))(x) + \Delta v(x). \quad (4.48)$$

*Beweis.* Diese Gleichheiten lassen sich zur Übung formal nachrechnen. □

### 4.3.1 Satz von Stokes

**Satz 4.37 (Integralsatz von Stokes für Rechtecke):** Sei  $Q := (a, b) \times (c, d) \subset \mathbb{R}^2$  und seien  $f, g : \mathcal{U} \rightarrow \mathbb{R}$  zwei  $\mathcal{C}^1$ -Funktion auf einer Umgebung  $\mathcal{U}$  von  $\overline{Q}$ . Dann gilt

$$\int_Q \frac{\partial}{\partial x} g(x, y) - \frac{\partial}{\partial y} f(x, y) \, d^2(x, y) = \int_{\partial Q} f(x, y) \, ds + \int_{\partial Q} g(x, y) \, ds. \quad (4.49)$$

*Beweis.* Zunächst gilt

$$\begin{aligned} \int_Q \frac{\partial}{\partial x} g(x, y) \, d(x, y) &= \int_Q \frac{\partial}{\partial x} g(x, y) \, d(x, y) \\ &= \int_c^d \left( \int_a^b \frac{\partial}{\partial x} g(x, y) \, dx \right) dy = \int_c^d \left( g(b, y) - g(a, y) \right) dy. \end{aligned} \quad (*)$$

Hierbei ist der Hauptsatz der Differential- und Integralrechnung anwendbar, da  $g$  auf einer Umgebung von  $\bar{Q}$  differenzierbar ist. Der Rand von  $Q$  lässt sich in die Kanten aufteilen, die wir mithilfe der Funktionen  $\alpha_i$ ,  $i = 1, \dots, 4$  parametrisieren. Dabei sind  $\alpha_1, \alpha_3 : [a, b] \rightarrow \mathbb{R}^2$  und  $\alpha_2, \alpha_4 : [c, d] \rightarrow \mathbb{R}$  durch

$$\alpha_1(t) := \begin{pmatrix} t \\ c \end{pmatrix}, \quad \alpha_2 := \begin{pmatrix} b \\ t \end{pmatrix}, \quad \alpha_3(t) := \begin{pmatrix} a + b - t \\ d \end{pmatrix} \quad \text{sowie} \quad \alpha_4(t) := \begin{pmatrix} a \\ c + d - t \end{pmatrix}$$

gegeben. Wir durchlaufen den Rand also beginnend beim Punkt  $(a, c)$  entgegen dem Uhrzeigersinn.<sup>3</sup> Da  $g$  skalarwertig ist, könnten wir nun die rechte Seite von (\*) umschreiben als skalares Kurvenintegral über die Kurven  $\alpha_1, \dots, \alpha_4$ . Dabei würden wir allerdings die Orientierung der Kurven verlieren, die ja nur bei Vektorfeldern sinnvoll definiert ist. Daher interpretieren wir die Integrale in (\*) als vektorielle Kurvenintegrale des Vektorfeldes  $v = (0, g)^T$ , sodass

$$\begin{aligned} \int_c^d (g(b, y) - g(a, y)) dy &= \int_c^d g(b, y) dy - \int_c^d g(a, y) dy \\ &= \int_c^d \left\langle \begin{pmatrix} 0 \\ g(b, y) \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle dy + \int_c^d \left\langle \begin{pmatrix} 0 \\ g(a, y) \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\rangle dy \\ &= \int_c^d \left\langle \begin{pmatrix} 0 \\ g \circ \alpha_2(t) \end{pmatrix}, \dot{\alpha}_2(t) \right\rangle dt + \int_c^d \left\langle \begin{pmatrix} 0 \\ g \circ \alpha_4(t) \end{pmatrix}, \dot{\alpha}_4(t) \right\rangle dt \\ &= \int_{\alpha_2} g ds + \int_{\alpha_4} g ds. \end{aligned} \tag{**}$$

In diesem Sinne können wir auch das Wegintegral von  $g$  über  $\alpha_1$  interpretieren, und wir sehen, dass

$$\int_{\alpha_1} g ds = \int_c^d \left\langle \begin{pmatrix} 0 \\ g \circ \alpha_1(t) \end{pmatrix}, \dot{\alpha}_1(t) \right\rangle dt = 0.$$

Ebenso verschwindet das Wegintegral von  $g$  über  $\alpha_3$ . Damit ergibt sich aus (\*) und (\*\*)

$$\int_Q \frac{\partial}{\partial x} g(x, y) d(x, y) = \sum_{i=1}^4 \int_{\alpha_i} g ds = \int_{\alpha} g ds = \int_{\partial Q} g ds. \tag{4.50}$$

Die Analyse des zweiten Integrals  $\int_Q \frac{\partial}{\partial y} f(x, y) d(x, y)$  verläuft ganz analog. □

Bemerkung: Dieser Satz ist auch als *Satz von Green* oder *Satz von Green-Riemann* bekannt. Bemerge, dass mit  $\partial Q$  nicht der topologische Rand des Quaders  $Q$ , sondern vielmehr seine Parametrisierung gemeint ist. Das Integral auf der rechten Seite von (4.49) steht also für das Kurvenintegral des Vektorfeldes  $v = (f, g)$  über die stückweise glatte Kurve  $\alpha$ .

Nachdem wir den Satz von Stokes für Rechtecke gezeigt haben, wollen wir ihn nun auf abstraktere Formen verallgemeinern. Das funktioniert, solange der Rand der Formen gewisse Voraussetzungen erfüllt. Hierfür führen wir sogenannte *zweidimensionale Pflaster* ein, und betrachten diese als Teilmengen des  $\mathbb{R}^3$ .

<sup>3</sup>Nach Bemerkung 4.6 können wir natürlich auch jede andere Parametrisierung wählen.

**Definition 4.38:** Sei  $U \subset \mathbb{R}^2$  ein Gebiet mit  $[0, 1]^2 \subset U$ . Ein *zweidimensionales Pflaster* im  $\mathbb{R}^3$  ist ein orientiertes Flächenstück  $M \subset \mathbb{R}^3$ , dass sich durch eine  $\mathcal{C}^2$ -Parametrisierung  $\Phi : U \rightarrow \mathbb{R}^3$  beschreiben lässt, wobei  $\Phi$  folgende Eigenschaften besitzen soll:

1. Die Einschränkung  $\Phi|_{(0,1)^2}$  ist eine positive Parametrisierung von  $M$ .
2. Für  $k \in \{1, 2, 3, 4\}$  sei die  $k$ -te *Seitenkurve*  $\varphi_k : (0, 1) \rightarrow \mathbb{R}^3$  mit

$$\varphi_k := (\Phi \circ \alpha_k)|_{(0,1)}$$

entweder konstant (entartet) oder injektiv und regulär. Hierbei seien  $\alpha_k$  analog zum Beweis von Satz 4.37 definiert.

3. Für die Seiten  $\varphi_k((0, 1))$  der Fläche  $M$  gilt: Jede nicht entartete Seite trifft höchstens eine weitere nicht entartete Seite.
4. Trifft die  $j$ -te Seite die  $k$ -te Seite für  $j \neq k$ , ist also  $\varphi_j((0, 1)) \cap \varphi_k((0, 1)) \neq \emptyset$ , so existiert eine  $\mathcal{C}^1$ -Parametertransformation  $h : (0, 1) \rightarrow (0, 1)$  mit  $\varphi_j = \varphi_k \circ h$  und  $\dot{h} < 0$ . Diese Transformation nennt man auch *Klebung*.

Existiert ein solches  $\Phi$ , so heißt es auch *Pflasterparametrisierung von  $M$* .

**Interpretation der Forderungen an Pflasterparametrisierungen** Bevor wir einige Beispiele für Pflasterparametrisierungen sehen, wollen wir die gestellten Forderungen an  $\Phi$  und ihre Bedeutung kurz diskutieren. Zunächst verknüpft die erste Forderung Pflasterparametrisierungen mit den Flächenparametrisierungen (vgl. Definition 4.21). Die zweite Bedingung erlaubt, dass die Seiten des Einheitsquadrats durch die Parametrisierung auf Punkte zusammengezogen werden (konstante/entartete Kurve). Ist nämlich  $\varphi_k$  konstant in  $\mathbb{R}^2$ , bildet  $\Phi$  alle Punkte entlang der Seite  $\alpha_k$  auf den gleichen Punkt in  $\mathbb{R}^2$  ab. Das erlaubt beispielsweise dreieckige Flächen und Geraden  $M \subset \mathbb{R}^3$ . Ist die Seitenkurve nicht entartet, so soll sie hinreichend angenehm sein, um die entsprechende Seite von  $M$  vernünftig parametrisieren zu können (vgl. Definition 4.2).

Weiter sollen nicht entartete Seiten auf höchstens eine weitere nicht entartete Seite treffen. Das bedeutet, dass höchstens zwei nicht entartete Seiten einen (oder mehrere) gemeinsame Punkte besitzen. Damit wird ausgeschlossen, dass zwei nicht entartete Seite eine dritte beispielsweise in deren Mittelpunkt treffen. Die letzte Bedingung garantiert nun, dass zwei Seiten, die sich außerhalb ihrer Eckpunkte treffen im Prinzip überall treffen. Sie liegen also übereinander. Wegen der dritten Bedingung können höchstens zwei Seiten übereinander liegen. Die Forderung  $\dot{h} < 0$  ist nun gleichbedeutend damit, dass beide aufeinanderliegenden Seitenkurven einen unterschiedlichen Richtungssinn haben. Die Seiten werden also entgegengesetzt durchlaufen. Bemerke, dass die Eckpunkte wegen der Offenheit der  $\varphi_k$  nicht zur Seitenkurve zählen.

**Bemerkung 4.39:** Wir schreiben in Zukunft  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  und meinen damit, dass  $\Phi$ , wie in der Definition 4.38 gefordert, bis zum Rand hin (und darüber hinaus) mindestens zweimal stetig differenzierbar ist. Bemerke, dass Differenzierbarkeit auf offenen Mengen definiert ist, die partiellen Ableitungen aber wegen ihrer vorausgesetzten Stetigkeit bis zum Rand hin stetig fortgesetzt werden können. Weiter ist  $M = \Phi((0, 1)^2)$ .

**Beispiele**

1. Sei  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  durch

$$\Phi(u_1, u_2) := \begin{pmatrix} u_1 \\ u_1 u_2 \\ 0 \end{pmatrix}$$

gegeben. Diese Pflasterparametrisierung beschreibt ein Dreieck. Bemerke, dass  $\varphi_4$  entartet ist. Die vierte Seitenfläche wurde also auf den Anfangspunkt  $\Phi(0, 0)$  zusammengezogen.

2. Sei  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  durch

$$\Phi(u_1, u_2) := \begin{pmatrix} u_1 \cos(2\pi u_2) \\ u_1 \sin(2\pi u_2) \\ 0 \end{pmatrix}$$

gegeben. Hier beschreibt  $\Phi$  den geschlitzten Einheitskreis in der  $xy$ -Ebene. Die Seitenkurve  $\varphi_1$  bildet  $(u_1, 0) \mapsto (u_1, 0, 0)^T$  ab. Die Seitenkurve  $\varphi_2$  beschreibt nun die Kreiskurve, und  $\varphi_3$  verläuft vom Punkt  $(1, 0)$  zurück zum Ursprung, also auf  $\varphi_1$  mit umgekehrten Richtungssinn. Die Kurve  $\varphi_4$  ist wieder entartet.

3. Sei  $\theta \in (0, 2\pi]$  fixiert und  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  durch

$$\Phi(u_1, u_2) := \begin{pmatrix} \cos(\theta u_1) \\ \sin(\theta u_1) \\ u_2 \end{pmatrix}$$

gegeben. Hier beschreibt  $\Phi$  einen offenen (im Fall  $\theta = 2\pi$  geschlossenen) Zylindermantel.

**Definition 4.40:** Sei  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  eine Pflasterparametrisierung für  $M \subset \mathbb{R}^3$ . Wir bezeichnen wie zuvor die Seitenkurven von  $M$  mit  $\varphi_k$ . Aus der Menge  $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$  entfernen wir alle entarteten Kurven und auch alle übereinander liegenden Kurven, also alle Paare  $(\varphi_j, \varphi_k)$ , die das gleiche Kurvenstück mit entgegengesetztem Richtungssinn parametrisieren. Die übrigen Seiten bezeichnen wir mit  $\bar{\varphi}_1, \dots, \bar{\varphi}_m$ , wobei  $0 \leq m \leq 4$ . Sei  $U$  eine Umgebung des Abschlusses  $\bar{M}$  und  $v : U \rightarrow \mathbb{R}^3$  ein Vektorfeld. Dann definieren wir das *Integral von  $v$  über den Rand  $\partial M$  von  $M$*  durch

$$\int_{\partial M} v \cdot ds := \sum_{l=1}^m \int_{\bar{\varphi}_l} v \cdot ds, \tag{4.51}$$

falls  $m > 0$ . Im Fall  $m = 0$  setzen wir

$$\int_{\partial M} v \cdot ds := 0. \tag{4.52}$$

**Bemerkung 4.41:** Beim Entfernen der entarteten und übereinander liegenden Kurven kann es sein, dass alle Kurven entfernt werden. Das ist beispielsweise beim Torus der Fall, oder auch bei der Einheitssphäre ohne Nullmeridian, also dem Oberfläche der Einheitskugel in drei Raumdimensionen ohne die durch die  $x$ -Achse gehende Verbindungskurve von  $(0, 0, 1)$  zu  $(0, 0, -1)$  auf der Kugeloberfläche. Der Hintergrund des Entfernen der Seitenkurven ist die Idee, dass

Kurvenintegrale über Punkte (auf Nullmengen) verschwinden, und sich zudem Integrale über die gleichen Kurven, die mit entgegengesetztem Richtungssinn durchlaufen werden, aufheben. Sie spielen also für das Integral, also beispielsweise die Arbeit, die entlang eines Weges verrichtet wird, keine Rolle.

**Satz 4.42 (Stokes):** Sei  $v : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ein gegebenes  $C^1$ -Vektorfeld und  $M$  ein Pflaster mit Rand  $\partial M$ . Dann gilt

$$\int_M \operatorname{rot} v \, \operatorname{do}(x) = \int_{\partial M} v \cdot ds. \quad (4.53)$$

*Beweis.* Wir bezeichnen mit  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  die Pflasterparametrisierung von  $M$ . Zunächst gilt nach Definition 4.40

$$\int_{\partial M} v \cdot ds = \sum_{l=1}^m \int_{\overline{\varphi}_l} v \cdot ds = \sum_{k=1}^4 \int_{\varphi_k} v \, dx.$$

Wir bezeichnen wie zuvor mit  $\alpha_k : [0, 1] \rightarrow \mathbb{R}^2$ ,  $\alpha_k = (\alpha_{k1}, \alpha_{k2})$  die Parametrisierung der Quadratseiten und mit  $\varphi_k := \Phi \circ \alpha_k$  die entsprechenden Seitenkurven. Aus der Kettenregel folgt

$$\dot{\varphi}_k(t) = \partial_1 \Phi(\alpha_k(t)) \circ \dot{\alpha}_{k1}(t) + \partial_2 \Phi(\alpha_k(t)) \circ \dot{\alpha}_{k2}(t).$$

Setzen wir das in das Integral ein, so erhalten wir für jedes  $k = 1, \dots, 4$

$$\begin{aligned} \int_{\varphi_k} v \cdot ds &= \int_0^1 \langle v \circ \varphi_k(t), \dot{\varphi}_k(t) \rangle dt \\ &= \int_0^1 \langle v \circ \Phi(\alpha_k(t)), \partial_1 \Phi(\alpha_k(t)) \rangle \dot{\alpha}_{k1}(t) \\ &\quad + \langle v \circ \Phi(\alpha_k(t)), \partial_2 \Phi(\alpha_k(t)) \rangle \dot{\alpha}_{k2}(t) dt \\ &= \int_0^1 \left\langle \begin{pmatrix} P_1(\alpha_k(t)) \\ P_2(\alpha_k(t)) \end{pmatrix}, \dot{\alpha}_k(t) \right\rangle dt \\ &= \int_{\alpha_k} P_1 \, du_1 + \int_{\alpha_k} P_2 \, du_2, \end{aligned}$$

wobei die Funktionen  $P_i : \mathbb{R}^2 \rightarrow \mathbb{R}$  durch  $P_i((u_1, u_2)) := \langle v \circ \Phi(u_1, u_2), \partial_i \Phi(u_1, u_2) \rangle$  definiert sind. Wir haben das Integral also als Integral des Vektorfeldes  $P = (P_1, P_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  über die Kurve  $\alpha_k$  umgeschrieben. Auf dieses Vektorfeld wenden wir nun Satz 4.37 an und erhalten

$$\begin{aligned} \int_{\partial M} v \cdot ds &= \sum_{k=1}^4 \int_{\alpha_k} P_1 \, du_1 + P_2 \, du_2 \\ &= \int_{\partial[0,1]^2} P_1 \, du_1 + P_2 \, du_2 = \int_{[0,1]^2} \partial_1 P_2 - \partial_2 P_1 \, d^2u \end{aligned} \quad (4.54)$$

Wir zerlegen unser ursprüngliches Vektorfeld  $v$  nun in seine Komponenten, schreiben also

$$v = v_1 e_1 + v_2 e_2 + v_3 e_3$$

mit den drei kanonischen Einheitsvektoren  $e_1, e_2, e_3 \in \mathbb{R}^3$ . Es reicht daher ohne Beschränkung der Allgemeinheit, im Weiteren nur noch Vektorfelder zu betrachten, deren zweite und dritte Komponente verschwindet. Die Argumentation lässt sich dann wegen Linearität auf allgemeine Vektorfelder übertragen.

Sei also  $v = f \cdot e_1$  für ein  $f \in C^1(\Omega, \mathbb{R})$ . Dann schreiben sich die obigen Funktionen  $P_i$ ,  $i = 1, 2$  als

$$P_i = \langle (f \circ \Phi)e_1, \partial_i \Phi \rangle = f \circ \Phi \cdot \partial_i \Phi_1.$$

Anwendung von Produkt- und Kettenregel liefert

$$\partial_1 P_2 = \partial_2 \Phi_1 \sum_{k=1}^3 \partial_k f(\Phi) \partial_1 \Phi_k + f(\Phi) \partial_1 \partial_2 \Phi_1.$$

und

$$\partial_2 P_1 = \partial_1 \Phi_1 \sum_{k=1}^3 \partial_k f(\Phi) \partial_2 \Phi_k + f(\Phi) \partial_2 \partial_1 \Phi_1$$

Bemerke, dass die zweiten Ableitungen von  $\Phi$  existieren, da  $\Phi$  eine Pflasterparametrisierung ist. Subtrahieren wir beide Gleichungen voneinander, fällt der letzte Term weg. Außerdem stimmen die Summanden für  $k = 1$  überein, sodass

$$\partial_1 P_2 - \partial_2 P_1 = \partial_2 f \left( \partial_2 \Phi_1 \partial_1 \Phi_2 - \partial_1 \Phi_1 \partial_2 \Phi_2 \right) - \partial_3 f \left( \partial_1 \Phi_3 \partial_2 \Phi_1 - \partial_1 \Phi_1 \partial_2 \Phi_3 \right) \quad (4.55)$$

übrig bleibt. Auf der anderen Seite ist für  $v = f \cdot e_1$  nach (4.40)

$$\text{rot } v = \text{rot}(f \cdot e_1) = \partial_3 f e_2 - \partial_1 f e_3.$$

Nun gilt wegen der positiven Orientierung von  $\Phi$  bezüglich  $M$ , dass das vektorielle Flächenelement nach (4.35) durch

$$\begin{aligned} \text{do}(x) &= \partial_1 \Phi \times \partial_2 \Phi \, d^2 u = (\partial_1 \Phi_2 \partial_2 \Phi_3 - \partial_2 \Phi_2 \partial_1 \Phi_3) e_1 \\ &\quad + (\partial_1 \Phi_3 \partial_2 \Phi_1 - \partial_2 \Phi_3 \partial_1 \Phi_1) e_2 \\ &\quad + (\partial_1 \Phi_1 \partial_2 \Phi_2 - \partial_2 \Phi_1 \partial_1 \Phi_2) e_3 \, d^2 u \end{aligned}$$

gegeben ist. Berechnen wir nun  $\text{rot } v \cdot \text{do}(x) = \langle \text{rot } v, \text{do}(x) \rangle$  und vergleichen das mit (4.55), so sehen wir

$$((\text{rot } v) \circ \Phi) \cdot \text{do}(x) = \langle (\text{rot } v) \circ \Phi, \text{do}(x) \rangle = \partial_1 P_2 - \partial_2 P_1 \, d^2 u.$$

Diese Argumentation gilt analog für Vektorfelder, deren erste und dritte oder erste und zweite Komponente verschwinden, mithin für allgemeine Vektorfelder aufgrund der Linearität der Zerlegung in seine Komponenten. Mit (4.54) und  $M = \Phi((0, 1)^2)$  folgt schließlich das Resultat

$$\begin{aligned} \int_{\partial M} v \cdot ds &= \int_{[0,1]^2} \partial_1 P_2 - \partial_2 P_1 \, d^2 u \\ &= \int_{[0,1]^2} \langle (\text{rot } v) \circ \Phi, \partial_1 \Phi \times \partial_2 \Phi \rangle \, d^2 u \\ &= \int_M \text{rot } v \cdot \text{do}(x). \quad \square \end{aligned}$$

**Beispiel** Sei  $M \subset \mathbb{R}^3$  ein Torus und  $\Phi$  eine positive Parametrisierung von  $M$ . Nach Bemerkung 4.41 bleiben beim Entfernen ungültiger Seitenkurven der Parametrisierung  $\Phi$  zur Definition des Integrals eines Vektorfelds  $v : M \rightarrow \mathbb{R}^3$  keine Seitenflächen übrig. Die Menge der zulässigen Seitenkurven ist also leer. Damit ist jedes Vektorfeld  $v$  über einem Torus rotationsfrei, da nach (4.52)

$$\int_M \operatorname{rot} v \cdot d\mathbf{o} = \int_{\partial M} v \cdot d\mathbf{s} = 0.$$

**Definition 4.43:** Eine Menge  $M \subset \mathbb{R}^3$  heißt *Pflasterkette in  $\mathbb{R}^3$* , wenn eine Menge von zwei-dimensionalen Pflasterparametrisierungen  $\Phi_1, \dots, \Phi_n : [0, 1]^2 \rightarrow \mathbb{R}^3$  existiert, sodass folgende Eigenschaften erfüllt sind:

1.  $M$  lässt sich in  $n$  disjunkte Teilmengen  $M_1, \dots, M_n$  zerlegen, die sich durch die entsprechenden Pflasterparametrisierungen parametrisieren lassen. Also

$$M_i \cap M_k = \emptyset \quad \text{für } i \neq k \quad \text{sowie} \quad \Phi_i((0, 1)^2) = M_i$$

Die Seitenkurven von  $M_i$  bezeichnen wir mit  $\varphi_{ik} = \Phi_i \circ \alpha_k \Big|_{(0,1)}$ .

2. Eine nicht-entartete Seite  $\varphi_{ij}((0, 1))$  darf höchstens eine weitere nicht-entartete Seite  $\varphi_{kl}$  treffen.
3. Treffen sich zwei beliebige nicht-entarteten Seiten  $\varphi_{ij}$  und  $\varphi_{kl}$ , so stimmen sie überein und sind entgegengesetzt zueinander orientiert. Es existiert also eine  $\mathcal{C}^1$ -Parametertransformation  $h : [0, 1] \rightarrow [0, 1]$  mit  $\varphi_{ij} = \varphi_{kl} \circ h$  und  $\dot{h} < 0$  (*Klebung*).

Bemerke, dass die letzten beiden Forderungen übergreifend über alle Teilpflaster  $M_i$  gelten.

Für eine Pflasterkette  $M$  schreiben wir auch  $M = M_1 + \dots + M_n$ . Für die Vereinigung aller Teilpflaster mit ihren geklebten Randseiten schreiben wir

$$|M| := \bigcup_{i=1}^n \overline{M}_i. \tag{4.56}$$

**Definition 4.44:** Sei  $M \subset \mathbb{R}^3$  eine Pflasterkette. Für ein Vektorfeld  $v : |M| \rightarrow \mathbb{R}^3$  setzen wir

$$\int_M v \cdot d\mathbf{o} := \sum_{i=1}^n \int_{M_i} v \cdot d\mathbf{o}. \tag{4.57}$$

Für das Integral über den Pflasterrand entfernen wir wie zuvor die entarteten und geklebten Seiten. Die übrigen Seitenkurven bezeichnen wir mit  $\overline{\varphi}_k, k = 1, \dots, m \leq n$  sodass  $\partial M = \overline{\varphi}_1 + \dots + \overline{\varphi}_m$ . Das Integral des Vektorfelds über den Pflasterrand definieren wir vormalig durch

$$\int_{\partial M} v \cdot d\mathbf{s} := \sum_{k=1}^m \int_{\overline{\varphi}_k} v \cdot d\mathbf{s}, \tag{4.58}$$

falls  $m > 0$  beziehungsweise

$$\int_{\partial M} v \cdot d\mathbf{s} := 0 \tag{4.59}$$

sonst.

**Satz 4.45 (Stokes für Pflasterketten):** Ist  $M = M_1 + \dots + M_n$  eine Pflasterkette im  $\mathbb{R}^3$  und  $v : \Omega \subset |M| \rightarrow \mathbb{R}^3$  ein  $\mathcal{C}^1$ -Vektorfeld, so gilt

$$\int_M \operatorname{rot} v \cdot d\mathbf{o} = \int_{\partial M} v \cdot d\mathbf{s}. \quad (4.60)$$

*Beweis.* Der Beweis überträgt sich direkt. □

### 4.3.2 Satz von Green

Wir betrachten nun ebene Flächenstücke und ebene Vektorfelder. Ein ebenes Flächenstück  $M \subset \mathbb{R}^3$  ist ein Flächenstück in der  $xy$ -Ebene, das durch das Normalenfeld  $e_3$  orientiert ist. Insbesondere verschwindet die dritte Komponente einer Pflasterparametrisierung  $\Phi : [0, 1]^2 \rightarrow M$ , also  $\Phi = (\Phi_1, \Phi_2, 0)$ . Wir können daher auch  $M \subset \mathbb{R}^2$  annehmen. Damit lässt sich der Satz von Stokes auf ebene Vektorfelder  $v : M \rightarrow \mathbb{R}^2$  anwenden, wenn wir diese natürlich im  $\mathbb{R}^3$  einbetten. Wir identifizieren dabei  $v : M \rightarrow \mathbb{R}^2$  und  $\bar{v} : M \rightarrow \mathbb{R}^3$ , wobei

$$v(x, y) = \begin{pmatrix} v_1(x, y) \\ v_2(x, y) \end{pmatrix} \quad \text{und} \quad \bar{v}(x, y, z) = \begin{pmatrix} v_1(x, y) \\ v_2(x, y) \\ 0 \end{pmatrix}.$$

Gleiches funktioniert natürlich auch für Pflasterketten. Insbesondere gilt für das vektorielle Oberflächenelement (4.35)

$$d\mathbf{o} = (\partial_1 \Phi \times \partial_2 \Phi) e_3 d^2u = |\det J_\Phi| e_3 d^2u = e_3 ds,$$

da das skalare Oberflächenelement nach Definition 4.28 und (4.29)

$$ds = |\partial_1 \Phi \times \partial_2 \Phi| d^2u = |\det J_\Phi| d^2u.$$

erfüllt. Auch lassen sich hier Integrale von skalarwertigen Funktionen (skalaren Feldern) über Pflastern auf gewöhnliche Integrale zurückführen, da

$$\int_M f d\mathbf{o} = \int_{[0,1]^2} f(\Phi(u)) |\det J_\Phi(u)| d^2u = \int_M f(x, y) dx dy \quad (4.61)$$

unter Verwendung des Transformationssatzes gilt.

**Satz 4.46 (Green):** Sei  $M \subset \mathbb{R}^3$  eine ebene Pflasterkette, also eine Pflasterkette in der  $xy$ -Ebene. Wir betrachten ein ebenes  $\mathcal{C}^1$ -Vektorfeld  $v = (v_1, v_2) : \Omega \subset M \rightarrow \mathbb{R}^2$ . Dann gilt

$$\int_{|M|} \left( \frac{\partial}{\partial x} v_2(x, y) - \frac{\partial}{\partial y} v_1(x, y) \right) d^2(x, y) = \int_{\partial M} v_1 dx + v_2 dy. \quad (4.62)$$

*Beweis.* Nach obigen Überlegungen ist die Rotation von  $v$  sinnvoll zu definieren und es gilt

$$\operatorname{rot} v = \operatorname{rot} \bar{v} \cdot e_3 = \left( \frac{\partial}{\partial x} v_2 - \frac{\partial}{\partial y} v_1 \right) e_3.$$



Auch stimmen entspricht das skalare Oberflächenelement der dritten Komponente des vektoriellen Oberflächenelements, also  $do = e_3 ds$  gilt. Es folgt mit Satz 4.45

$$\int_{|M|} \partial_x v_2 - \partial_y v_1 d^2(x, y) = \int_M \operatorname{rot} \bar{v} \cdot e_3 ds = \int_M \operatorname{rot} v do = \int_{\partial M} v \cdot ds = \int_{\partial M} v_1 dx + v_2 dy. \quad \square$$

**Korollar 4.47:** Sei  $M \subset \mathbb{R}^2$  eine ebene Pflasterkette. Dann lässt sich der Flächeninhalt von  $M$  durch

$$A(M) := \frac{1}{2} \int_{\partial M} x dy - y dx \quad (4.63)$$

berechnen.

*Beweis.* Wir betrachten das ebene Vektorfeld  $v : M \rightarrow \mathbb{R}^2$  mit  $v(x, y) := (-y, x)$ . Dann gilt nach Satz 4.46

$$\frac{1}{2} \int_{\partial M} x dy - y dx = \frac{1}{2} \int_{|M|} \partial_x x + \partial_y y d^2(x, y) = \int_{|M|} d^2(x, y). \quad \square$$

### Beispiele

1. Wir verwenden den Satz von Green zur Berechnung des Flächeninhalts der Hypozykloide. Sei hierfür  $a > 0$  fixiert und  $M := \{(x, y) \in \mathbb{R}^2, x^{3/2} + y^{3/2} = a^{3/2}\}$ . Wir verwenden für die Parametrisierung des Randes die Funktion  $\sigma : [0, 2\pi] \rightarrow \partial M \subset \mathbb{R}^2$  mit

$$\sigma(\theta) := \begin{pmatrix} a \cos^3(\theta) \\ \sin^3(\theta) \end{pmatrix}.$$

Dann berechnen wir unter Anwendung von (4.4) auf das Vektorfeld  $v = (y, -x)$

$$\begin{aligned} A &= \frac{1}{2} \int_{\partial M} x dy - y dx = \frac{1}{2} \int_0^{2\pi} a \cos^3 \theta \cdot 3a \sin^2 \theta \cos \theta - a \sin^3 \theta \cdot (-3a \cos^2 \theta \sin \theta) d\theta \\ &= \frac{3}{2} a^2 \int_0^{2\pi} \sin^2(\theta) \cos^4(\theta) + \cos^2(\theta) \sin^4(\theta) d\theta \\ &= \frac{3}{2} a^2 \int_0^{2\pi} \sin^2(\theta) \cos^2(\theta) d\theta \\ &= \frac{3}{8} a^2 \int_0^{2\pi} \sin^2(2\theta) d\theta \\ &= \frac{3}{8} a^2 \int_0^{2\pi} \frac{1 - \cos(4\theta)}{2} d\theta = \frac{3}{8} a^2 \left[ \frac{\theta}{2} - \frac{\sin(4\theta)}{8} \right]_0^{2\pi} = \frac{3}{16} a^2 \cdot 2\pi = \frac{3}{8} a^2 \pi. \end{aligned}$$

2. Sei nun das Vektorfeld  $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $v(x, y) := (xy^2, y + x)$  und das zweidimensionale Pflaster  $M := \{(x, y) \in \mathbb{R}^2, 0 \leq x \leq 1, x^2 \leq y \leq x\}$  gegeben. Wir berechnen das Integral der Rotation des Vektorfelds über  $M$ . Hierbei bieten sich zwei mögliche Wege an.

Erstens: Zur Bestimmung der Rotation betten wir das Vektorfeld wieder im  $\mathbb{R}^3$  ein und sehen, dass

$$\operatorname{rot} v(x, y) = \begin{pmatrix} 0 \\ 0 \\ 1 - 2xy \end{pmatrix}.$$

Damit ergibt sich

$$\int_M \operatorname{rot} v \cdot d\mathbf{o} = \int_{\partial M} v \cdot d\mathbf{s} = \int_0^1 \int_{x^2}^x (1 - 2xy) dy dx = \int_0^1 x - x^3 - x^2 + x^5 dx = \frac{1}{12}.$$

Zweitens können wir auch den Rand parametrisieren und orientieren. Mithilfe des Satzes von Green schreiben wir

$$\int_M \operatorname{rot} v \cdot d\mathbf{o} = \int_M \partial_x v_2 - \partial_y v_1 d^2(x, y) = \int_{\partial M} v_1 dx + v_2 dy.$$

Wir teilen nun den Rand  $\partial M$  auf in die Kurven  $y = x$  und  $y = x^2$  und parametrisieren diese Kurven durch  $\sigma_1, \sigma_2 : \mathbb{R} \rightarrow \mathbb{R}^2$  mit

$$\sigma_1(t) := \begin{pmatrix} t \\ t \end{pmatrix} \quad \text{und} \quad \sigma_2(t) := \begin{pmatrix} t \\ t^2 \end{pmatrix}.$$

Damit berechnen sich die Kurvenintegrale als

$$\int_{y=x} v_1 dx + v_2 dy = \int_{y=x} xy^2 dy + (x + y) dy = \int_0^1 t^3 + 2t dt = \frac{5}{4}$$

und

$$\int_{y=x^2} v_1 dx + v_2 dy = \int_0^1 t^5 \cdot (t + t^2) \cdot 2t dt = \frac{4}{3}.$$

Da wir natürlich beim Durchlaufen des Randes einen der beiden Wege entgegen der Parametrisierung durchlaufen, hat das entsprechende Integral das entgegengesetzte Vorzeichen. Es ergibt sich damit

$$\int_{\partial M} v \cdot d\mathbf{s} = \pm \left( \int_{\sigma_1} v \cdot d\mathbf{s} - \int_{\sigma_2} v \cdot d\mathbf{s} \right) = \pm \frac{1}{12}.$$

### 4.3.3 Satz von Gauß

Wir betrachten zunächst den Einheitswürfel in  $\mathbb{R}^3$

$$Q_3 := \{(x_1, x_2, x_3) \in \mathbb{R}^3, x_i \in (0, 1) \text{ für } i = 1, 2, 3\}.$$

Wir bezeichnen die Seitenflächen des Würfels mit  $E_{ij}$ ,  $i \in \{1, 2, 3\}$ ,  $j \in \{0, 1\}$ , wobei

$$E_{1j} := \{(j, x_2, x_3) \in \mathbb{R}^3, x_2, x_3 \in (0, 1)\}$$

$$E_{2j} := \{(x_1, j, x_3) \in \mathbb{R}^3, x_1, x_3 \in (0, 1)\}$$

$$E_{3j} := \{(x_1, x_2, j) \in \mathbb{R}^3, x_1, x_2 \in (0, 1)\}.$$

Zusätzlich legen wir die Orientierung durch

$$n = -e_j \quad \text{auf } E_{i0} \qquad \text{und} \qquad n = e_j \quad \text{auf } E_{i1}$$

fest. Die Normalenvektoren zeigen also immer nach außen. Dadurch ist ein Einheitsnormalenfeld  $n$  auf dem Rand  $\partial Q_3$  festgelegt.

**Satz 4.48 (Gauß):** Sei  $v : \Omega \subset \overline{Q_3}$  ein Vektorfeld, das stetig differenzierbar auf einer Umgebung des abgeschlossenen Einheitswürfels ist. Dann gilt

$$\int_{Q_3} \operatorname{div} v \, d^3x = \int_{\partial Q_3} v \cdot d\mathbf{o}, \tag{4.64}$$

wobei das vektorielle Oberflächenintegral als

$$\int_{\partial Q_3} v \, d\mathbf{o} := \sum_{i=1}^3 \sum_{j=0}^1 \int_{E_{ij}} v \cdot d\mathbf{o} = \sum_{i=1}^3 \sum_{j=0}^1 \int_{E_{ij}} v \cdot n \, ds$$

zu verstehen ist.

*Beweis.* Sei  $Q_2 = (0, 1)^2$  der zweidimensionale Einheitswürfel. Dann lassen sich die Würfelseiten von  $Q_3$  kanonisch durch  $\Lambda_{ij} : \overline{Q_2} \rightarrow \overline{E_{ij}}$  mit

$$\begin{aligned} \Lambda_{10}(u_1, u_2) &:= (0, u_1, u_2), & \Lambda_{11}(u_1, u_2) &:= (1, u_1, u_2) \\ \Lambda_{20}(u_1, u_2) &:= (u_1, 0, u_2), & \Lambda_{21}(u_1, u_2) &:= (u_1, 1, u_2) \\ \Lambda_{30}(u_1, u_2) &:= (u_1, u_2, 0), & \text{sowie} & \Lambda_{31}(u_1, u_2) &:= (u_1, u_2, 1) \end{aligned}$$

parametrisieren. Hierbei ist jedes  $\Lambda_{ij}|_{(0,1)^2}$  eine positive Parametrisierung der Seitenfläche  $E_{ij}$ .

Nach Wahl des Einheitsnormalenfeldes ist

$$\int_{E_{ij}} v \cdot d\mathbf{o} = \int_{E_{ij}} v \cdot n \, ds = (-1)^{j+1} \int_{E_{ij}} v \cdot e_j \, ds = (-1)^{j+1} \int_{E_{ij}} v_j \, ds.$$

Weiterhin gilt

$$\begin{aligned} \int_{Q_3} \frac{\partial}{\partial x_1} v_1 \, d^3x &= \int_0^1 \int_0^1 \int_0^1 \frac{\partial}{\partial x} v_1(x_1, x_2, x_3) \, dx_1 \, dx_2 \, dx_3 \\ &= \int_0^1 \int_0^1 v_1(1, x_2, x_3) - v_1(0, x_2, x_3) \, dx_2 \, dx_3 \\ &= \int_0^1 \int_0^1 (v_1 \circ \Lambda_{11})(x_2, x_3) \, dx_2 \, dx_3 - \int_0^1 \int_0^1 (v_1 \circ \Lambda_{10})(x_2, x_3) \, dx_2 \, dx_3 \\ &= \int_{E_{11}} v_1 \, ds + \int_{E_{10}} v_1 \, ds = \int_{E_{11}} v \cdot d\mathbf{o} + \int_{E_{10}} v \cdot d\mathbf{o}. \end{aligned}$$

Hierbei lässt sich genauso argumentieren wie im Beweis von Satz 4.37. Analog folgt für die anderen Ableitungen

$$\int_{Q_3} \frac{\partial}{\partial x_2} v_2 \, d^3x = \int_{E_{21}} v \cdot do + \int_{E_{20}} v \cdot do$$

und

$$\int_{Q_3} \frac{\partial}{\partial x_3} v_2 \, d^3x = \int_{E_{31}} v \cdot do + \int_{E_{30}} v \cdot do.$$

Addieren wir all diese Gleichungen, führt das auf die zu beweisende Aussage.  $\square$

**Bemerkung 4.49:** Der Satz gilt auch unter der schwächeren Voraussetzung, dass  $v$  stetig auf dem gesamten abgeschlossenen Einheitswürfel und im Inneren stetig differenzierbar mit beschränkten partiellen Ableitungen ist.

**Definition 4.50:** Ein Gebiet  $\Omega \subset \mathbb{R}^3$  heißt *dreidimensionales Pflaster*, wenn es eine Umgebung von  $U$  von  $[0, 1]^3$  und eine  $\mathcal{C}^2$ -Abbildung  $\Phi : U \rightarrow \mathbb{R}^3$  gibt, sodass

1. Die Einschränkung  $\Phi|_{(0,1)^3}$  ist ein positiver Diffeomorphismus ( $\det h > 0$ ).
2. Für  $i \in \{1, 2, 3\}$  und  $j \in \{0, 1\}$  die Einschränkung der Funktion  $\varphi_{ij} : [0, 1]^2 \rightarrow \mathbb{R}^3$  mit  $\varphi_{ij} := h \circ \Lambda_{ij}$  auf das offene Einheitsquadrat  $(0, 1)^2$  entweder eine Parametrisierung der Seite  $M_{ij} := \varphi_{ij}((0, 1)^2)$  oder entartet. Hierbei bedeutet entartet, dass die Jacobimatrix  $D\varphi_{ij}$  von Rang kleiner gleich 1 ist. Die  $\Lambda_{ij}$  seien analog zum Beweis von Satz 4.48 definiert.
3. Falls zwei nicht-entartete Seiten nicht disjunkt sind, also  $M_{ij} \cap M_{kl} \neq \emptyset$ , so liegen sie übereinander (stimmen überein). Die dazugehörigen Parametrisierungen  $\varphi_{ij}$  und  $\varphi_{kl}$  seien entgegengesetzt orientiert. Es existiert also ein  $\mathcal{C}^2$ -Diffeomorphismus  $h$  mit  $\det Dh < 0$  und  $\varphi_{ij} = \varphi_{kl} \circ h$ .

Existiert ein solches  $\Phi$ , so heißt es auch *Pflasterparametrisierung von  $\Omega$* .

**Bemerkung 4.51:** Wie zuvor auch werden entartete Seitenkurven zugelassen. Damit können auch Figuren mit weniger als sechs Seitenflächen pflasterparametrisiert werden. Hierfür werden wie im zweidimensionalen Fall zwei Seitenflächen durch den  $\mathcal{C}^2$ -Diffeomorphismus  $h$  miteinander verklebt.

Bevor wir den Satz von Gauß auf dreidimensionale Objekte verallgemeinern, erweitern wir die zulässigen Mengen auf sogenannte *Gaußsche Gebiete*. Dabei gehen wir ähnlich vor wie beim Übergang zu Pflastern: Sei  $\Omega \subset \mathbb{R}^3$  ein dreidimensionales Pflaster und bezeichne  $\varphi_{ij}((0, 1)^2)$  die entsprechenden Seitenflächen. Wir entfernen die entarteten Seitenflächen und alle Paare  $(\varphi_{ij}, \varphi_{kl})$  von verklebten Seitenflächen. Die Menge der übrigen Seitenflächen bezeichnen wir mit  $M_1, \dots, M_m$ ,  $1 \leq m \leq 6$  und nennen sie *zulässige Teilmengen*.

**Definition 4.52:** Ein beschränktes Gebiet  $\Omega \subset \mathbb{R}^3$  heißt *Gaußsches Gebiet*, wenn es ein dreidimensionales Pflaster  $\tilde{\Omega} \subset \Omega$  und eine dazugehörige Pflasterparametrisierung  $\Phi : [0, 1]^3 \rightarrow \mathbb{R}^3$ ,  $\Phi((0, 1)^3) = \tilde{\Omega}$  gibt, sodass die folgenden Eigenschaften erfüllt sind

1. Für den topologische Rand  $\partial M$  gilt:  $\partial M = \cup_{i=1}^m |M|_i$ . Der topologische Rand stimmt also mit der Vereinigung der Pflasterkette überein, die aus den nicht entarteten und nicht verklebten Seitenflächen von  $\tilde{\Omega}$  besteht.
2. Die Menge  $\Omega \setminus \tilde{\Omega}$  ist Teilmenge der Pflasterkette, die aus den entarteten und verklebten Seitenflächen besteht.
3. Die Seitenparametrisierung  $\varphi_{ij}$  ist so gewählt, dass das Normalenfeld auf den zulässigen Teilmengen  $M_i$  nach außen orientiert ist.

**Beispiele** Die Kugel  $K_R(0) = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 \leq R^2\}$  ist ein Gaußsches Gebiet. Ebenso ist jedes dreidimensionale Pflaster ohne entartete und ohne verklebte Seitenflächen ein Gaußsches Gebiet.

**Definition 4.53:** Sei  $\Omega \subset \mathbb{R}^3$  ein Gaußsches Gebiet und  $v : \Omega \rightarrow \mathbb{R}^3$  ein  $\mathcal{C}^1$ -Vektorfeld. Wir definieren das Oberflächenintegral von  $v$  auf  $\Omega$  als

$$\int_{\partial\Omega} v \cdot \text{do} = \sum_{k=1}^m \int_{M_k} v \cdot \text{do}. \quad (4.65)$$

Da das Integral über entartete/verklebte Seitenflächen verschwindet, gilt ebenso

$$\int_{\partial\Omega} v \cdot \text{do} = \sum_{i=1}^3 \sum_{j=0}^1 \int_{[0,1]^2} \langle v \circ \varphi_{ij}, \partial_1 \varphi_{ij}(u_1, u_2) \times \partial_2 \varphi_{ij}(u_1, u_2) \rangle d(u_1, u_2) \quad (4.66)$$

nach der bekannten Formel für vektorielle Oberflächenintegrale (4.34). Bemerke, dass aufgrund der Beschränktheit des Gebiets  $\Omega$  die Integrale auf der rechten Seite für stetige Vektorfelder endlich sind.

**Satz 4.54 (Gauß):** Sei  $\Omega \subset \mathbb{R}^3$  ein Gaußsches Gebiet und  $v : \Omega \rightarrow \mathbb{R}^3$ , das auf einer Umgebung von  $\bar{\Omega}$  stetig-differenzierbar ist, so gilt

$$\int_{\Omega} \text{div } v \, d^3x = \int_{\partial\Omega} v \cdot \text{do} = \int_{\partial\Omega} v \cdot n \, \text{do}. \quad (4.67)$$

Hierbei ist  $n$  das nach außen gerichtete Einheitsnormalenfeld auf den Randseiten von  $\Omega$ .

*Beweisskizze.* Wir verwenden erneut die Parametrisierung  $\Phi : [0, 1]^3 \rightarrow \Omega$  von  $\Omega$  und die entsprechenden Parametrisierungen  $\varphi_{ij} = h \circ \Lambda_{ij}$  der Seitenflächen  $M_{ij} = \varphi_{ij}((0, 1)^2)$ .

Wir setzen

$$w := \begin{pmatrix} \langle v \circ \Phi, \partial_2 \Phi \times \partial_3 \Phi \rangle \\ \langle v \circ \Phi, \partial_3 \Phi \times \partial_1 \Phi \rangle \\ \langle v \circ \Phi, \partial_1 \Phi \times \partial_2 \Phi \rangle \end{pmatrix}.$$

Dann lässt sich mittels der Kettenregel zeigen, dass für festes  $(i, j)$

$$\langle v \circ \varphi_{ij}, \partial_1 \varphi_{ij} \times \partial_2 \varphi_{ij} \rangle = \langle w \circ \Lambda_{ij}, \partial_1 \Lambda_{ij} \times \partial_2 \Lambda_{ij} \rangle = \langle w \circ \Lambda, N \rangle$$

gilt, wobei  $N$  das nach außen gerichtete Einheitsnormalenfeld auf dem Rand  $\partial Q_3$  des abgeschlossenen dreidimensionalen Einheitswürfels  $Q_3$  ist. Unter Verwendung des Satzes 4.48 lässt sich sehen, dass

$$\int_{\partial\Omega} v \cdot d\mathbf{o} = \int_{\partial\Omega} v \cdot n \, ds = \int_{\partial Q_3} w \cdot N \, ds = \int_{Q_3} \operatorname{div} w \, d^3x$$

ist. Wir haben also Vektorfelder über Gaußsche Gebiete mithilfe der Parametrisierung zurück auf Vektorfelder über Einheitswürfel geführt. Es lässt sich weiter zeigen, dass

$$\operatorname{div} w = ((\operatorname{div} v) \circ \Phi) \cdot \det J_\Phi.$$

Nun nutzt man noch aus, dass  $\Phi$  ein orientierungstreuer Diffeomorphismus ist, also  $\det J_\Phi > 0$  gilt, sodass eine Anwendung des Transformationssatzes

$$\int_{\partial\Omega} v \cdot n \, ds = \int_{Q_3} \operatorname{div} w \, d^3x = \int_{Q_3} ((\operatorname{div} v) \circ \Phi) \cdot |\det J_\Phi| \, d^3u = \int_{\Omega} \operatorname{div} v \, d^3x$$

liefert. □

### Beispiele

1. Sei  $\Omega := \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 \leq 1\}$  zunächst die Einheitskugel. Wir möchten das Integral

$$\int_{\partial\Omega} x^2 + y + z \, ds$$

mithilfe des Gaußschen Satzes berechnen. Dafür suchen wir ein Vektorfeld  $v$ , sodass

$$v \cdot n = x^2 + y + z$$

und  $n$  das nach außen gerichtete Einheitsnormalenfeld auf der Kugeloberfläche ist. Dieses ist (offensichtlich) durch  $n : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit  $n(x, y, z) := (x, y, z)^T$  gegeben. Daher Forderung an das Vektorfeld wird daher zum Beispiel von  $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit  $v(x, y, z) := (x, 1, 1)^T$  erfüllt. Die Divergenz dieses Vektorfeldes ist konstant 1. Unter Anwendung des Satzes von Gauß finden wir also

$$\int_{\partial\Omega} x^2 + y + z \, ds = \int_{\partial\Omega} v \cdot n \, ds = \int_{\Omega} \operatorname{div} v \, d^3x = \int_{\Omega} 1 \, d^3x = \operatorname{vol}(\Omega) = \frac{4}{3}\pi.$$

2. Sei nun  $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  durch  $v(x, y, z) := (x^2, y^2, z^2)$  gegeben. Diesmal integrieren wir über das Gaußsche Gebiet (den Einheitswürfel)  $Q_3$ . Hierbei ergibt sich

$$\int_{\partial Q_3} v \cdot d\mathbf{o} = \int_{\Omega} \operatorname{div} v \, d^3x = \int_0^1 \int_0^1 \int_0^1 2x + 2y + 2z \, dx \, dy \, dz$$

$$= \int_0^1 \int_0^1 1 + 2y \, dy \, dz = \int_0^1 2 + 2z \, dz = 3.$$

Alternativ können wir die Integrale über die Seitenfläche auch direkt berechnen. Wir zeigen das beispielhaft für die Fläche  $E_{11} = \{(x, y, z) \in \mathbb{R}^3, x = 1, 0 < y, z < 1\}$ . Die Einheitsnormale auf dieser Fläche ist durch  $n = e_1$  gegeben. Damit ist

$$\begin{aligned} \int_{E_{11}} v \cdot d\mathbf{o} &= \int_{E_{11}} v \cdot n \, ds = \int_{E_{11}} v_1 \, ds \\ &= \int_{(0,1)^2} v_1 \circ \Lambda_{11}(u_1, u_2) \cdot \sqrt{\det G(u)} \, d^2u \\ &= \int_0^1 \int_0^1 1 \, d(u_1, u_2) = 1, \end{aligned}$$

wobei  $G = I_3$  hier die zur Parametrisierung  $\Lambda_{11}$  gehörende Grammatrix ist. Für die Seitenfläche  $E_{10}$  erhalten wir ganz analog mit dem entsprechenden nach außen gerichteten Normalenvektor  $n = -e_1$

$$\begin{aligned} \int_{E_{10}} v \cdot d\mathbf{o} &= \int_{E_{10}} v \cdot n \, ds = \int_{E_{10}} -v_1 \, ds \\ &= \int_{(0,1)^2} v_1 \circ \Lambda_{10}(u_1, u_2) \cdot \sqrt{\det G(u)} \, d^2u = 0, \end{aligned}$$

da  $\Lambda_{10}(u_1, u_2) = (0, u_1, u_2)$ . Auf die gleiche Art und Weise erhalten wir

$$\int_{E_{21}} v \cdot d\mathbf{o} = \int_{E_{31}} v \cdot d\mathbf{o} = 1 \quad \text{sowie} \quad \int_{E_{20}} v \cdot d\mathbf{o} = \int_{E_{30}} v \cdot d\mathbf{o} = 0,$$

und somit insgesamt

$$\int_{\partial\Omega} v \cdot d\mathbf{o} = \sum_{i,j} \int_{E_{ij}} v \cdot d\mathbf{o} = 3.$$

# Literatur

- [1] Herbert Amann und Joachim Escher. *Analysis I*. 3. Aufl. Grundstudium Mathematik. Basel, Boston, Berlin: Birkhäuser, 2006. ISBN: 978-3-764-37755-7.
- [2] Herbert Amann und Joachim Escher. *Analysis II*. 2. Aufl. Grundstudium Mathematik. Basel, Boston, Berlin: Birkhäuser, 2006. ISBN: 978-3-764-37105-0.
- [3] Helga Baum. „Analysis 1 und 2. Analysis für Lehramtsstudiengänge“. 2015. URL: <https://www.mathematik.hu-berlin.de/~baum/Skript/Analysis-LA-14-15-Summe.pdf> (besucht am 29.04.2019).
- [4] Otto Forster. *Analysis 1. Differential- und Integralrechnung einer Veränderlichen*. 11., erweiterte Auflage. Grundkurs Mathematik. Wiesbaden: Springer Fachmedien, 2013. ISBN: 978-3-658-00316-6.
- [5] Otto Forster. *Analysis 2. Differentialrechnung im  $\mathbb{R}^n$ , gewöhnliche Differentialgleichungen*. 10., verbesserte Auflage. Grundkurs Mathematik. Wiesbaden: Springer Fachmedien, 2013. ISBN: 978-3-658-02356-0.
- [6] Helmut Fischer und Helmut Kaul. *Mathematik für Physiker. Band 1*. 8. Aufl. Springer Spektrum, 2018. ISBN: 978-3-662-56560-5.
- [7] Klaus Jänich. *Vektoranalysis*. 5. Aufl. Berlin, Heidelberg: Springer, 2005. ISBN: 978-3-540-23741-9.
- [8] Jiří Lebl. *Introduction to Real Analysis. Volume 1*. 5. Aufl. CreateSpace Independent Publishing Platform, 2018. ISBN: 978-1-718-86240-1. URL: <https://www.jirka.org/ra/realanal.pdf> (besucht am 02.05.2019).
- [9] Jiří Lebl. *Introduction to Real Analysis. Volume 2*. 2. Aufl. CreateSpace Independent Publishing Platform, 2018. ISBN: 978-1-718-86548-8. URL: <https://www.jirka.org/ra/realanal2.pdf> (besucht am 02.05.2019).